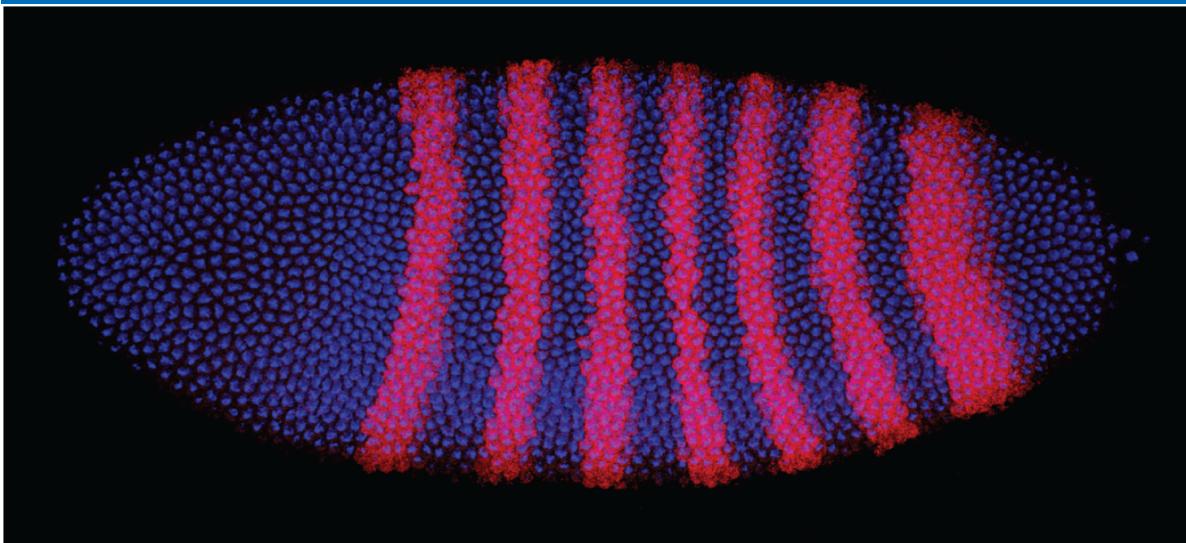


CHAPTER 13

The Genetic Control of Development



Dave Kosman, Ethan Bier, and Bill McGinnis.

Gene expression in a developing fruit-fly embryo. The seven magenta stripes mark the cells expressing the mRNA of a gene encoding a regulatory protein that controls segment number in the *Drosophila* embryo. The spatial regulation of gene expression is central to the control of animal development.

CHAPTER OUTLINE AND LEARNING OBJECTIVES

13.1 THE GENETIC APPROACH TO DEVELOPMENT

LO 13.1 Outline experimental approaches to identify and characterize members of the genetic toolkit for development in different animal phyla.

13.2 THE GENETIC TOOLKIT FOR DROSOPHILA DEVELOPMENT

LO 13.2 Differentiate members of the genetic toolkit for development from other genes.

13.3 DEFINING THE ENTIRE TOOLKIT

LO 13.3 Predict both the phenotypic effects of mutations in toolkit genes based on their expression during development as well as the expression patterns of toolkit genes based on the phenotypic effects of mutations in toolkit genes.

13.4 SPATIAL REGULATION OF GENE EXPRESSION IN DEVELOPMENT

LO 13.4 Infer how spatially and temporally restricted patterns of gene expression are generated during development from analyses of genetic mutations.

13.5 POST-TRANSCRIPTIONAL REGULATION OF GENE EXPRESSION IN DEVELOPMENT

13.6 FROM FLIES TO FINGERS, FEATHERS, AND FLOOR PLATES: THE MANY ROLES OF INDIVIDUAL TOOLKIT GENES

LO 13.5 Summarize the evidence that the genetic toolkit for development is conserved across animal phyla.

13.7 DEVELOPMENT AND DISEASE

CHAPTER OBJECTIVE

In this chapter, we will see that the genetic toolkit that controls the development of complex structures in multicellular organisms comprises a small number of genes that are highly conserved across species. The broad objective for this chapter is to learn how the general principles governing the spatial and temporal regulation of these toolkit genes can be inferred from studies of genetic model organisms such as *Drosophila melanogaster*.

Of all the phenomena in biology, few if any inspire more awe than the formation of a complex animal from a single-celled egg. In this spectacular transformation, unseen forces organize the dividing mass of cells into a form with a distinct head and tail, various appendages, and many organs. The great geneticist Thomas Hunt Morgan was not immune to its aesthetic appeal:

A transparent egg as it develops is one of the most fascinating objects in the world of living beings. The continuous change in form that takes place from hour to hour puzzles us by its very simplicity. The geometric patterns that present themselves at every turn invite mathematical analysis. ... This pageant makes an irresistible appeal to the emotional and artistic sides of our nature.¹

Yet, for all its beauty and fascination, biologists were stumped for many decades concerning how biological form is generated during development. Morgan also said that “if the mystery that surrounds embryology is ever to come within our comprehension, we must ... have recourse to other means than description of the passing show.”

The long drought in embryology lasted well beyond Morgan’s heyday in the 1910s and 1920s, but it was eventually broken by geneticists working very much in the tradition of Morgan-style genetics and with his favorite, most productive genetic model, the fruit fly *Drosophila melanogaster*.

The key catalysts to understanding the making of animal forms were the discoveries of genetic “monsters”—mutant fruit flies with dramatic alterations of body structures ([Figure 13-1](#)). In the early days of *Drosophila* genetics, rare mutants arose spontaneously or as by-products of other experiments with spectacular transformations of body parts. In 1915, Calvin Bridges, then Morgan’s student, isolated a fly having a mutation that caused the tiny hind wings (halteres) of the

fruit fly to resemble the large forewings. He dubbed the mutant *bithorax*. The transformation in *bithorax* mutants is called *homeotic* (Greek *homeos*, meaning same or similar) because one part of the body (the hind wing) is transformed to resemble another (the forewing), as shown in [Figure 13-1b](#). Subsequently, several more homeotic mutants were identified in *Drosophila*, such as the dramatic *Antennapedia* mutant in which legs develop in place of the antennae ([Figure 13-1c](#)).

Homeotic mutants of *Drosophila melanogaster*

(a)



(b)



(c)



Sean Carroll.

FIGURE 13-1 In homeotic mutants, the identity of one body structure has been changed into another. (a) Normal fly with one pair of forewings on the second thoracic segment and one pair of small hind wings on the third thoracic segment. (b) Mutations in the *Ultrabithorax* gene lead to loss of *Ubx* function in the posterior thorax, which causes the development of forewings in place of the hind wings. (c) *Antennapedia* mutant in which the antennae are transformed into legs.

The spectacular effects of homeotic mutants inspired what would become a revolution in embryology, once the tools of molecular biology became available to understand what homeotic genes encoded and how they exerted such enormous influence on the development of entire body parts. Surprisingly, these strange fruit-fly genes turned out to be a passport to the study of the entire animal kingdom, as counterparts to these genes were discovered that played similar roles in

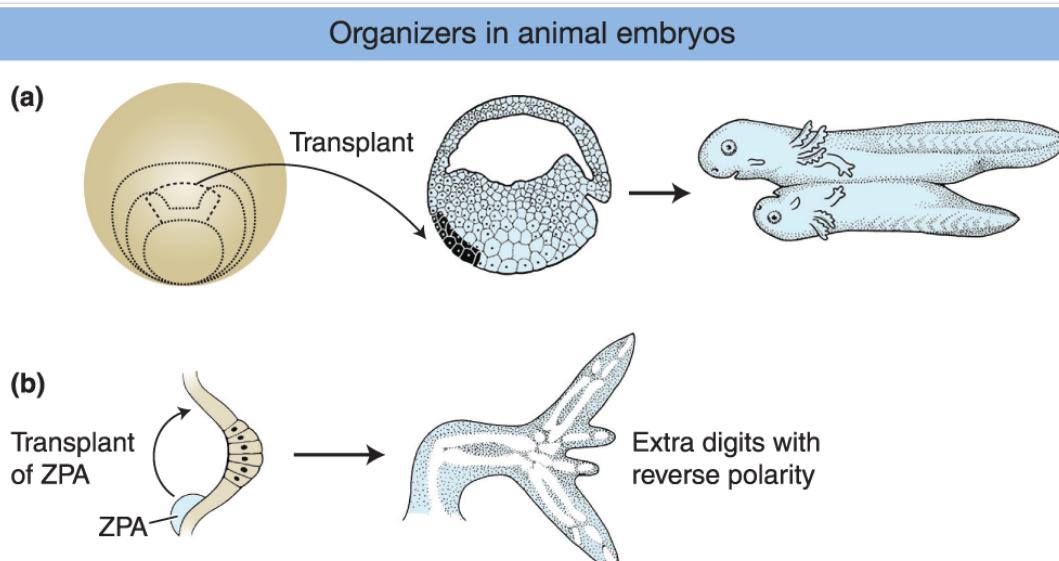
almost all animals. Furthermore, the same regulatory logic that underpins development in animals is also used to control development in plants.

