

12. Reconceiving Animals and Their Evolution: On Some Consequences of New Research on the Modularity of Development and Evolution¹

Richard M. Burian

The last fifteen years have seen an ongoing synthesis among developmental biology, evolutionary biology, and molecular genetics.² A new discipline, “evolutionary developmental biology,” is forcing biologists to reconceive evolutionary history, evolutionary processes, and the ways in which animals are constructed. In this chapter I examine some work bearing on how animals (including humans) are put together. A key claim is that evolution deploys ancient modular processes and tinkers with multi-leveled modular parts, many also ancient, yielding organisms whose relationships, because of modular construction, are far more complex and interesting than had been suspected until very recently. For example, all segmented animals share regulatory machinery that demarcates and specifies identities of body segments and switches on the formation of some organs, e.g., eyes. Some processes, bits of machinery, and parts are recycled and reused repeatedly, both in evolution and in development of a single animal. Such claims require major rethinking of how animals are put together and raise issues about how – and the extent to which – animals are harmoniously integrated. Our understanding of how synchronic and sequential developmental processes are controlled to yield an organism is still far from complete. We do not yet understand the philosophical implications of this new work, but I suggest that they include a limited, non-vitalist form of holism.

¹ This is the submission version of the last chapter of my *Epistemological Essays on Development, Genetics, and Evolution* (New York: Cambridge University Press, 2004). Earlier versions of this chapter were presented at the International Society for History, Philosophy, and Social Studies of Biology and at Ohio University in 2001 and at a Colloquium sponsored by the Virginia Tech Department of Philosophy and Center for Science and Technology Studies in 2002. I am grateful to all three audiences and to numerous colleagues for comments on this and a companion presentation. Colleagues deserving particular thanks are Bill FitzPatrick, Scott Gilbert, Marjorie Grene, Paul Siegel, Günter Wagner, and Lee Zwanziger. Scott Gilbert generously supplied electronic copies of seven of the figures in this chapter. Thanks, Scott!

² The best single resource for the biology that goes into the new discipline is the last chapter of either of the last two editions of Scott Gilbert’s *Developmental Biology* (Gilbert 2000, 2003). Gilbert’s book does an extraordinary job of covering the substantial content of present-day developmental biology. This chapter was built using the sixth edition since the seventh arrived too recently for me to take revisions into account. Some other important books written in the last 15 years include (Arthur 1988, 1997, Carroll, Grenier and Weatherbee 2001, Gerhart and Kirschner 1997, Hall 1992, Hall, 1994, Hall 1999, Hall and Olson 2003, Jablonka and Lamb 1995, Keller 2003, Müller and Newman 2003, Pigliucci 2001, Raff 1996, Schlichting and Pigliucci 1998, Wagner 2001, Wilkins 2002).

Introduction

The task of this chapter is somewhat complicated. I shall devote most of it to some philosophically interesting developments in evolutionary developmental biology and then begin to extract from them some issues that they raise about the nature of animals. To do this, I must cover a great deal of biology rather sketchily. I shall try to provide enough exposition of the technicalities to raise the most important issues calling for clarification and substantive investigation.

To start, I show that animals are put together very differently than we imagined until very recently. Here are a few salient points. The genome as such does not constitute a blueprint for the organism. Rather, a variety of control systems are involved. Some of these are deployed in eggs before fertilization. The control systems are crucial in dividing animals into segments and, in a few known cases, they initiate the formation of particular organs, e.g., brain, eyes, heart, and kidneys. The operation of these control systems are typically staged, coordinated, or synchronized by exogenous, hormonal, and other signals. A great deal of self-assembly is involved; many of the control systems and organs are built in the middle of the process of development, combining resources found in the egg and sperm (including the genome) with signals provided by the environment. Although copies of the control systems are built during development, many of the controls deployed in building animal bodies are very ancient – they are indirect copies of controls that already existed before the lineages that yielded insects and vertebrates diverged, i.e., they are well over 300 million years old. The control systems that regulate gene expression are exquisitely context sensitive and have been recycled, both on an evolutionary scale and within the development of a single organism, to accomplish a great variety of tasks.

We do not yet understand how the actions of these control systems are integrated to yield the very different organisms that fall within these enormous lineages. The organization of the genome has a major role in this integration, a point that I will touch on below. Nonetheless, it is open to question whether there is a central control unit of any sort responsible for the integration of the organism. It is just possible that a series of self-assembly units accomplish much of what goes into making an animal – i.e., that animals are put together as the result of cascades of self-assembly processes that run without much informational input from any sort of central control unit. And within any animal there are parts with conflicting interests that must have reached some sort of stable, but imperfect, equilibrium. If all of this is correct, one of the key problems of developmental biology, potentially tractable to the new molecular and computational methods, will be how integration of the organism is achieved in development. In any event, however this is done, animals have been cobbled together by evolutionary tinkering, in which the reshuffling of miscellaneous pieces and what I call control modules and is at least as important as genetic novelties.

Most of the developments I will discuss are quite recent. The biological findings are rather specialized, and have been obtained largely thanks to the power of new technologies. Accordingly, they are not yet widely known to non-specialists. Accordingly, I will devote most of the chapter to an overview of some findings of interest. Only then can I take on, all too briefly, at some of the larger issues that they raise.

Some Basic Biological Surprises Regarding the Regulation of Development

Our topic is the regulation of development in multicellular animals. For ease of exposition, I will speak only about segmented animals, using examples from insects (primarily the fruit fly, *Drosophila*, whose segmentation is fairly well understood) and mammals (primarily rats and mice, secondarily humans). As will soon become clear, the results discussed here apply to a much wider range of organisms than anyone expected as recently as fifteen years ago. These results could not have been obtained without use of new biological technologies, some of which have been developed in the Human Genome Project. I hope to convince you of the power and importance of a few key findings.

To begin, consider one of the more startling claims made by the teams that mapped the human genome (Consortium 2001, Venter and al. 2001, and other articles in the same journal issues). Although these numbers are still contentious (Hogenesch, et al. 2001, Wright, et al. 2001) using conventional molecular means for delimiting genes, human beings seem to have only about 26,000 to 38,000 genes (probably 30,000 to 35,000), which is approximately 1/3 of the number predicted about a decade ago and less than double the number identified in *Drosophila*. This surprisingly low number of genes is counteracted, however, by a corresponding complexity of the regulatory apparatus that makes far more than 200,000 proteins from these roughly 30,000 genes – more than a 1,200,000 if one includes antibody proteins. *So there is no such thing as a 1:1 relationship between genes and the proteins they produce.*

A second point: depending how one counts and on some uncertainties about the interpretation of the molecular findings, *only about 1.5% to at most 4% of the DNA in our genome codes for proteins.* A significant portion, possibly about half of the rest, is highly repetitive DNA, sometimes called selfish DNA, much of which has no known function. *But something like 40% of the genome has regulatory significance*, i.e., it includes material that in one way or another makes, serves as, modulates, or interacts with, signals affecting the transcription of DNA or the processing of RNA made from that DNA.

Many mechanisms intervene between the transcription of protein-coding sequences to RNA and the eventual production of a protein. For simplicity, I shall concentrate on only two mechanisms, although there are at least two dozen of them. The two I discuss both involve post-transcriptional modification of the sequences transcribed from DNA to RNA. To facilitate the discussion, I use Scott Gilbert's simplified diagram of a real case, illustrating some key steps involved in producing β -globin and hemoglobin (See Fig. 1).³ The stretch of DNA that

encodes the sequence information for the protein begins with a promoter region, required for making an RNA transcript, ultimately processed into the messenger RNA that specifies the sequence of amino acids in the protein. As the first step in the diagram shows, the DNA in question is transcribed to a shorter length of RNA, begin with a so-called leader sequence and ending with a tail composed of a poly-A string (repeated adenines attached to the sugar chain of the RNA). This nuclear RNA, however, contains "introns" – intervening sequences that must be excised, with the remaining RNA spliced

³ This and most of the other figures are taken from (Gilbert 2000). I am grateful for his generous permission to utilize these figures here.

at the excision joints – if the resulting (“mature”) messenger RNA (mRNA) is to pass through a nuclear pore into the cytoplasm.⁴ (Some of the excised material forms “small nuclear RNAs,” some of which play further roles in the processing of RNAs in the nucleus.) The mature mRNA contains a shorter tail, none of the introns, and a signal (the sequence AUG) that marks the end of the leader and the beginning of the RNA that will be translated in making the β -globin protein. This protein, in turn, may go through “post-translational processing” in which it is altered in various ways and combined with some other similarly produced products to make a hemoglobin molecule.