The study of animal and plant development is a very large and still-growing discipline. As such, we do not attempt a comprehensive overview. Rather, in this chapter, we will focus on a few general concepts that illustrate the logic of the genetic control of animal development. We will explore how the information for building complex structures is encoded in the genome. In contrast to the control of gene regulation in single bacterial or eukaryotic cells, the genetic control of body formation and body patterning is fundamentally a matter of gene regulation in three-dimensional *space* and over *time*. Yet we will see that the principles governing the genetic control of development are connected to those already presented in [Chapters 11](#) and [12](#), governing the physiological control of gene expression in bacteria and single-celled eukaryotes.

13.1 THE GENETIC APPROACH TO DEVELOPMENT

LO 13.1 Outline experimental approaches to identify and characterize members of the genetic toolkit for development in different animal phyla.

For many decades, the study of embryonic development largely entailed the physical manipulation of embryos, cells, and tissues. Several key concepts were established about the properties of developing embryos through experiments in which one part of an embryo was transplanted into another part of the embryo. For example, the transplantation of a part of a developing amphibian embryo to another site in a recipient embryo was shown to induce the surrounding tissue to form a second complete body axis ([Figure 13-2a](#)). Similarly, transplantation of the posterior part of a developing chick limb bud to the anterior could induce extra digits, but with reversed polarity with respect to the normal digits ([Figure 13-2b](#)). These transplanted regions of the amphibian embryo and chick limb bud were termed **organizers** because of their remarkable ability to organize the development of surrounding tissues. The cells in the organizers were postulated to produce **morphogens**, molecules that induced various responses in surrounding tissue in a concentration-dependent manner.



Griffiths et al., *Introduction to Genetic Analysis*, 12e, © 2020 W. H. Freeman and Company

FIGURE 13-2 Transplantation experiments played a central role in early embryology and demonstrated the long-range organizing activity of embryonic tissues. (a) The Spemann-Mangold organizer. The dorsal blastopore “lip” of an early amphibian embryo can induce a second embryonic axis and embryo when transplanted to the ventral region of a recipient embryo. (b) In the developing chick limb bud, the zone of polarizing activity (ZPA) organizes pattern along the

anteroposterior axis. Transplantation of the ZPA from a donor limb bud to the anterior position in a recipient limb bud induces extra digits with reverse polarity.

KEY CONCEPT Organizers are groups of cells in an embryo that have the remarkable ability to instruct the development of other cells in an embryo via the production of morphogens, which are molecules that act in a concentration-dependent manner. Cells in close proximity to the organizer are exposed to high concentrations of morphogens and therefore develop into different structures from cells located further from the organizer.

Although these experimental results were spectacular and fascinating, further progress in understanding the nature of organizers and morphogens stalled after their discovery in the first half of the 1900s. It was essentially impossible to isolate the molecules responsible for these activities by using biochemical separation techniques. Embryonic cells make thousands of substances—proteins, glycolipids, hormones, and so forth. A morphogen could be any one of these molecules but would be present in minuscule quantities—one needle in a haystack of cellular products.

The long impasse in defining embryology in molecular terms was broken by genetic approaches—mainly the systematic isolation of mutants with discrete defects in development and the subsequent characterization and study of the gene products that they encoded. The genetic approach to studying development presented many advantages over alternative, biochemical strategies. First, the geneticist need not make any assumptions about the number or nature of molecules required for a process. Second, the (limited) quantity of a gene product is no impediment: all genes can be mutated regardless of the amount of product made by a gene. And, third, the genetic approach can uncover phenomena for which there is no biochemical or other bioassay.

MODEL ORGANISM

Drosophila melanogaster

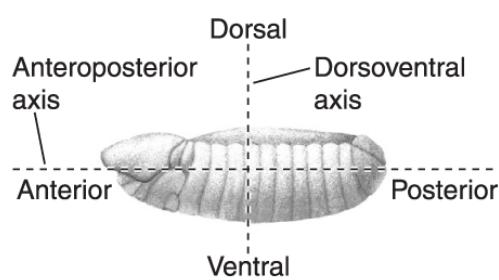
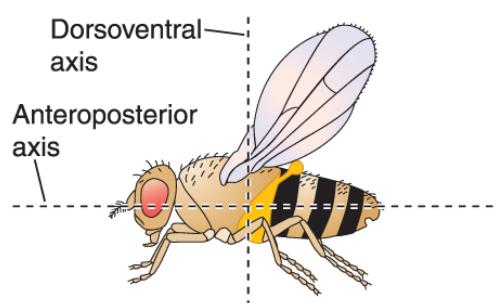
Mutational Analysis of Early *Drosophila* Development

The initial insights into the genetic control of pattern formation emerged from studies of the fruit fly *Drosophila melanogaster*. *Drosophila* development has proved to be a gold mine to researchers because developmental problems can be approached by the use of genetic and molecular techniques simultaneously.

The *Drosophila* embryo has been especially important in understanding the formation of the basic animal body plan. One important reason is that an abnormality in the body plan of a mutant is easily identified in the larval

exoskeleton in the *Drosophila* embryo. The larval exoskeleton is a noncellular structure, made of a polysaccharide polymer called chitin that is produced as a secretion of the epidermal cells of the embryo. Each structure of the exoskeleton is formed from epidermal cells or cells immediately underlying that structure. With its intricate pattern of hairs, indentations, and other structures, the exoskeleton provides numerous landmarks to serve as indicators of the fates assigned to the many epidermal cells (see [Figure 13-13](#)). In particular, there are many distinct anatomical structures along the anteroposterior (A–P) and dorsoventral (D–V) axes (see the figure below). Furthermore, because all the nutrients necessary to develop to the larval stage are prepackaged in the egg, mutant embryos in which the A–P or D–V cell fates are drastically altered can nonetheless develop to the end of embryogenesis and produce a mutant larva in about 1 day (see the figure on the next page). The exoskeleton of such a mutant larva mirrors the mutant fates assigned to subsets of the epidermal cells and can thus identify genes worthy of detailed analysis.

The development of the *Drosophila* adult body pattern takes a little more than a week (see the figure on the next page). Small populations of cells set aside during embryogenesis proliferate during three larval stages (instars) and differentiate in the pupal stage into adult structures. These set-aside cells include the *imaginal disks*, which are disk-shaped regions that give rise to specific appendages and tissues in each segment as the leg, wing, eye, and antennal disks. Imaginal disks are easy to remove for analysis of gene expression (see [Figure 13-7](#)).

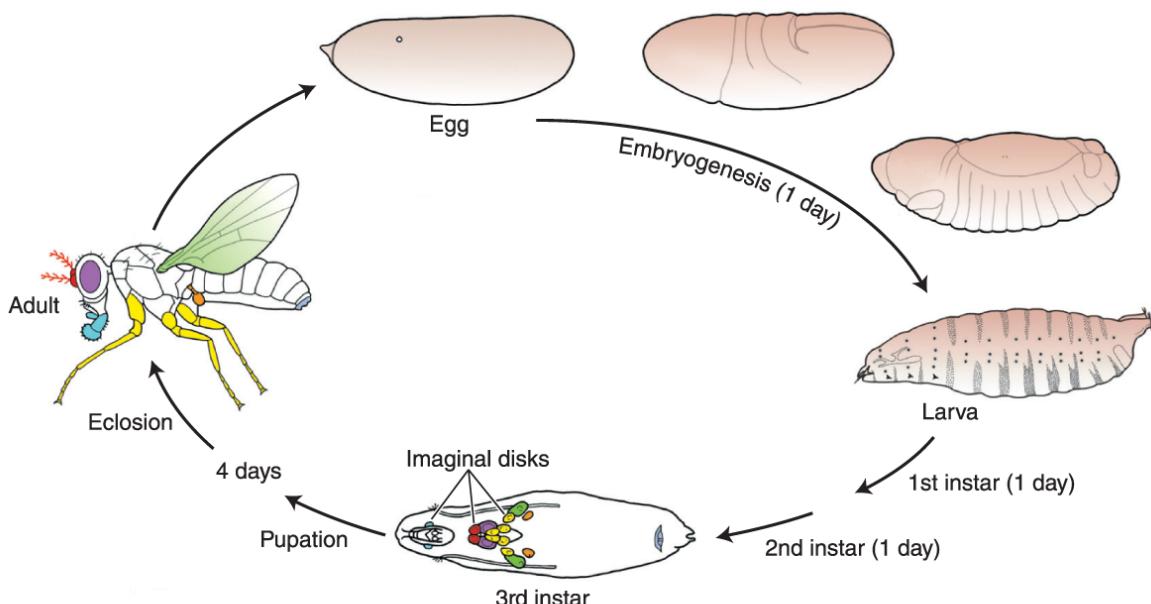


Griffiths et al., *Introduction to Genetic Analysis*, 12e, © 2020 W. H. Freeman and Company

The relationship between adult and embryonic body axes. Note that most images of *Drosophila* embryos in this chapter are oriented so that anterior is to the left, and dorsal is at the top.

Once a mutant with an effect on the *Drosophila* body plan has been identified, the underlying gene can be cloned and characterized at the molecular level with ease. The analysis of the cloned genes often provides valuable information on the function of the protein product—usually by identifying close relatives in amino acid sequence of the encoded polypeptide through comparisons with all the protein sequences stored in public databases. In addition, one can investigate the spatial and temporal patterns of expression of (1) an mRNA. by using

histochemically tagged single-stranded DNA sequences complementary to the mRNA to perform RNA *in situ* hybridization, or (2) a protein, by using histochemically tagged antibodies that bind specifically to that protein (see [Figure 13-5](#)).



Griffiths et al., *Introduction to Genetic Analysis*, 12e, © 2020 W. H. Freeman and Company

Overview of *Drosophila* development. The larva forms in 1 day and then undergoes several stages of growth during which the imaginal disks and other precursors of adult structures proliferate. These structures differentiate during pupation, and the adult fly hatches (eclosion) and begins the cycle again.

St:

ANIMATED ART  **Sapling Plus**

Drosophila embryonic development

Using Knowledge from One Model Organism to Fast-Track Developmental Gene Discovery in Others

With the discovery of genes that regulate development within the *Drosophila* genome, similarities among the DNA sequences of these genes could be exploited in treasure hunts for other members of the gene family. These hunts depend on DNA base-pair complementarity. For this purpose, DNA hybridizations were carried out under *moderate stringency conditions*, in which there could be some mismatch of bases between the hybridizing strands without disrupting the proper hydrogen bonding of nearby base pairs. Some of these treasure hunts were carried out in the *Drosophila* genome itself, in looking for more family members. Others searched for similar genes in other animals, by means of *zoo blots* (Southern blots of restriction-enzyme-digested DNA from different animals), by using radioactive *Drosophila* DNA as the probe (see [Chapter 10](#)). This approach led to the discovery of homologous gene sequences in many different animals, including humans and mice. Now homologous genes are typically identified by computational searches of genome sequences (see [Chapter 14](#)).

From the genetic viewpoint, there are four key questions concerning the number, identity, and function of genes taking part in development:

1. Which genes are important in development?
2. Where in the developing organism and at what times are these genes active?
3. How is the expression of developmental genes regulated?
4. Through what molecular mechanisms do gene products affect development?

To address these questions, strategies had to be devised to identify, catalog, and analyze genes that control development. One of the first considerations in the genetic analysis of animal development was which animal to study. Of the millions of living species, which offered the most promise? The fruit fly *Drosophila melanogaster* emerged as the leading genetic model of animal development because its ease of rearing, rapid life cycle, cytogenetics, and decades of classical genetic analysis (including the isolation of many very dramatic mutants) provided important experimental advantages (see [the Model Organism box on *Drosophila melanogaster*](#) above). The nematode worm *Caenorhabditis elegans* also presented many attractive features, most particularly its simple construction and well-studied cell lineages (see [the Model Organism box on *Caenorhabditis elegans* on page 451](#)). Among vertebrates, the development of targeted gene disruption techniques opened up the laboratory mouse *Mus musculus* to more systematic genetic study, and the zebrafish *Danio rerio* has recently become a favorite model owing to the transparency of the embryo and to advances in its genetic study. Among plants, *Arabidopsis thaliana* has played a similar role as *Drosophila* in illuminating fundamental mechanisms in plant development. More information about the most common model organisms can be found in “A Brief Guide to Model Organisms” at the end of this book.

Through systematic and targeted genetic analysis, as well as comparative genomic studies, much of the [genetic toolkit](#)—the set of genes that control the development of the bodies, body parts, and cell types of several different animal species—has been defined. We will first focus on the genetic toolkit of *Drosophila melanogaster* because its identification was a source of major insights into the genetic control of development; its discovery catalyzed the identification of the genetic toolkit of other animals, including humans.

KEY CONCEPT Genetic model organisms, particularly *Drosophila melanogaster*, have played a key role in the identification of the genetic toolkit for development. Remarkably, many of the toolkit genes discovered in model organisms play fundamental roles in human development and disease.

13.2 THE GENETIC TOOLKIT FOR DROSOPHILA DEVELOPMENT

LO 13.1 Outline experimental approaches to identify and characterize members of the genetic toolkit for development in different animal phyla.

LO 13.2 Differentiate members of the genetic toolkit for development from other genes.

Animal genomes typically contain about 13,000 to 22,000 genes. Many of these genes encode proteins that function in essential processes in all cells of the body (for example, in cellular metabolism or the biosynthesis of macromolecules). Such genes are often referred to as **housekeeping genes**. Other genes encode proteins that carry out the specialized tasks of various organ systems, tissues, and cells of the body such as the globin proteins in oxygen transport or antibody proteins that mediate immunity. Here, we are interested in a different set of genes, those concerned with the building of organs and tissues and the specification of cell types—the genetic toolkit for development that determines the overall body plan and the number, identity, and pattern of body parts.

Toolkit genes of the fruit fly have generally been identified through the monstrosities or catastrophes that arise when they are mutated. Toolkit-gene mutations from two sources have yielded most of our knowledge. The first source consists of spontaneous mutations that arise in laboratory populations, such as those found in the Morgan lab. The second source comprises mutations induced at random by treatment with mutagens (such as chemicals or radiation) that greatly increase the frequency of damaged genes throughout the genome. Elegant refinements of the latter approach have made possible systematic searches, called **genetic screens**, in which organisms are treated with a mutagen and allowed to reproduce, and then the offspring are examined for visible defects in a phenotype of interest. Such screens have identified many members of the fly's genetic toolkit. The members of this toolkit constitute only a small fraction, perhaps several hundred genes, of the roughly 14,000 genes in the fly genome.

KEY CONCEPT The genetic toolkit for animal development is composed of a small fraction of all genes. Only a small subset of the entire complement of genes in the genome affect development in discrete ways.

Classification of genes by developmental function

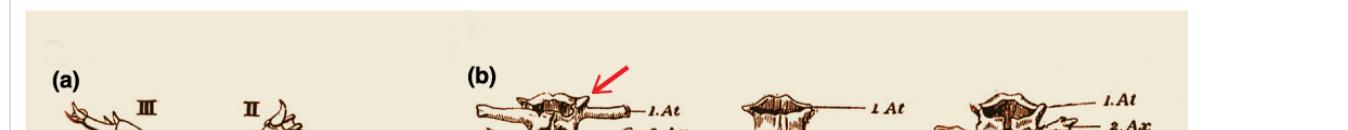
One of the first tasks following the execution of a genetic screen for mutations is to sort out those of interest. Many mutations are lethal when hemi- or homozygous because cells cannot survive without products affected by these mutations. The more interesting mutations are those that cause some discrete defect in either the embryonic or the adult body pattern, or both. It has proved useful to group the genes affected by mutations into several categories based on the nature of their mutant phenotypes. Many toolkit genes can be classified according to their function in controlling the identity of body parts (for example, of different segments or appendages), the formation of body parts (for example, of organs or appendages), the number of body parts, the formation of cell types, and the organization of the primary body axes (the anteroposterior, or A–P, and dorsoventral, or D–V, axes; see [the Model Organism Box on page 430](#)).

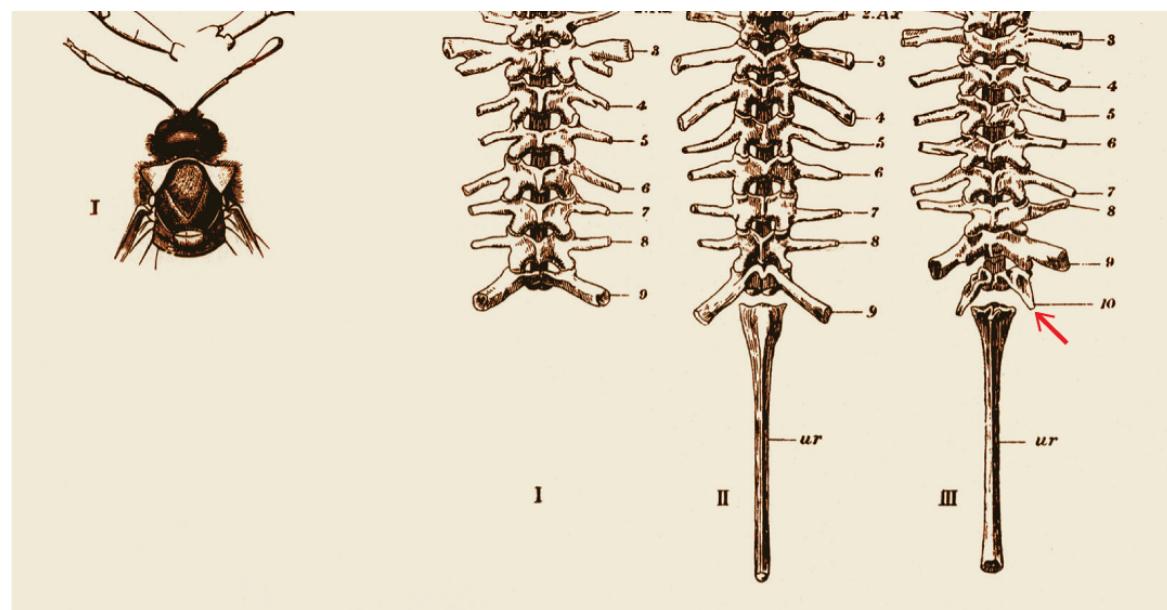
We will begin our inventory of the *Drosophila* toolkit by examining the genes that control the identity of segments and appendages. We do so for both historical and conceptual purposes. The genes controlling segmental and appendage identity were among the very first toolkit genes identified. Subsequent discoveries about their nature were sources of profound insights into not just how their products work, but also the content and workings of the toolkits of most animals. Furthermore, their spectacular mutant phenotypes indicate that they are among the most globally acting genes that affect animal form.