The excision of the introns – non-coding segments of the raw RNA transcript⁵ – joins together previously separated coding segments (called exons) that are now contiguous in the processed RNA transcript. The joined pieces contain the actual signal for the sequence of amino acids produced on ribosomes in the cytoplasm. Virtually all protein-producing genes in animals have from one to dozens of introns. The gene for β -globin is simpler than average in this respect – it has only two introns. Nearly all the DNA that codes for proteins in plants and animals is not contiguous on the DNA molecule. So, point three, *one reason that DNA sequence does not correspond exactly with the protein product of an animal gene is that the signal for the amino acid sequence of the protein is typically put together out of disjoint pieces of DNA*. Thus, the sequence of DNA nucleotides *within* genes does not correspond neatly with the sequence of amino acids in the proteins that our bodies make. Genes are made of modular pieces, some of which correspond to particular domains performing particular functions in the ultimate protein product – and in evolutionary history many of them have been duplicated and/or shuffled around from one gene to another. So sub-genic modular units in the DNA are important in building proteins and in evolution – pieces of genes are put together in building proteins.

The second mechanism to be discussed is alternative splicing of RNA transcripts (see Fig. 2). Different cells utilize variants of closely related proteins. Some protein domains perform particular functions – e.g., binding the rest of the

protein to some particular sort of cell membrane. Other domains perform different functions – e.g., in the case of an immune molecule, recognizing a particular antigen. If what counts as genes are stretches of DNA that produce raw RNA transcripts like those illustrated in the image of the hemoglobin gene, then many animal genes yield RNA that is spliced together differently in different circumstances or in different cells. Fig. 3 presents a schematic diagram of a real case – the rat α -tropomyosin gene, which is regulated so that it yields different

⁴ One of the many hypotheses about the evolutionary reason for this constraint on getting the RNA through the nuclear pore is that the processing involved helps protect against RNAs made by retroviruses from getting out into the cytoplasm and doing damage to cells and the organisms that they have infected.

⁵ Rarely, introns actually contain coding material that belongs to another gene. Thus, strictly speaking, they need not be non-coding, but if they contain coding segments, those segments are part of the code for an entirely different protein.

proteins in different cellular contexts by means of alternative splicing. So, the very same gene can yield quite different proteins, often by putting alternative modular pieces performing particular functions together. This process often yields closely related proteins – e.g., a soluble immune protein that recognizes a particular antigen and a membrane-bound protein that recognizes the same antigen. In the course of evolution, exons can be duplicated or shuffled so that the signal coding for a domain performing a particular function can be spliced into an entirely different protein.⁶

About 40% of human genes are known to be alternatively spliced. Alternative splicing must be highly controlled so that the “right” protein ends up, by and large, being produced in the right place at the right time. However, our understanding of the system of controls that do this job is still quite poor. Suffice to say that those controls are extremely complex. They often involve several layers of double and triple assurance, inhibitors that block misexpression, inhibitors of those inhibitors to allow expression, and so on. These devices affect whether, when and where the product is made, which of the alternative products is made, the rate at which the product is synthesized, and its total quantity. In short, gene expression is in general tightly controlled – and the agents that do the controlling typically are not genes, but gene products that are assembled in the relevant cells or tissues into control devices – devices that are sensitive to external signals. When those controls break down, all kinds of problems arise. One result may be cancer – which happens when some part of the control system malfunctions so that manufacture of one or more particular products and/or one or more cell types is not shut down when it normally would be. (You may have noticed that the α -tropomyosin gene makes products that are components of hepatomas (liver cancers) and of brain tissue – so that gene ought not be defined simply as a gene for striated muscle or as a cancer gene.) The control systems are quite baroque, as is clear from the fact that some genes produce an enormous number of distinct proteins. Evelyn Keller, for instance, cites a study showing that one particular gene may in fact yield 576 splicing variants (Black 1998, cited by, Keller 2000).

As the α -tropomyosin case illustrates, the products of a single coding region are often quite distinct proteins that are used in quite different contexts – such as brain, liver, and skeletal muscle. So regulatory systems control what products a gene yields and which alternative products are made in which circumstances. The success or failure of those systems in controlling what is made, in what circumstances, by particular coding regions (“genes” as counted in the Human Genome Project papers) is crucial in building organisms and required for the well-being of those organisms.

Indeed, a consensus has already emerged that the unexpectedly low increase in the number of genes in mammals (including humans) in comparison with much simpler organisms means that differences in the genes, as such, cannot account for our biological differences from other animals. Rather, the differences between mammals and insects, or among primates, rest primarily on the regulation of gene expression and on what happens in the construction of the organism after the first genes are expressed in the embryo. Differences in the regulatory apparatus that determine which genes are used, when and

⁶ Thus another evolutionary advantage obtained from the modular exon-intron structure is that it makes it possible to move domains independently of one another, combining segments of proteins that serve distinct functions without having to build novel proteins from scratch.

how, are at least as important as the differences in the coding segments of the genes in question. Thus, in their triumphal article announcing the draft sequence of the human genome, (Venter and al. 2001) argued that the reductionist paradigm that drove molecular biology – and much else in biology – is seriously deficient (on this point see also Laubichler 2000). It is impossible to determine from the exact sequence of nucleotides in DNA the function(s) of many, perhaps most, genes. And the ambiguities caused by alternative splicing mean that the genes in question typically do not have a unique protein-encoding structure and do not encode a unique protein. It follows, I believe, that genes alone do not suffice to provide a blueprint for the construction of an animal. Rather, regulatory systems, integrated in ways that are not yet understood, but not wholly specified by the genes are at the heart of the construction of an organism.

This dramatic departure from genetic reductionism provides one of the major conclusions of this chapter and of the entire book. *Genes alone do not provide a blueprint for the organism.* During the last half of the twentieth century reductionism, specifically genetic reductionism, rode high in biology. *During the next few decades, I believe, biologists will highlight the roles played in constructing organisms by dynamic regulatory systems above the level of the genome.* The result will be a non-vitalist, but much more holistic vision of the organism, one that places the integration of the organism at the focus of attention. In short, our new understanding of the apparatus regulating gene expression has undermined classical genetic determinism. The remainder of this chapter will help make this claim clearer and amplify on some of its ramifications.

On Building an Organism by Tinkering

The next biological topic concerns some findings about how body segments are laid down, how segment identities are determined, and how some organs are generated. I argue for a sharp separation of the processes that regulate development and the products of those processes. There are strong indications that homologous segments can be formed by distinct processes, that homologous processes can yield utterly distinct products, and that homologous processes can yield organs that perform similar functions, but that do not have homologous properties. It will require one more round of expository work to explain and justify these fundamental claims.

The biological findings also support the claim that animals are cobbled together by a relatively small number of processes that are very highly conserved and that draw on relatively standard materials, many of which are found in most animals. As early as 1977, François Jacob suggested that evolution is a tinkerer – a *bricoleur* to use the more descriptive French term (Jacob 1977, 1982; see also Duboule and Wilkins 1998). For almost twenty years most evolutionists treated this as a pretty metaphor, but paid it little heed. By now, however, this image seems to be exactly right. In the course of evolution, animals are put together in miscellaneous ways, by jerrybuilt combinations of ancient processes, coopting whatever components are at hand. Evolution *is* a tinkerer.