Homeotic genes and segmental identity

Among the most fascinating abnormalities to be described in animals are those in which one normal body part is replaced by another. Such [homeotic transformations](#) have been observed in many species in nature, including sawflies in which a leg forms in place of an antenna and frogs in which a thoracic vertebra forms in place of a cervical vertebra ([Figure 13-3](#)). Whereas only one member of a bilateral pair of structures is commonly altered in many naturally occurring variants, both members of a bilateral pair of structures are altered in homeotic mutants of fruit flies (see [Figure 13-1](#)). In the former case, the alteration is not heritable, but homeotic mutants breed true from generation to generation.

In a homeotic transformation, one body part is replaced by another





Griffiths et al., *Introduction to Genetic Analysis*, 12e, © 2020 W. H. Freeman and Company
From W. Bateson, Material for the Study of Variation. Macmillan, 1894.

FIGURE 13-3 A late-nineteenth-century drawing from one of the first studies of homeotic transformations in nature. (a) Homeosis in a sawfly, with the left antenna transformed into a leg. (b) Homeosis in a frog. The middle specimen is normal. The specimen on the left has extra structures growing out of the top of the vertebral column, transforming a cervical vertebra into a thoracic vertebra (red arrow). The specimen on the right has an extra set of vertebrae (red arrow).

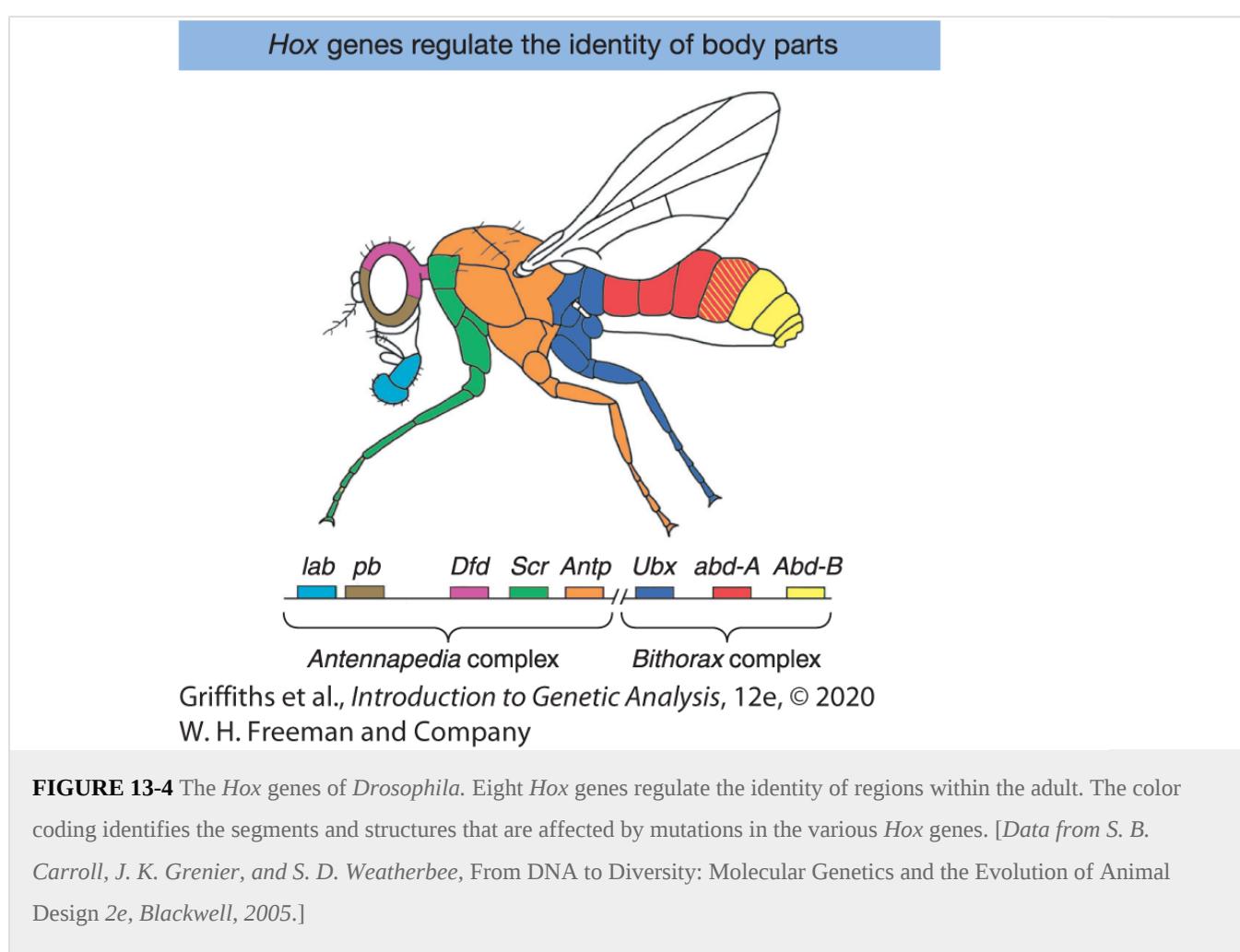
The scientific fascination with homeotic mutants stems from three properties. First, it is amazing that a single gene mutation can alter a developmental pathway so dramatically. Second, it is striking that the structure formed in the mutant is a well-developed likeness of another body part. And, third, it is important to note that homeotic mutations transform the identity of serially reiterated structures. Insect and many animal bodies are made of repeating parts of similar structure, like building blocks, arranged in a series. The forewings and hind wings, the segments, and the antennae, legs, and mouthparts of insects are sets of serially reiterated body parts. Homeotic mutations transform identities within these sets.

A mutation may cause a loss of homeotic gene function where the gene normally acts, or it may cause a gain of homeotic function where the homeotic gene does not normally act. For example, the *Ultrabithorax (Ubx)* gene acts in the developing hind wing to promote hind-wing development and to repress forewing development. Loss-of-function mutations in *Ubx* transform the hind wing into a forewing (see [Figure 13-1](#)). Dominant gain-of-function mutations in *Ubx* transform the forewing into a hind wing. Similarly, the antenna-to-leg transformations of *Antennapedia (Antp)* mutants are caused by the dominant gain of *Antp* function in the antenna (see [Figure 13-1](#)). In addition to these transformations in appendage identity, homeotic mutations can transform segment identity, causing one body segment of the adult or larva to resemble another.

Although homeotic genes were first identified through spontaneous mutations affecting adult flies, they are required throughout most of a fly's development. Systematic searches for homeotic genes have led to the identification of eight loci, now referred to as **Hox genes**, that affect the identity of segments and their associated appendages in *Drosophila*. Generally, the complete loss of any *Hox*-gene function is lethal in early development. The dominant mutations that transform adults are viable in heterozygotes because the wild-type allele provides normal gene function to the developing animal.

Organization and expression of *Hox* genes

A most intriguing feature of *Hox* genes is that they are clustered together in two **gene complexes** that are located on the third chromosome of *Drosophila*. The *Bithorax* complex contains three *Hox* genes, and the *Antennapedia* complex contains five *Hox* genes. Moreover, the order of the genes in the complexes and on the chromosome corresponds to the order of body regions, from head to tail, that are influenced by each *Hox* gene (**Figure 13-4**).



The relation between the structure of the *Hox*-gene complexes and the phenotypes of *Hox*-gene mutants was illuminated by the molecular characterization of the genes. Molecular cloning of the sequences encompassing each *Hox* locus provided the means to analyze where in the developing animal each gene is expressed. These spatial aspects of gene expression and gene regulation are crucial to understanding the logic of the genetic control of development. In regard to the *Hox* genes and other toolkit genes, the development of technology that made possible the visualization of gene and protein expression was crucial to understanding the relation among gene organization, gene function, and mutant phenotypes.

Two principal technologies for the visualization of gene expression in embryos or other tissues are (1) the expression of RNA transcripts visualized by *in situ* hybridization and (2) the expression of proteins visualized by immunological methods. Each technology depends on the isolation of cDNA clones representing the mature mRNA transcript and protein ([Figure 13-5](#)).

Methods for visualizing gene expression in developing organisms

In situ hybridization for visualization of mRNA transcripts

Transcribe to obtain single-stranded RNA probe with modified nucleotides, complementary to mRNA sequence.

Incubate embryos with RNA probe (probe hybridizes to mRNA).

Wash away unbound probe.

Add enzyme-conjugated antibody to specific modified nucleotide.

Wash away unbound antibody.

Add enzyme

cDNA clone of a gene

Fixed embryos or dissected tissue

Immunolocalization of protein expression

Express protein in bacteria.

↓
Inject protein into vertebrate host.

↓
Extract antibodies (IgG) to protein.

Incubate embryos with antibody (antibody binds to protein).

Wash away unbound antibody.

Add fluorochrome-conjugated antibody to IgG of host species.

Wash away unbound antibody.