Much of the relevant work was done on the fruit fly *Drosophila*. About fifteen years ago new molecular tools made it possible to visualize the products of many of the genes whose products act during development. Thanks to these tools and the immense store of knowledge of *drosophila* genetics, *drosophila* quickly became the best-explored insect in developmental biology and a workhorse for developmental genetics.

I shall explain briefly what I mean by homologous segments and then turn to the regulation of developmental genes. There are two kinds of homology to consider. The first is serial homology, the second homology by descent. Fig. 11-3 (redrawn from Raff and Kaufman 1983) diagrams one well-established evolutionary lineage, from annelids to pterygote insects. It reminds us of a number of points. Homology is a matter of degree. The segments of the first animal shown are clearly serially homologous. They very probably originated from duplication of segments – a phenomenon I'll illustrate shortly. Considered just as segments, all the segments of insects are serially homologous in that they are descended from serially homologous segments. But, since the segments acquired more specific identities in the course of evolution, finer-grained homologies arose. Thus, the order *Insecta* arose in this lineage when the three thoracic segments, each bearing a pair of legs, became distinct from the others and became the only leg-bearing segments. The fact that all insects have only six legs,⁷ all belonging by descent to these segments shows that the limitation to three thoracic leg-bearing segments is, for some reason, deeply entrenched and that this homology is of considerable importance to what an insect is and how it is built. The segments of the last insect in this lineage are all homologous by descent to segments of the most remote ancestor shown. The contrast between thoracic segments and abdominal segments is due to the descent of these insects from an ancestor that had locked in a developmental system that yields only three thoracic segments and restricted leg-making to those segments. Similar things apply, of course to vertebrates, as is shown by the fact that a bat wing is homologous *as a forelimb* to a cow's leg, an orangutan's arm, and a human's arm, but only the latter two are homologous *as arms*.

Let me now tie all this to the new work in development by reference to one of the most famous homeotic mutants of drosophila. The wild type has one pair of wings on its second thoracic segment and small balancing organs called halteres on its third thoracic segment. The mutant, called bithorax (described briefly in chap. 11), is a striking and unusual case of serial homology (see Fig. 4). Two mutations have converted what should be the third thoracic segment of this drosophila into a second thoracic segment. Thus, the fly has an extra second thoracic segment, and, accordingly, an extra pair of wings, no third thoracic segment, and thus no halteres.

At the end of the 19th century, William Bateson coined the term *homeosis* for the phenomenon of having a segment or part with the wrong identity; early in the twentieth century he named a mutation that yielded homeosis a homeotic mutation. So this fly has a homeotic mutation, which we now know took place in a complex of three genes called *ultrabithorax*.

So-called hox genes, discussed briefly in the previous chapter, often produce homeotic mutants. They all possess a homeobox, which encodes a 60 amino acid long homeodomain in the corresponding protein. These protein domains are very highly conserved in *all* animals (and even in yeasts), and they attach to DNA at highly specific sites. Proteins with homeodomains regulate the transcription and expression of genes

⁷ The fact that leg formation is sometimes blocked does not alter the point; only three segments have leg-making apparatus, whether or not the production of legs is blocked in some way.

downstream from those attachment sites, most often by turning on transcription, sometimes by affecting its rate or turning it off. In short, *hox* genes produce proteins that alter the transcription of specific regions of DNA. Those proteins are *signal transduction factors*; they are components of *signal transduction modules*. These modules are composed of interacting proteins (and signals that encode those proteins) that act on stretches of DNA. These modules are crucial elements in the regulatory apparatus that controls development and they are typically triggered by the presence, or the removal, of signaling molecules (sometimes environmental, sometimes made by the organism). The mutations that cause the *bithorax* phenotype, for example, block some downstream genes from being transcribed in the normal manner during development. The products of those downstream genes are required to specify the normal identity of the third thoracic segment.

At least a score of *hox* genes play crucial roles in laying down segment boundaries and establishing segment identities in the early *drosophila* embryo. Fig. 5 shows staining for the products of just two *hox* genes at two very early

stages in the development of the *drosophila* embryo. As you can see, the formation of boundaries is quite dramatic. Fig. 6 shows a particularly elegant relationship of seven of the *hox* genes in *Drosophila*. To a first approximation, the expression of each gene is required not only to help establish segment boundaries, but also to activate the next gene in the series. That product is required to block the expression of the preceding gene, to delimit and help specify the identity of the next segment, and to activate the genes required for the succeeding segment. This apparently explains the fact that the physical position of these genes on the chromosome is highly conserved and corresponds with the geographical position of the segments in the embryo in which they are expressed. The second approximation takes into account the controls that affect the timing of gene expression and the fact that the proteins involved interact with each other and with the machinery for transcribing genes in establishing boundaries, but fortunately we do not need to go into the messy details for present purposes.

Let me connect these findings to my claims about homology at the beginning of this section and in the previous chapter. I claimed that distinct processes can trigger formation of homologous segments. This is true for formation of a second thoracic segment as the *bithorax* mutation shows. (Roughly speaking, formation of the first copy of the second thoracic segment is triggered by expression of the *ultrabithorax* gene and formation of the second copy is triggered by preventing the ultrabithorax protein from triggering expression of the *abdominal a* gene.) The point is even clearer from a comparison of grasshoppers and *drosophila*. Grasshoppers first make their cells, and then, afterward, sequentially establish the segment identities for groups of cells, by means of (among other things) cell-to-cell signaling. *Drosophila* define their body segments before they make any cell walls – when the nuclei of the *drosophila* cells are in a common cytoplasm, in which the boundaries are laid down *before* the cells are formed. So cell-to-

cell signaling, crucial in grasshopper segmentation, plays no role at all in establishing drosophila segments. Thus grasshoppers build segments by a different process than drosophila, even though the same hox genes and gene products determine the homologous segment boundaries. Conclusion: *homologous segments can be formed by distinct processes.*

In Drosophila, eight hox genes are found on two separate regions of chromosome 3, but in some insects and primitive chordates they are found on a single chromosome (see Fig. 7). Sometime in chordate evolution, in the lineage

leading to vertebrates, the chromosome carrying those genes doubled two or three times, eventually yielding four chromosomes, on some of which there were local doublings that resulted in a few extra hox genes. Thus *all* vertebrates have at least four clusters of hox genes (though none is complete). But it is still possible to identify safely which vertebrate hox genes are homologous with which drosophila hox genes. The hox genes are all laid out in the same order as in drosophila on all four vertebrate chromosomes, though with some genes omitted on each chromosome. All these genes are activated at some time in development. In some cases, if one of them is mutated or deleted in the laboratory, another one, usually on one of the other chromosomes, can perform some or all of its functions.