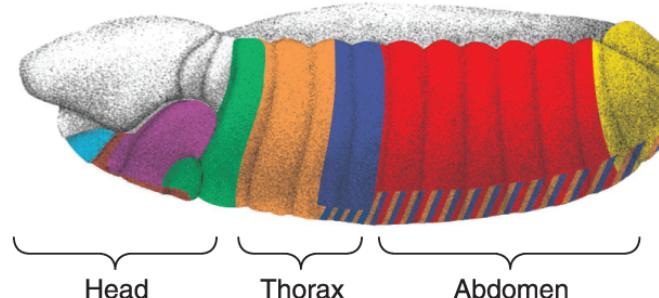
Griffiths et al., *Introduction to Genetic Analysis*, 12e, © 2020 W. H. Freeman and Company

FIGURE 13-5 The two principal technologies for visualizing where a gene is transcribed or where the protein that it encodes is expressed are (*left*) *in situ* hybridization of complementary RNA probe to mRNA and (*right*) immunolocalization of protein expression. The procedures for each method are outlined. Expression patterns may be visualized as the product of an enzymatic reaction or of a chromogenic substrate or with fluorescently labeled compounds.

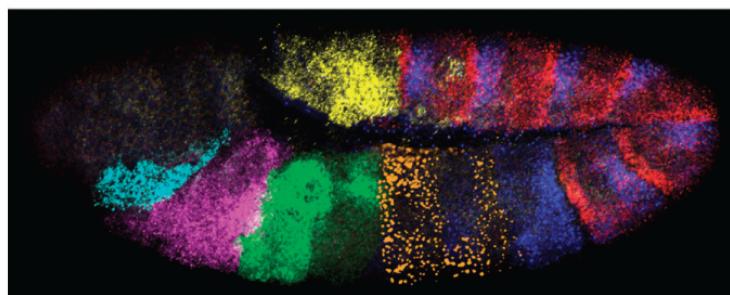
In the developing embryo, the *Hox* genes are expressed in spatially restricted, sometimes overlapping domains within the embryo (Figure 13-6). The genes are also expressed in the larval and pupal tissues that will give rise to the adult body parts.

Hox genes are expressed in spatially restricted domains

(a)



(b)



Griffiths et al., *Introduction to Genetic Analysis*, 12e, © 2020

W. H. Freeman and Company

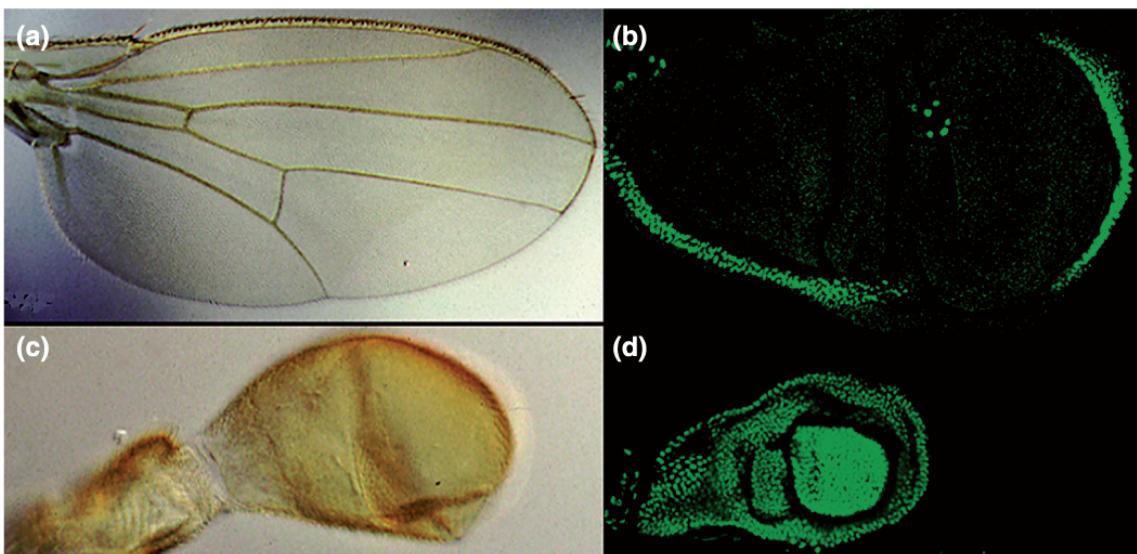
(b) Dave Kosman, Ethan Bier, and Bill McGinnis.

FIGURE 13-6 Expression of *Hox* genes in the *Drosophila* embryo. (a) Schematic representation of *Drosophila* embryo showing regions where eight individual *Hox* genes are expressed. (b) Actual image of the expression of seven *Hox* genes

visualized by *in situ* hybridization. Colors indicate expression of *labial* (turquoise), *Deformed* (lavender), *Sex combs reduced* (green), *Antennapedia* (orange), *Ultrabithorax* (dark blue), *Abdominal-A* (red), and *Abdominal-B* (yellow). The embryo is folded so that the posterior end (yellow) appears near the top center.

The patterns of *Hox*-gene expression (and other toolkit genes) generally correlate with the regions of the animal affected by gene mutations. For example, the dark blue shading in [Figure 13-6](#) indicates where the *Ubx* gene is expressed. This *Hox* gene is expressed in the posterior thoracic and most of the abdominal segments of the embryo. The development of these segments is altered in *Ubx* mutants. *Ubx* is also expressed in the developing hind wing but not in the developing forewing ([Figure 13-7](#)), as one would expect knowing that *Ubx* promotes hind-wing development and represses forewing development in this appendage.

Hox genes are expressed in structures affected by *Hox*-gene mutations



Scott Weatherbee.

FIGURE 13-7 An example of *Hox*-gene expression. (a) The adult forewing of *D. melanogaster*. (b) *Ubx* protein is not expressed in cells of the developing imaginal disk that will form the forewing. Cells enriched in *Hox* proteins are stained green; in this image, the green-stained cells are cells that do *not* form the wing. (c) The adult hind wing (haltere). (d) The *Ubx* protein is expressed at high levels in all cells of the developing hind-wing imaginal disk.

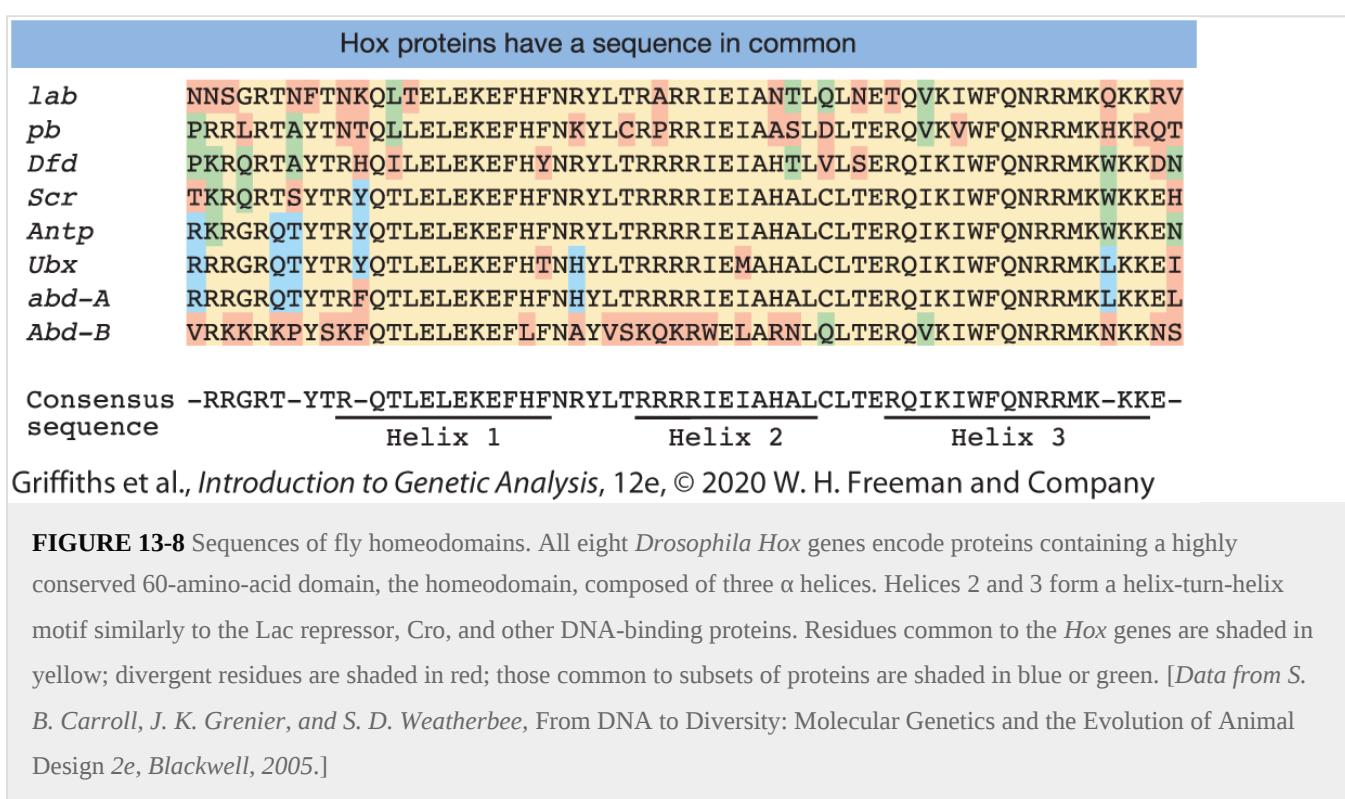
KEY CONCEPT The spatial expression of toolkit genes is usually closely correlated with the regions of the animal affected by gene mutations.