The findings about the functions of the control modules that regulate production of the products of these homologous genes yield a startling conclusion. In comparing the expression of the segment-defining genes we saw in drosophila with the generalized picture of the expression of the homologous genes in vertebrates (Fig. 7), we see that, considered at an appropriate level, strongly homologous control systems are at play here. Without going into further detail, one can see that there is genuine segmentation in vertebrates and that homologous controls determine the segment identities of insects and vertebrates. Yet there is no risk of confusing the segments of a chicken, a giraffe, a mouse, or a human with each other or with those of a fruit fly. The controls that determine segment boundaries and segment identities yield quite different products. In short, *homologous processes can yield utterly distinct products.*

I shall take two more small steps before turning again to the controls for making eyes introduced in the previous chapter. First, a slightly finer grained example reinforces the idea that tinkering with the control systems utilizing these hox genes plays a key role in evolution (see Fig. 8). Part of what allows the formation of the extra cervical vertebrae formed during ontogeny of chickens is the shifting the position at which *Hox6* is expressed further toward the end of the organism. Similarly, to produce specialized thoracic vertebrae like those required

for wings in just the right region of the body the *Hox9/Hox10* expression boundary is shifted backward and the boundary between thoracic and lumbar slightly forward. Although we have not examined how to change the adult structures produced at a given segment and the control system is baroquely complicated in ways I have not described,

we can see that (and glimpse how) altering segment identities facilitates evolutionary change. And an account rather similar to the one I have given for the axis from front to back also applies, though with considerable differences of detail, to the proximal to distal axis in segments and limbs.⁸ This means that the tinkering idea looks plausible there as well. Thus the account I have offered for assigning identities along the major body axis by use of control modules may be extended, with modifications, to cover the assignment of identities on other axes and to many body compartments, and is likely to be central to an account of a great deal of development. Repetition of similar controls at many levels and in novel combinations contributes to the regulation of gene expression, and helps explain how particular gene products are made only in certain biochemical milieux, in certain compartments of the body, and, in some cases, at certain stages of development.

We are ready to take another look at some of the recent work on the switches that turn on the process of making eyes. Within the last decade, various labs have shown that so-called switch genes can turn on the elaborate cascade of events to make certain organs in *Drosophila*. This was first shown for eyes (Gehring 1998, Halder, Callaerts and Gehring 1995) and chapter 11 above and recently for wings (Guss, et al. 2001). The story is very interesting, and it depends on a variety of biotechnological tricks. The combination of tricks allows one to express the gene of interest virtually at pleasure by using a normally irrelevant signal to turn on the gene in a specific patch of tissue to learn what happens when it is expressed in the ‘wrong’ place or at the ‘wrong’ time.

In the laboratory of Walter Gehring in Basel, this was done in 1995 with a gene originally named *eyeless* in *Drosophila* because, when mutated, it yielded a fly without eyes. When the unmutated version of the *eyeless* gene was activated in various tissues, supernumerary eyes were produced (Fig. 5 of chap. 11). But the situation proved to be yet more interesting – and raised a problem that will occupy us for years. The sequence of the *eyeless* gene showed it to be a *hox* gene. So the people in the Gehring lab searched in the sequence databases for apparently homologous *hox* genes and they found them to be ubiquitous in animals. Indeed, the follow-up has shown that the genes *are* homologous (for their extraordinary similarity is due to being derived from a common source) and they play a crucial role in eye formation in almost anything that has eyespots or eyes.⁹ (The homologous gene is also expressed in the anterior parts of the heads of primitive animals that have no eyes, so it probably was not originally involved in making eyes.)

In 1995, the gene had recently been renamed *Pax6*.¹⁰ In an already classic experiment, Gehring’s group employed the mouse gene just as they had employed the

⁸ For example, a homeotic mutation in insects that replaces an antenna with a leg orders the parts of the leg, proximal to distal, correctly.

⁹ There is a clear consensus that the circuitry of the regulatory module has been reasonably well elucidated. For example, Pineda et al. (Pineda, et al. 2000) have shown that the same regulatory circuit, or process module, drawing on homologous genes and gene products, deployed in the same order as in vertebrates and insects, is present in platyhelminths (flat worms), where it is required for the formation of eyes. They argue in considerable molecular detail that the regulatory circuit is evolutionarily conserved and is found throughout triploblastic animals (animals with three-layered embryos, which include flatworms, arthropods, and vertebrates).

¹⁰ As was indicated in the previous chapter, it was originally known as *aniridia* in mice, because, when mutated, it yielded mice whose eyes had no iris, as *small eye* in humans, because that was the phenotype produced when the gene mutated.

eyeless gene in drosophila – that is, they activated it in various tissues of a developing drosophila. Fig. 6 of the previous chapter shows what they got. Expression of the mouse *Pax6* gene in drosophila acted just like expression of *eyeless* (the drosophila *Pax6* gene), switching on the cascade of events that make a drosophila eye. Recently the drosophila *Pax6* gene was used to make ectopic eyes in the frog *Xenopus* (Chow, et al. 1999)¹¹. Arguably, these results show that *homologous processes can trigger the formation of organs that perform similar functions, but (because the structures produced have many features that are not shared and not derived from a common source) are not closely homologous*. Of course, this result has many other consequences and raises many more questions. We will return to some of them in the concluding discussion.

Before concluding, I turn to one last figure to make a final point (see Fig. 9). Gehring's group described the *Pax6* gene as a master control gene for making eyes. I think that description is seriously misleading. The gene itself is not a control device; its expression is very tightly controlled. Here are four arguments for this claim. (1) As this schematic diagram shows, in mice the gene is expressed in at least four different places, one of which is the pancreas. When the gene is expressed in the pancreas, it does not make eyes. Rather, it and its gene product interact with other regulatory genes and proteins to regulate development of pancreatic cells and transcription of the insulin, glucagon, and somatostatin genes of the pancreas (Andersen, et al. 1999, Hussain and Habener 1999, Sander, et al. 1997, reviewed in Gilbert 2003, pp. 116-117).

(2) What switches on the cascade that makes the eye or that produces pancreatic proteins is a group of interacting signal transduction modules hooked into the right context. And the multiple modules required to initiate the different processes are composed not only of genes, but also of their own gene products and a series of additional proteins that must interact correctly, with each other and with the nucleotide sequences that respond to the signals. Thus, neither a gene by itself, nor its products by themselves, are controllers. The genes regulated by *Pax6* will not initiate the eye-making cascade or production of pancreatic proteins unless they are operating within the context of the relevantly related groups of signal transduction modules. The result of signal transduction depends on what signal is transduced and on the cellular and biochemical context into which the signal transducing module is introduced. We have seen that the particular module involving *Pax6* serves to switch on the eye-making cascade only in some cellular and subcellular contexts, not in others, so that expression of *Pax6* does not universally switch on the eye-making apparatus. *Pax6* can be used in a variety of ways. Depending on how its expression is regulated and what context its product is put into when that product is expressed. Recently, many biologists and philosophers have made related claims about genes in general (see, e.g., Griffiths and Neumann-Held 1999, Keller 2000, 2003, Morange 2001, Moss 2001, 2003, Neumann-Held 2001).

¹¹ However, the drosophila gene appears not to produce ectopic eyes in mice. According to Gilbert (Gilbert 2003, 120), this is because the drosophila *Pax6* gene product does not repress the mouse genes that are used to construct the organs that would normally be produced in locations downstream from eyes, whereas the mouse *Pax6* gene product does repress the downstream triggers for formation of “downstream” organs in drosophila.

(3) Even when genes or their direct products do perform sharply defined functions – e.g., switching on the process of making eyes – they may not be the only genes that can perform those functions. Consider the impact of the sorts of gene duplication we saw with the hox genes. In the case of eye making, at least four distinct genes in drosophila are now known whose products can turn on the cascade of events that normally result in an eye.

(4) The *Pax6* gene may be considered singular among the genes that initiate the eye-making cascade in that it is usually expressed before (“upstream of”) the other genes that can initiate that cascade. But this does not make *Pax6* a master controller for two reasons. For one thing, there are cases in which it is not the most upstream gene in the cascade regulating eye-making. A gene named “twin of eyeless” (abbreviated “*Toy*”) turns out to be the initiator of the eye making regulatory process in drosophila (Czerny, et al. 1999). Secondly, the expression of the controls that turn on the eye-making cascade is, in any case, tightly controlled by a higher order system. It takes but a moment of thought to realize that expression of this cascade (and thus of the genes that initiate it) *has* to be rigorously controlled – otherwise there would be organisms with eyes in all sorts of surprising places. What switches on the cascade is an integrated module or group of interacting modules, i.e., numerous interacting transducing signals set into the appropriate contexts, not the *Pax6* gene or its expression.