It is crucial to distinguish the role of *Hox* genes in determining the *identity* of a structure from that governing its *formation*. In the absence of function of all *Hox* genes, segments form, but they all have the same identity; limbs also can form, but they have antennal identity; and, similarly, wings

can form, but they have forewing identity. Other genes control the formation of segments, limbs, and wings and will be described later. First, we must understand how *Hox* genes exert their dramatic effects on fly development.

The homeobox

Because *Hox* genes have large effects on the identities of entire segments and other body structures, the nature and function of the proteins that they encode are of special interest. Edward Lewis, a pioneer in the study of homeotic genes, noted early on that the clustering of *Bithorax* complex genes suggested that the multiple loci had arisen by tandem duplication of an ancestral gene. This idea led researchers to search for similarities in the DNA sequences of *Hox* genes. They found that all eight *Hox* genes of the two complexes have a short region of sequence similarity, 180 bp in length. Because this stretch of DNA sequence similarity is present in homeotic genes, it was dubbed the **homeobox**. The homeobox encodes a protein domain, the **homeodomain**, containing 60 amino acids. The amino acid sequence of the homeodomain is very similar among the *Hox* proteins ([Figure 13-8](#)).



Re Although the discovery of a common protein motif in each of the *Hox* proteins was very exciting, further analysis of the structure of the homeodomain revealed that it forms a helix-turn-helix

motif—the structure common to the Lac repressor, the λ repressor, Cro, and the α 2 and α 1 regulatory proteins of the yeast mating-type loci! This similarity suggested immediately (and it was subsequently borne out) that Hox proteins are sequence-specific DNA-binding proteins and that they exert their effects by controlling the expression of genes within developing segments and appendages. Thus, the products of these remarkable genes function through principles that are already familiar from [Chapters 11](#) and [12](#)—by binding to regulatory elements of other genes to activate or repress their expression. We will see that it is also true of many other toolkit genes: a significant fraction of these genes encode transcription factors that control the expression of other genes.

KEY CONCEPT Homeotic transformations result from mutations in *Hox* genes, which are genes that contain a conserved sequence called the homeobox. This sequence encodes a protein domain called the homeodomain, which is similar to the helix-turn-helix motif found in many other transcription factors.

We will examine how Hox proteins and other toolkit proteins orchestrate gene expression in development a little later. First, there is one more huge discovery to describe, which revealed that what we learn from fly *Hox* genes has very general implications for the animal kingdom.

Clusters of *Hox* genes control development in most animals

When the homeobox was discovered in fly *Hox* genes, it raised the question whether this feature was some peculiarity of these bizarre fly genes or was more widely distributed, in other insects or segmented animals, for example. To address this possibility, researchers searched for homeoboxes in the genomes of other insects, as well as earthworms, frogs, cows, and even humans. They found many homeoboxes in each of these animal genomes.

The similarities in the homeobox sequences from different species were astounding. Over the 60 amino acids of the homeodomain, some mouse and fish Hox proteins were identical with the fly sequences at as many as 54 of the 60 positions ([Figure 13-9](#)). In light of the vast evolutionary distances between these animals, more than 500 million years since their last common ancestor, the extent of sequence similarity indicates very strong pressure to maintain the sequence of the homeodomain.

Drosophila and vertebrate Hox proteins show striking similarities

Fly <i>Dfd</i>	PKRQRTAYTRHQILELEKEFHYNRYLTRRRRIEIAHTLVLSERQIKIWFQNRRMKWKKDN	KLPNTKNVR
Amphibian <i>Hox4</i>	TKRSRTAYTRQQVLELEKEFHFNRYLTRRRRIEIAHSLGLTERQIKIWFQNRRMKWKKDN	RLPNTKTRS
Mouse <i>HoxB4</i>	PKRSRTAYTRQQVLELEKEFHYNRYLTRRRVEIAHALCLSERQIKIWFQNRRMKWKKDH	KLPNTKIRS
Human <i>HoxB4</i>	PKRSRTAYTRQQVLELEKEFHYNRYLTRRRVEIAHALCLSERQIKIWFQNRRMKWKKDH	KLPNTKIRS
Chick <i>HoxB4</i>	PKRSRTAYTRQQVLELEKEFHYNRYLTRRRVEIAHSLCLSERQIKIWFQNRRMKWKKDH	KLPNTKIRS
Frog <i>HoxB4</i>	AKRSRTAYTRQQVLELEKEFHYNRYLTRRRVEIAHTLRLSERQIKIWFQNRRMKWKKDH	KLPNTKIKS
Fugu <i>HoxB4</i>	PKRSRTAYTRQQVLELEKEFHYNRYLTRRRVEIAHTLCLSERQIKIWFQNRRMKWKKDH	KLPNTKVRS
Zebrafish <i>HoxB4</i>	AKRSRTAYTRQQVLELEKEFHYNRYLTRRRVEIAHTLRLSERQIKIWFQNRRMKWKKDH	KLPNTKIKS

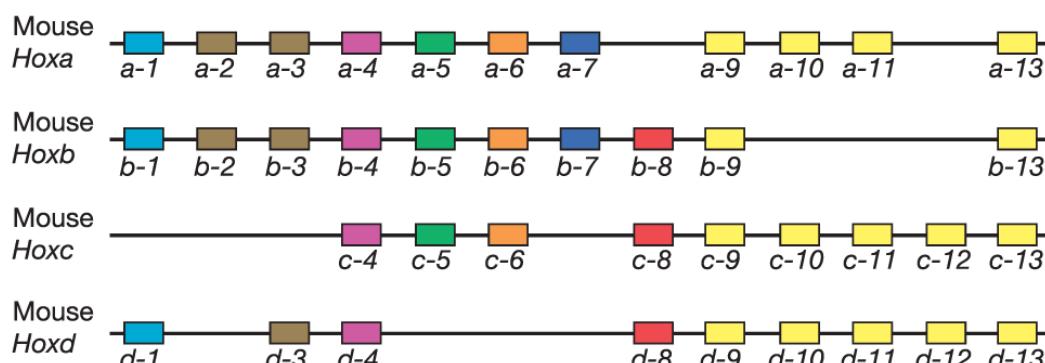
Griffiths et al., *Introduction to Genetic Analysis*, 12e, © 2020 W. H. Freeman and Company

FIGURE 13-9 The sequences of the *Drosophila* Deformed protein homeodomain and of several members of the vertebrate *Hox* group 4 genes are strikingly similar. Residues in common are shaded in yellow; divergent residues are shaded in red; residues common to subsets of proteins are shaded in blue. The very similar C-terminal flanking regions outside of the homeodomain are shaded in green. [Data from S. B. Carroll, J. K. Grenier, and S. D. Weatherbee, From DNA to Diversity: Molecular Genetics and the Evolution of Animal Design 2e, Blackwell, 2005.]

The existence of *Hox* genes with homeoboxes throughout the animal kingdom was entirely unexpected. Why different types of animals would possess the same regulatory genes was not obvious, which is why biologists were further surprised by the results when the organization and expression of *Hox* genes was examined in other animals. In vertebrates, such as the laboratory mouse, the *Hox* genes also are clustered together in four large gene complexes on four different chromosomes. Furthermore, the order of the genes in the mouse *Hox* complexes parallels the order of their most related counterparts in the fly *Hox* complexes, as well as in each of the other mouse *Hox* clusters (**Figure 13-10a**). This correspondence indicates that the *Hox* complexes of insects and vertebrates are related and that some form of *Hox* complex existed in their distant common ancestor. The four *Hox* complexes in the mouse arose by duplications of entire *Hox* complexes (perhaps of entire chromosomes) in vertebrate ancestors.

The order of *Hox* genes parallels the order of body parts in which they are expressed

(a)



(b)





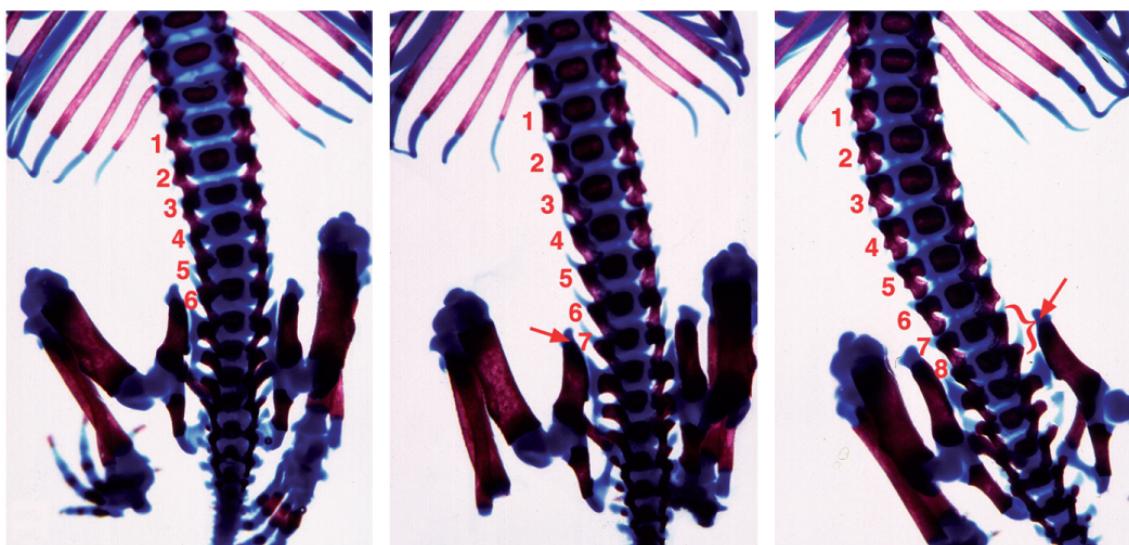
Griffiths et al., *Introduction to Genetic Analysis*, 12e, © 2020 W. H. Freeman and Company
S. B. Carroll, "Homeotic Genes and the Evolution of Arthropods and Chordates," *Nature* 376, 1995, 479–485.

FIGURE 13-10 Like those of the fruit fly, vertebrate *Hox* genes are organized in clusters and expressed along the anteroposterior axis. (a) In the mouse, four complexes of *Hox* genes, comprising 39 genes in all, are present on four different chromosomes. Not every gene is represented in each complex; some have been lost in the course of evolution. (b) The *Hox* genes are expressed in distinct domains along the anteroposterior axis of the mouse embryo. The color shading represents the different groups of genes shown in part a.

Why would such different animals have these sets of genes in common? Their deep, common ancestry indicates that *Hox* genes play some fundamental role in the development of most animals. That role is apparent from analyses of how the *Hox* genes are expressed in different animals. In vertebrate embryos, adjacent *Hox* genes also are expressed in adjacent or partly overlapping domains along the anteroposterior body axis. Furthermore, the order of the *Hox* genes in the complexes corresponds to the head-to-tail order of body regions in which the genes are expressed ([Figure 13-10b](#)).

The *Hox*-gene expression patterns of vertebrates suggested that they also specify the identity of body regions, and subsequent analyses of *Hox*-gene mutants have borne this suggestion out. For example, mutations in the *Hoxa11* and *Hoxd11* genes cause the homeotic transformation of sacral vertebrae to lumbar vertebrae ([Figure 13-11](#)). Thus, as in the fly, the loss or gain of function of *Hox* genes in vertebrates causes transformation of the identity of serially repeated structures. Such results have been obtained in several classes, including mammals, birds, amphibians, and fish. Furthermore, clusters of *Hox* genes have been shown to govern the patterning of other insects and to be deployed in regions along the anteroposterior axis in annelids, molluscs, nematodes, various arthropods, primitive chordates, flatworms, and other animals. Therefore, despite enormous differences in anatomy, the possession of one or more clusters of *Hox* genes that are deployed in regions along the main body axis is a common, fundamental feature of at least all bilateral animals. Indeed, the surprising lessons from the *Hox* genes portended what turned out to be a general trend among toolkit genes; that is, most toolkit genes are common to different animals.

Hox genes regulate the identity of serially repeated structures in vertebrates



(a) Wild type

(b) $Hoxa11^+/Hoxa11^-$;
 $Hoxd11^-/Hoxd11^-$ (c) $Hoxa11^-/Hoxa11^-$;
 $Hoxd11^-/Hoxd11^-$

Photographs courtesy of Dr. Ann Boulet, HHMI, University of Utah; from S. B. Carroll, J. K. Grenier, S. D. Weatherbee, from DNA to Diversity; Molecular Genetics and the Evolution of Animal Design, 2nd ed. Blackwell, 2005.

FIGURE 13-11 The morphologies of different regions of the vertebral column are regulated by *Hox* genes. (a) In the mouse, six lumbar vertebrae (numbers in red) form just anterior to the sacral vertebrae. (b) In mice lacking the function of the posteriorly acting *Hoxd11* gene and possessing one functional copy of the *Hoxa11* gene, seven lumbar vertebrae form and one sacral vertebra is lost. (c) In mice lacking both *Hoxa11* and *Hoxd11* function, eight lumbar vertebrae form and two sacral vertebrae are lost.

KEY CONCEPT Despite great differences in anatomy, a broad array of different animal phyla have many toolkit genes in common.

Now let's take an inventory of the rest of the toolkit to see what other general principles emerge.

13.3 DEFINING THE ENTIRE TOOLKIT

LO 13.1 Outline experimental approaches to identify and characterize members of the genetic toolkit for development in different animal phyla.

LO 13.2 Differentiate members of the genetic toolkit for development from other genes.

LO 13.3 Predict both the phenotypic effects of mutations in toolkit genes based on their expression during development as well as the expression patterns of toolkit genes based on the phenotypic effects of mutations in toolkit genes.

The *Hox* genes are perhaps the best-known members of the toolkit, but they are just a small family in a much larger group of genes required for the development of the proper numbers, shapes, sizes, and kinds of body parts. Little was known about the rest of the toolkit until the late 1970s and early 1980s, when Christiane Nüsslein-Volhard and Eric Wieschaus, working at the European Molecular Biology Laboratory in Heidelberg, Germany, set out to find the genes required for the formation of the segmental organization of the *Drosophila* embryo and larva.

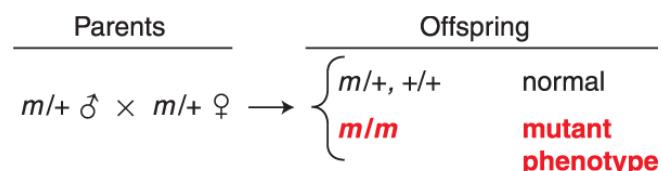
Until their efforts, most work on fly development focused on viable adult phenotypes and not the embryo. Nüsslein-Volhard and Wieschaus realized that the sorts of genes that they were looking for were probably lethal to embryos or larvae in homozygous mutants. So, they came up with a scheme to search for genes that were required in the **zygote** (the product of fertilization; [Figure 13-12](#), bottom). They also developed genetic screens to identify those genes with products that function in the egg, before the zygotic genome is active, and that are required for the proper patterning of the embryo. Genes with products provided by the female to the egg are called **maternal-effect genes**. Mutant phenotypes of strict maternal-effect genes depend only on the genotype of the mother ([Figure 13-12](#), top).

Genetic screens for maternally and zygotically required toolkit genes

MATERNALLY REQUIRED GENES

Parents	Offspring
$m/+ \text{♂} \times m/+ \text{♀}$ → $m/m, m/+, +/+$	all normal
$m/m \text{♂} \times m/+ \text{♀}$ → $m/m, m/+$	all normal
$+/, m/+, \text{ or } m/m \text{♂} \times \text{m/m ♀}$ → $m/+, m/m$	all mutant phenotype

ZYGOTICALLY REQUIRED GENES



Griffiths et al., *Introduction to Genetic Analysis*, 12e, © 2020
W. H. Freeman and Company

FIGURE 13-12 Genetic screens identify whether a gene product functions in the egg or in the zygote. The phenotypes of offspring depend on either (*top*) the maternal genotype for maternal-effect genes or (*bottom*) the offspring (zygotic) genotype for zygotically required genes (*m*, mutant; *+*, wild type).

In these screens, genes were identified that were necessary to make the proper number and pattern of larval segments, to make its three tissue layers (ectoderm, mesoderm, and endoderm), and to pattern the fine details of an animal's anatomy. The power of the genetic screens was their systematic nature. By saturating each of a fly's chromosomes (except the small fourth chromosome) with chemically induced mutations, the researchers were able to identify most genes that were required for the building of the fly. For their pioneering efforts, Nüsslein-Volhard, Wieschaus, and Lewis shared the 1995 Nobel Prize in Physiology or Medicine.



Christiane Nüsslein-Volhard.

Christiane Nüsslein-Volhard and Eric Wieschaus at the European Molecular Biology Laboratory.

The most striking and telling features of the newly identified mutants were that they showed dramatic but discrete defects in embryo organization or patterning. That is, the dead larva was not an amorphous carcass but exhibited specific, often striking patterning defects. The *Drosophila* larval body has various features whose number, position, or pattern can serve as landmarks to diagnose or classify the abnormalities in mutant animals. Each locus could thus be classified according to the body axis that it affected and the pattern of defects caused by mutations. Each class of genes appeared to represent different steps in the progressive refinement of the embryonic body plan—from those that affect large regions of the embryo to those with more limited realms of influence.

KEY CONCEPT Genetic screens are a powerful and unbiased approach to systematically identify genes that affect a biological process, such as embryonic development.

For any toolkit gene, three pieces of information are key toward understanding gene function: (1) the mutant phenotype, (2) the pattern of gene expression, and (3) the nature of the gene product. Extensive study of a few dozen genes has led to a fairly detailed picture of how each body axis is established and subdivided into segments or germ layers.

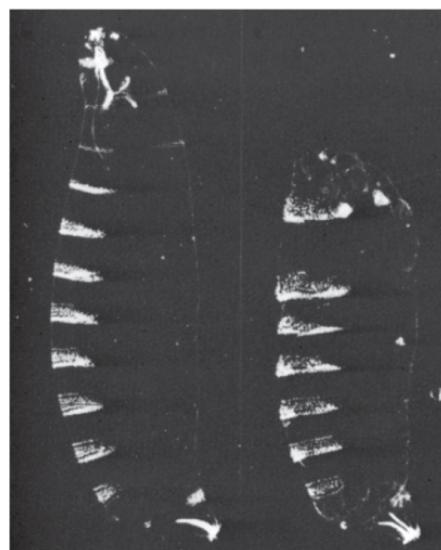
The anteroposterior axis

To illustrate the principles of toolkit genes, we will focus on the anteroposterior body axis in *Drosophila*. However, the same principles apply to the making of the dorsoventral body axis of *Drosophila*, and indeed to the establishment of body axes in both animals and plants. Genetic screens have shown that only a few dozen genes are required for proper organization of the anteroposterior body axis of the fly embryo. The genes are grouped into five classes on the basis of their realm of influence on embryonic pattern.