Thus, there is no determinate answer to what this gene does – or, more generally, what any gene does. What genes do typically depends on when and where they are expressed and how they are spliced, and this depends on the system of controls that activates or represses them and regulates the splicing. Until one knows in what contexts they are expressed, how they are spliced in those contexts, how their products are post-transcriptionally or post-translationally modified, and what processes they enter into in those contexts, one’s account is deficient. In the life of an individual organism, the answers to these questions depend on a great variety of contingencies. These include what other genes are nearby and (since regulatory systems respond to feedback) what biochemical and other modifications have already occurred in response to the environment that the organism and the relevant tissues have already encountered. Thus the regulatory controls are at least partly epigenetic.

These arguments are quite general. They show (what T. H. Morgan and his colleagues already knew about genes in 1915, though not sequences!) that the salient effects by means of which we identify Mendelian genes need not be the most important effects of the sequences we pick out by reference to the associated phenotype. They also show that the precise effects of genes interpreted as sequences cannot be determined by their sequences or even the sequence of the entire genome. To put it strongly: without knowing an enormous amount about the contexts in which a sequence and its products are placed – contexts that vary enormously in ways that cannot be predicted from complete knowledge of the genome – there is in general no determinate answer to the question “What does this sequence of nucleotides do for this organism?”. Thus, if one tries to proceed strictly from the genome up, one cannot predict, in general, what effects sequence-identified genes will have on the organism and how they will affect its fitness. At best we can obtain contextually restricted propensities or ex post facto averages.

Philosophical Coda

To close, I will first summarize the major points for which I have provided a direct argument and then turn to some larger speculations. I believe I have made a persuasive case for the following claims.

1. Signal transduction processes commonly regulate the contexts in which many proteins are produced and the amino acid sequences of the proteins,¹² and (indirectly thereby) what those proteins do. The signal transduction pathways are themselves multipurpose modules that have been recycled through evolutionary time into many different contexts. Since a given sequence of amino acids can take on radically different roles in different tissues and at different stages in ontogeny, the pathway from genotype to phenotype is not determined by nucleotide sequence.
2. Since the pathways from genes to proteins are far more complex than was traditionally expected, the pathway from a gene to the totality of phenotypic traits it causes or affects is even more difficult to analyze. In general, if genes are identified with nucleotide sequences, it is not possible to determine in any straightforward way the totality of effects of a gene on the particular animal within which it occurs.
3. Genes, taken by themselves, are not central agents in any of the processes we considered. Activation and control of transcription depends thoroughly on complex systems for making, activating, and regulating the action of the enzymes that produce RNA transcripts. This complex process is not controlled by the genes that are transcribed, though it is often modulated by signals that are, themselves, sequences of nucleotides or derived from sequences of nucleotides.

Given these and related considerations, no serious version of genetic determinism can be true. The sequence of nucleotides in the genome of an egg, sperm, or fertilized egg is not a blueprint for the organism.

These considerations are closely allied to the view that evolution is a process of tinkering. We can now see that animals are composites in surprising ways. Their ontogeny is controlled by a miscellany of processes that come from different sources and work in different ways. To appreciate the full depth and strength of this point, one has to go far beyond the material covered in this talk. Evolutionary tinkering means that items from different sources have been put together in novel, often awkward ways.

This completes the summary of what I claim to have shown. I close with some broad speculative comments.

The cascades of events that follow on the activation of particular modules exhibit considerable autonomy – witness the formation of an “extra” second thoracic segment or a supernumerary eye. All the processes we have considered are surprisingly independent of each other once they begin – formation of the primary body axis, formation of limb axes, formation of a specific organ (not just the eye – also a wing, a bone, a limb, a heart, a kidney, or a brain). Other things being equal, each of these is carried out to completion, whatever the fate of the others. Similar things are true for process on many scales that we haven’t considered here. Two examples, chosen simply to illustrate that similar claims hold at different levels illustrate the point: construction and assembly of the intracellular

¹² My direct argument turns on the regulation of DNA transcription, alternative splicing, and intra-nuclear processing of RNA. Other controls of post-transcriptional and post-translational processing, not discussed here, only strengthen the argument.

microtubule spindle apparatus required for mitosis and making the branches of nerves in the nervous system. All of the processes I named are under *local* control, which is how they can carry on independently of other processes on the same scale. The signals that limit an organ or that limit limb bud growth, that stop membrane growth, microtubule extension, and cell proliferation all seem to be based on short-range signals or local biochemistry. These considerations suggest that there are a very large number of fairly autonomous processes that are somehow coordinated in making a complex multicellular animal.

The quasi-independence of processes, just suggested, makes good evolutionary sense. In effect, for evolution to occur traits must be capable of varying independently. The modules that control those traits, therefore, must be mutually independent. After all, if forelimbs could not vary independently from hindlimbs, a single lineage could not produce opossums, koalas, and kangaroos. Quasi-autonomy at many levels allows tinkering. But quasi-autonomy of this sort means that there cannot be too close a connection between the controls that regulate one part and the controls that regulate other parts, even closely related parts. Speculatively, I suggest that distant signals – e.g., the signals of circulating hormones – coordinate fairly independent cascades of events that are “locally” controlled by particular modules or module cascades that draw on immediately available resources, including, of course, the nucleotide sequences of the genes inside the relevant cells. The idea here is that coordination of ontogeny is achieved by dependence on, or integration of, a cascade of distant signals that allows the start of a process that depends on distant assembly processes having already been completed.

But this does not require that there be any “master plan” in virtue of which the body is constructed. Rather, the integration of the body is the evolutionary result of quasi-autonomous modules locking in and starting irreversible processes in ways that are coordinated by distant signals. The coordination is under the long-term but rigorous constraints imposed by natural selection. Those organisms that are not as well integrated as their closest competitors are very likely to be culled in the long run. If this is right, coordination of autonomous modules is the key to the integration of the organism, and not a master plan.

References

- Andersen, F. G., R. S. Heller, H. V. Petersen, J. Jensen, O. D. Madsen and P. Serup 1999. "Pax6 and Cdx2/3 form a functional complex on the rat glucagon gene promoter G1-element." *FEBS Letters* 445: 306-310.
- Arthur, W. 1988. *A Theory of the Evolution of Development*. Chichester: Wiley.
- Arthur, W. 1997. *The Origin of Animal Body Plans: A Study in Evolutionary Developmental Biology*. Cambridge: Cambridge University Press.
- Black, D. L. 1998. "Splicing in the inner ear: A familiar tune, but what are the instruments." *Neuron* 20: 165-168.
- Breitbart, R. A., A. Andreadis and B. Nadal-Ginard 1987. "Alternative splicing: A ubiquitous mechanism for the generation of multiple protein isoforms from single genes." *Annual Review of Biochemistry* 56: 481-495.
- Carroll, S. B., J. K. Grenier and S. D. Weatherbee 2001. *From DNA to Diversity: Molecular Genetics and the Evolution of Animal Design*. Malden, MA: Blackwell Science.
- Chow, R. L., C. R. Altmann, R. A. Lang and A. Hemmati-Brivanlou 1999. "Pax6 induces ectopic eyes in a vertebrate." *Development* 126: 4213-4222.
- Consortium, I. G. S. 2001. "Initial sequencing and analysis of the human genome." *Nature* 409: 860-941.
- Czerny, T., G. Halder, U. Kloter, A. Souabni, W. J. Gehring and M. Busslinger 1999. "Twin of eyeless, a second Pax-6 gene of Drosophila, acts upstream of eyeless in the control of eye development." *Molecular Cell* 3: 297-307.
- Dessain, S., C. T. Gross, M. A. Kuziora and W. McGinnis 1992. "Antp-type homeodomains have distinct DNA-binding specificities that correlate with their different regulatory functions in embryos." *EMBO Journal* 11: 991-1002.
- Duboule, D. and A. S. Wilkins 1998. "The evolution of 'bricolage'." *Trends in Genetics* 14, no. 2: 54-59.
- Gehring, W. J. 1998. *Master Control Genes in Development and Evolution: The Homeobox Story*. New Haven: Yale University Press.
- Gerhart, J. and M. Kirschner 1997. *Cell, Embryos, and Evolution: Toward a Cellular and Developmental Understanding of Phenotypic Variation and Evolutionary Adaptability*. Malden, MA: Blackwell Science.
- Gilbert, S. F. 2000. *Developmental Biology*. Sunderland, MA: Sinauer.
- Gilbert, S. F. 2003. *Developmental Biology*. Sunderland, MA: Sinauer.
- Griffiths, P. E. and E. M. Neumann-Held 1999. "The many faces of the gene." *BioScience* 49: 656-662.
- Guss, K. A., C. E. Nelson, A. Hudson, M. E. Kraus and S. B. Carroll 2001. "Control of a genetic regulatory network by a selector gene." *Science* 292: 1164-1167.
- Halder, G., P. Callaerts and W. J. Gehring 1995. "Induction of ectopic eyes by targeted expression of the eyeless gene in Drosophila." *Science* 267, no. 5205: 1788-92.
- Hall, B. K. 1992. *Evolutionary Developmental Biology*. London: Chapman and Hall.
- Hall, B. K., ed. 1994. *Homology: The Hierarchical Basis of Comparative Anatomy*. New York and San Diego: Academic Press.
- Hall, B. K. 1999. *Evolutionary Developmental Biology*. Dordrecht: Kluwer.
- Hall, B. K. and W. M. Olson 2003. *Keywords & Concepts in Evolutionary Developmental Biology*. Cambridge, MA: Harvard University Press.

- Hirth, F. and H. Reichert 1999. "Conserved genetic programs in insect and mammalian brain development." *BioEssays* 21: 677-684.
- Hogenesch, J. B., K. A. Ching, S. Batalov, A. I. Su, J. R. Walker, Z. Y., S. A. Kay, P. G. Schultz and M. P. Cooke 2001. "A comparison of the Celera and Ensembl predicted gene sets reveals little overlap in novel genes." *Cell* 106: 413-415.
- Holland, P. W. H. and J. Garcia-Fernández 1996. "Hox genes and chordate evolution." *Developmental Biology* 173: 382-395.
- Horowitz, D. S. and A. R. Kraianer 1995. "Mechanisms for selecting 5' splice sites in mammalian pre-mRNA splicing." *Trends in Genetics* 10: 100-106.
- Hussain, M. A. and J. Habener, F. 1999. "Glucagon gene transcription activation mediated by synergistic interactions of pax-6 and cdx-2 with the p300 co-activator." *Journal of Biological Chemistry* 274, no. 41: 28950-28957.
- Jablonka, E. and M. J. Lamb 1995. *Epigenetic Inheritance and Evolution: The Lamarckian Dimension*. Oxford and New York: Oxford University Press.
- Jacob, F. 1977. "Evolution and tinkering." *Science* 196: 1161-1166.
- Jacob, F. 1982. *The Possible and the Actual*. Seattle, WA: University of Washington Press.
- Kammandel, B., A. Chowdhury, A. Stoykova, S. Aparicio, S. Brenner and P. Gruss 1998. "Distinct cis-essential modules direct the time-space pattern of *Pax6* gene activity." *Developmental Biology* 205: 79-97.
- Kaufman, T. C., M. A. Seeger and G. Olsen 1990. "Molecular and genetic organization of the Antennapedia gene complex of *Drosophila melanogaster*." *Advances in Genetics* 27: 309-362.
- Keller, E. F. 2000. *The Century of the Gene*. Cambridge, MA: Harvard University Press.
- Keller, E. F. 2003. *Making Sense of Life*. Cambridge, MA: Harvard University Press.
- Laubichler, M. D. 2000. "The organism is dead. Long live the organism!" *Perspectives on Science* 8: 286-315.
- Lawrence, P. A. 1992. *The Making of a Fly: The Genetics of Animal Design*. Oxford: Blackwell Scientific Publications.
- Morange, M. 2001. *The Misunderstood Gene*. Transl. by M. Cobb. Cambridge, MA: Harvard University Press.
- Moss, L. 2001. "Deconstructing the gene and reconstructing molecular developmental systems." In *Cycles of Contingency: Developmental Systems and Evolution*, ed. by S. Oyama, P. E. Griffith and R. D. Gray. Cambridge, MA: MIT Press, 85-97.
- Moss, L. 2003. *What Genes Can't Do*. Cambridge, MA: MIT.
- Müller, G. B. and S. A. Newman, eds. 2003. *Origination of Organismal Form: Beyond the Gene in Developmental and Evolutionary Biology*. Cambridge, MA: MIT Press.
- Neumann-Held, E. M. 2001. "Let's talk about genes: The process molecular gene concept and its context." In *Cycles of Contingency: Developmental Systems and Evolution*, ed. by S. Oyama, P. E. Griffith and R. D. Gray. Cambridge, MA: MIT Press, 69-84.
- Pigliucci, M. 2001. *Phenotypic Plasticity: Beyond Nature and Nurture*. Baltimore, MD: The Johns Hopkins University Press.
- Pineda, D., J. Gonzalez, P. Callaerts, K. Ikey, W. J. Gehring and E. Salo 2000. "Searching for the prototypic eye genetic network: Sine oculis is essential for eye

- regeneration in planarians.” *Proceedings of the National Academy of Sciences, USA* 97: 4525 - 4529.
- Raff, R. A. 1996. *The Shape of Life: Genes, Development, and the Evolution of Animal Form*. Chicago: University of Chicago Press.
- Raff, R. A. and T. C. Kaufman 1983. *Embryos, Genes, and Evolution: The Developmental-Genetic Basis of Evolutionary Change*. New York: Macmillan.
- Sander, M., A. Neubuser, J. Kalamaras, H. C. Ee, G. R. Martin and M. S. German 1997. “Genetic analysis reveals that PAX6 is required for normal transcription of pancreatic hormone genes and islet development.” *Genes and Development* 11: 1662-1673.
- Schlichting, C. D. and M. Pigliucci 1998. *Phenotypic Evolution: A Reaction Norm Perspective*. Sunderland, MA: Sinauer Associates.
- Venter, J. C. and e. al. 2001. “A high-resolution radiation hybrid map of the human genome draft sequence.” *Science* 291: 1204-1351.
- Wagner, G. P., ed. 2001. *The Character Concept in Evolutionary Biology*. San Diego: Academic Press.
- Wilkins, A. S. 2002. *The Evolution of Developmental Pathways*. Sunderland, MS: Sinauer.
- Williams, S. C., C. R. Altmann, R. L. Chow, A. Hemmati-Brivanlou and R. A. Lang 1998. “A highly conserved lens transcriptional control element from the *Pax-6* gene.” *Mechanics of Development* 73: 225-229.
- Wright, F. A., W. J. Lemon, W. D. Zhao, R. Sears, D. Zhuo, J. P. Wang, H. Yang, Y., T. Baer, D. Stredney, J. Spitzner, et al. 2001. “A draft annotation and overview of the human genome.” *Genome Biology* 2, no. 7: 1-18.

Figure 1

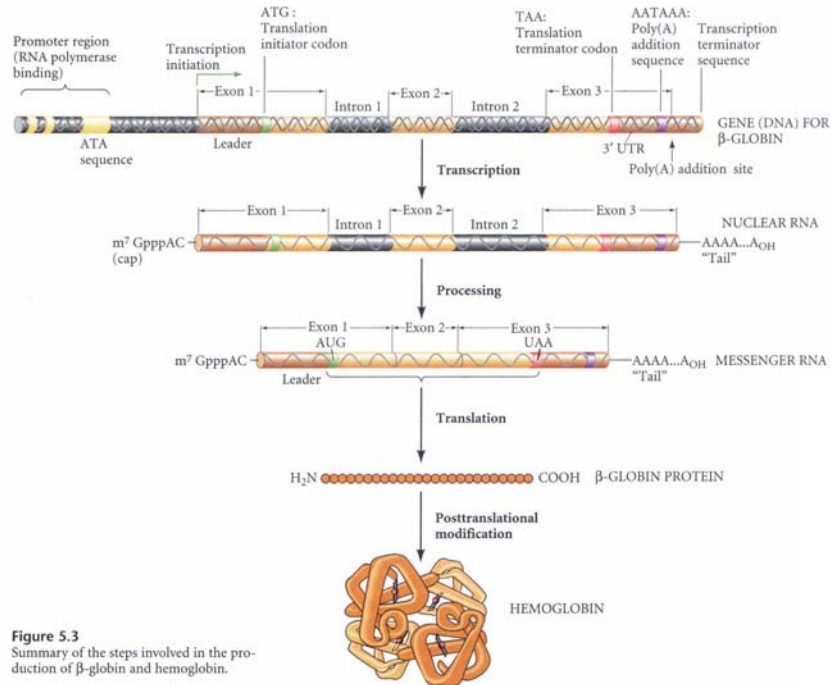


Figure 5.3
Summary of the steps involved in the production of β -globin and hemoglobin.

Fig. 1. A schematic summary of steps involved in producing β -globin and hemoglobin. Top: DNA. Second: Transcription yields nuclear RNA with a cap, exons, introns, and untranslated regions at both ends. Third: removal of introns in the nucleus yields mRNA. Fourth: Translation yields the polypeptide (string of amino acids) of β -globin. Final: post-translational processing may join (and modify) two β -globins and two α -globins, configured with four hemes to yield hemoglobin. Reproduced with permission from (Gilbert 2003, fig. 5.3). I am grateful to Prof. Gilbert for supplying me with electronic copies of his figures.

Figure 2

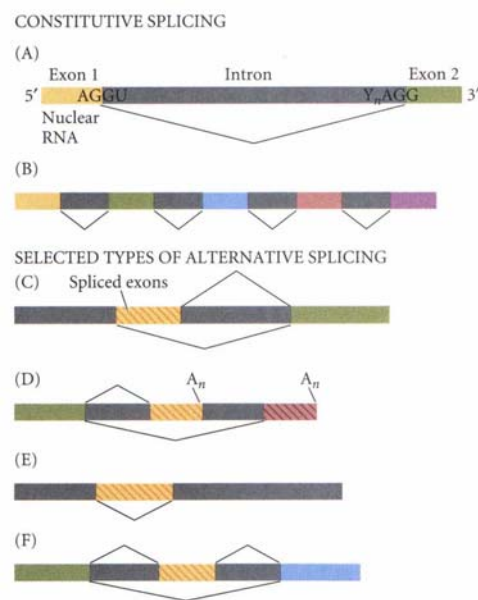


Fig. 2. Schematic Diagram of Alternative Splicing of nuclear RNA. Reproduced with permission from (Gilbert 2003, fig. 5.27), after (Horowitz and Kraianer 1995).

Figure 3

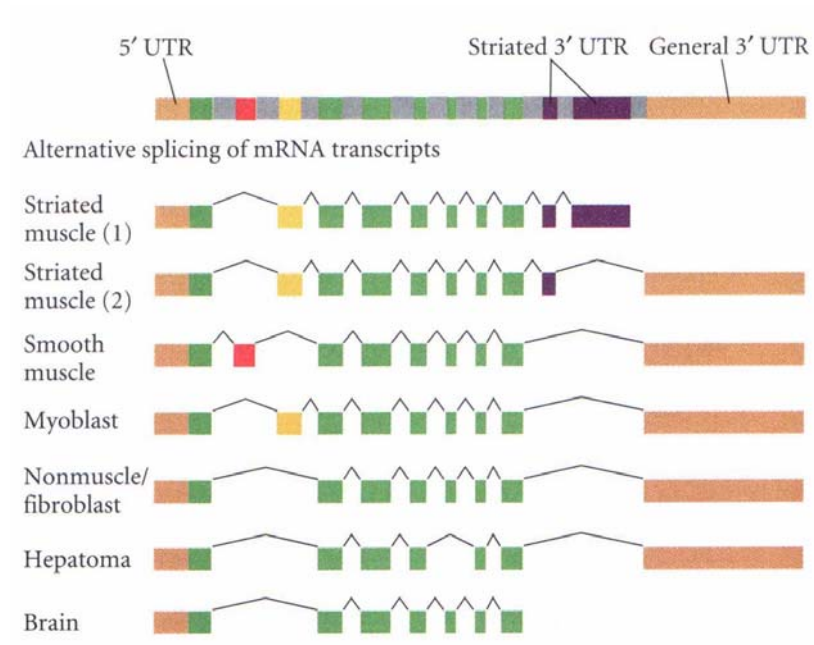


Fig. 3. Alternative RNA splicing to form a family of rat α -tropomyosin proteins. The DNA sequence is represented at the top. Thin lines represent sequences that become introns and are spliced out in forming mature mRNA. Reproduced with permission from (Gilbert 2003, fig. 5.28), after (Breitbart, Andreadis and Nadal-Ginard 1987).

Figure 4

A Four-Winged Fly (*Dipteran??*)



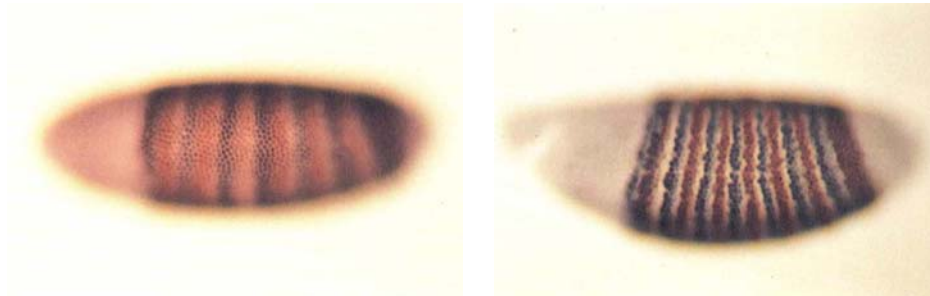
Plate 5.1 Four-winged fly, produced by combining *bithorax* and *postbithorax* mutations [see Figure 5.1].

From Peter A. Lawrence, *The Making of a Fly* (1992)

Fig. 4. Four-winged fly, produced by combining *bithorax* and *postbithorax* mutations. Reproduced with permission from (Lawrence 1992, Plate 5.1).

Figure 5

Expression of two Hox genes in *Drosophila* at Different Times of Development



From Peter A. Lawrence, *The Making of a Fly* (1992)

Fig. 5. The expression of the *fushi tarazu* and *even skipped* gene products at two times in the early development of the *drosophila* embryo. Reproduced with permission from (Lawrence 1992, Plate 4.1).

Figure 6

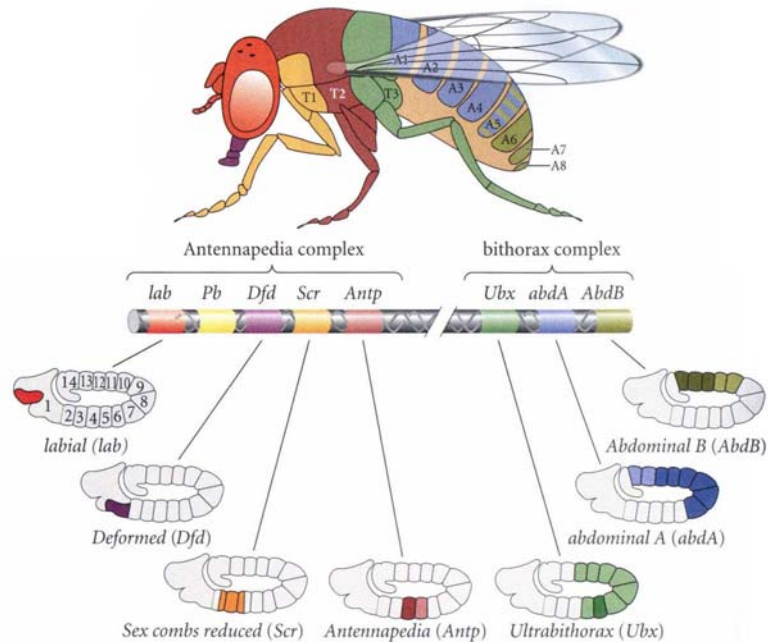


Fig. 6. Homeotic Gene Expression in *Drosophila*. The center bar represents two parts of chromosome three, with two disjoint regions containing the antennapedia and bithorax gene complexes. Above and below are the corresponding regions in which the products of these genes are expressed and in which mutations of the genes yield homeotic changes. Above is the adult, below the blastoderm of the embryo. Reproduced with permission from (Gilbert 2003, fig. 9.28), after (Dessain, et al. 1992) and (Kaufman, Seeger and Olsen 1990).

Figure 7

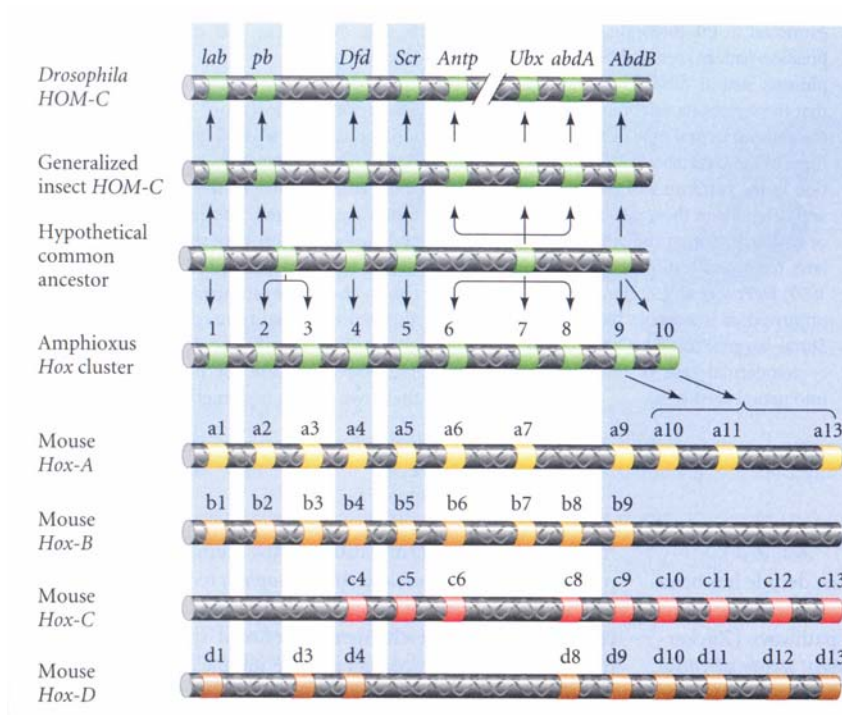


Fig. 7. Postulated ancestry of the homeotic genes from a hypothetical distant ancestor of both insects and mammals. Amphioxus, considered a direct ancestor of vertebrates still has only one chromosome with these homeotic genes. Mammals have four such chromosomes. Reproduced with permission from (Gilbert 2003, fig. 23.11), after (Holland and Garcia-Fernández 1996).

Figure 8

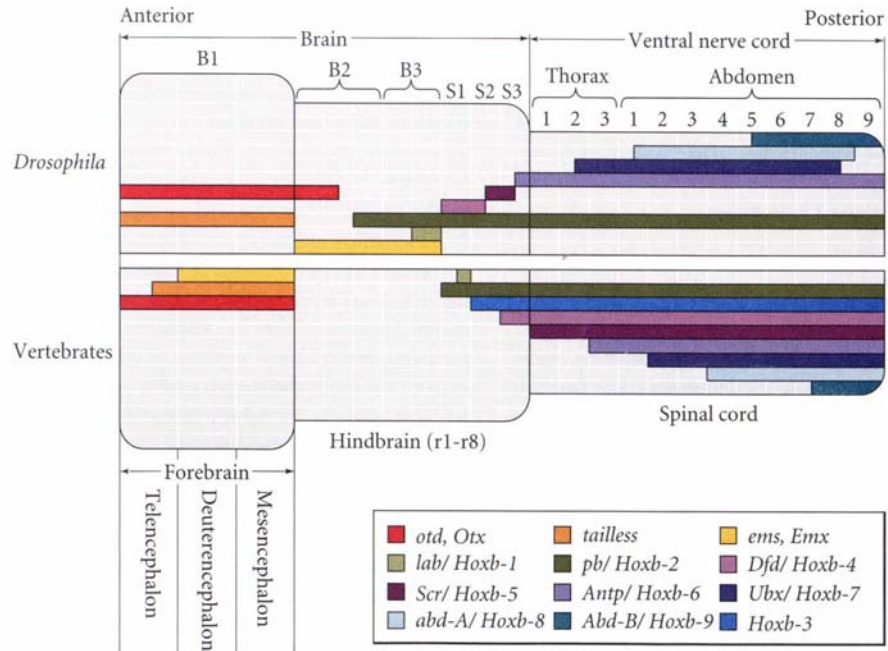


Fig. 8. Expression of regulatory transcription factors in *Drosophila* and in vertebrates along the anterior-posterior axis. Reproduced with permission from (Gilbert 2003, fig. 23.2), after (Hirth and Reichert 1999).

Figure 9

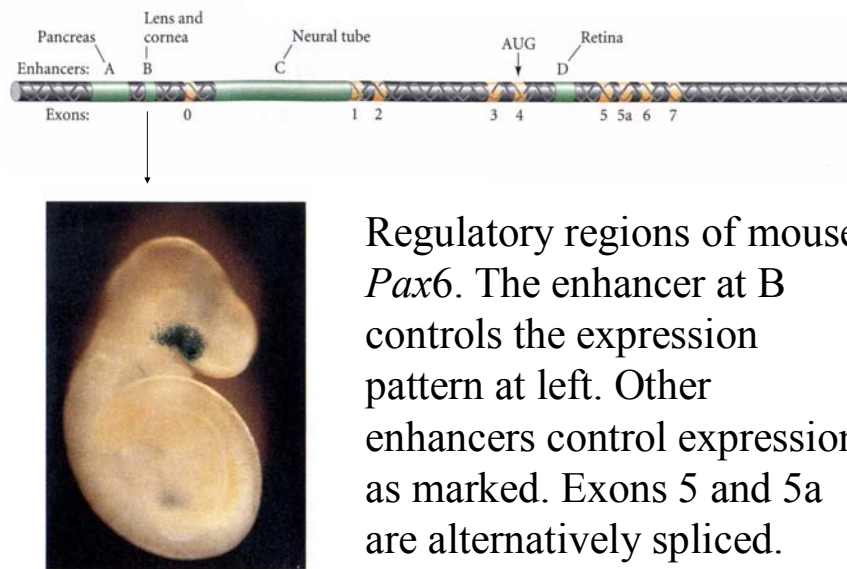


Fig. 9. Expression of a mouse *Pax6* gene when the lens and cornea enhancer is activated in embryogenesis. The gene was altered by biotechnology so that its product could be visualized. The schematic bar shows the mouse *Pax6* gene with its regulatory regions. Reproduced with permission from (Gilbert 2003, fig. 5.7). (After Williams, et al. 1998 and Kammandel, et al. 1998).