KEY CONCEPT Toolkit genes can be classified by their roles in development; that is, where and when they function during the development of an organism.

- The first class sets up the anteroposterior axis and consists of the maternal-effect genes. A key member of this class is the *Bicoid* gene. Embryos from *Bicoid* mutant mothers are missing the anterior region of the embryo (**Figure 13-13**), telling us that the gene is required for the development of that region.

Bicoid mutants are missing the anterior region



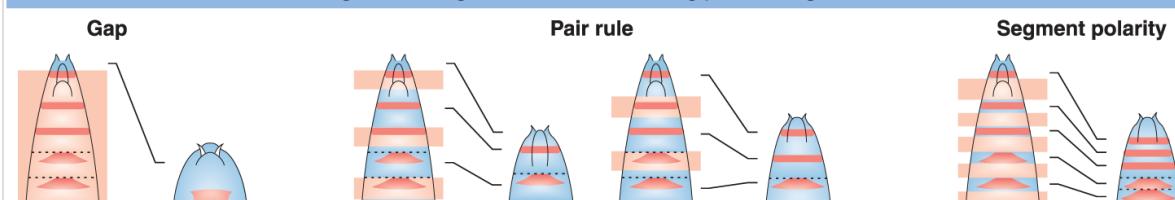
Republished with permission of the American Association for the Advancement of Science, from C.H. Nusslein-Volhard, G. Frohnhofer, and R. Lehmann, "Determination of anteroposterior polarity in *Drosophila*" *Science* Vol. 238, Issue 4834 (1987) 1678, Figure 4. Permission conveyed through Copyright Clearance Center, Inc.

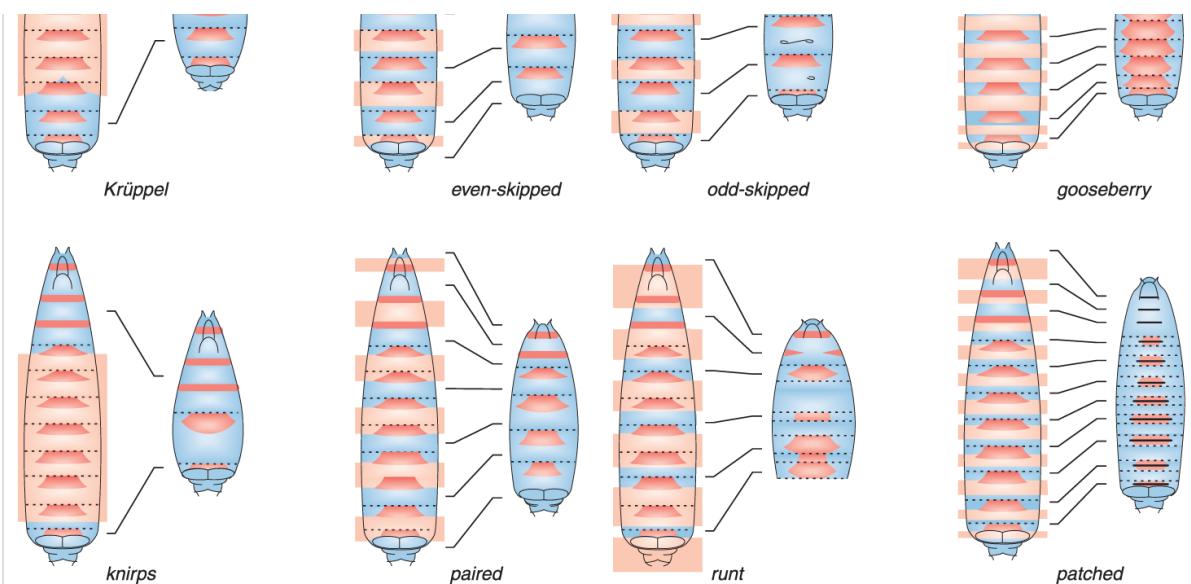
FIGURE 13-13 The *Bicoid* (*bcd*) maternal-effect gene affects the anterior part of the developing larva. These photomicrographs are of *Drosophila* larvae that have been prepared to show their hard exoskeletons. Dense structures, such as the segmental denticle bands, appear white. (Left) A normal larva. (Right) A larva from a homozygous *bcd* mutant female. Head and anterior thoracic structures are missing.

The next three classes are zygotically active genes required for the development of the segments of the embryo.

- The second class contains the **gap genes**. Each of these genes affects the formation of a contiguous block of segments; mutations in gap genes lead to large gaps in segmentation (**Figure 13-14**, left).

Segmentation-gene mutants are missing parts of segments





Griffiths et al., *Introduction to Genetic Analysis*, 12e, © 2020 W. H. Freeman and Company

FIGURE 13-14 Classes of *Drosophila* segmentation-gene mutants. These diagrams depict representative gap, pair-rule, and segment-polarity mutants. The red trapezoids are the dense bands of exoskeleton seen in [Figure 13-13](#). The boundary of each segment is indicated by a dotted line. The left-hand diagram of each pair depicts a wild-type larva, and the right-hand diagram depicts the pattern formed in a given mutant. The shaded light orange regions on the wild-type diagrams indicate the domains of the larva that are missing or affected in the mutant.



ANIMATED ART Sapling Plus

Drosophila embryonic development

- The third class comprises the **pair-rule genes**, which act at a double-segment periodicity. Pair-rule mutants are missing part of each pair of segments, but different pair-rule genes affect different parts of each double segment. For example, the *even-skipped* gene affects one set of segmental boundaries, and the *odd-skipped* gene affects the complementary set of boundaries ([Figure 13-14](#), middle).
- The fourth class consists of the **segment-polarity genes**, which affect patterning within each segment. Mutants of this class display defects in segment polarity and number ([Figure 13-14](#), right).

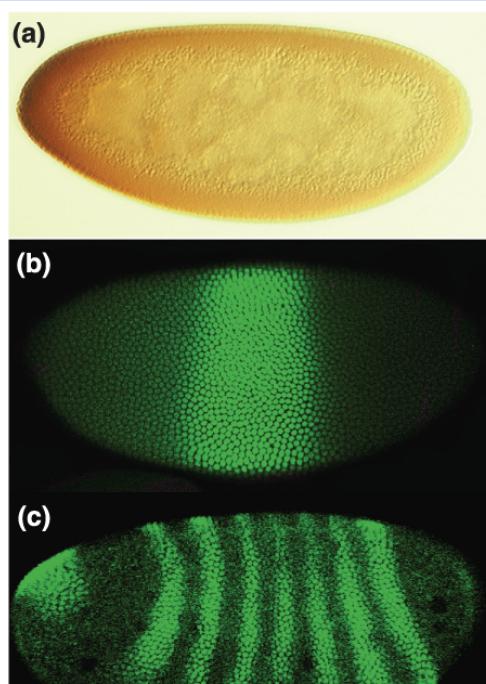
The fifth class of genes determines the fate of each segment.

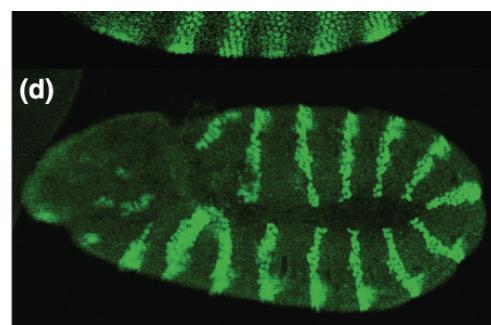
- The fifth class includes the *Hox* genes already discussed; *Hox* mutants do not affect segment number, but they alter the appearance of one or more segments.

Expression of toolkit genes

To understand the relation between genes and mutant phenotype, we must know the timing and location of gene-expression patterns and the molecular nature of the gene products. The patterns of expression of the toolkit genes turn out to vividly correspond to their phenotypes, inasmuch as they are often precisely correlated with the parts of the developing body that are altered in mutants. Each gene is expressed in a region that can be mapped to specific coordinates along either axis of the embryo. For example, the maternal-effect Bicoid protein is expressed in a graded pattern emanating from the anterior pole of the early embryo, the section of the embryo missing in mutants ([Figure 13-15a](#)). Similarly, the gap proteins are expressed in blocks of cells that correspond to the future positions of the segments that are missing in respective gap-gene mutants ([Figure 13-15b](#)). The pair-rule proteins are expressed in striking striped patterns: one transverse stripe is expressed per every 2 segments, in a total of 7 stripes covering the 14 future body segments (the position and periodicity of the stripes correspond to the periodicity of defects in mutant larvae), as shown in [Figure 13-15c](#). Many segment-polarity genes are expressed in stripes of cells within each segment, 14 stripes in all corresponding to 14 body segments ([Figure 13-15d](#)). Note that the domains of gene expression become progressively more refined as development proceeds: genes are expressed first in large regions (gap proteins), then in stripes from three to four cells wide (pair-rule proteins), and then in stripes from one to two cells wide (segment-polarity proteins).

Expression of anteroposterior-axis-patterning proteins





(a) Photomicrographs courtesy of Ruth Lehmann, (b), (c), (d)
Photomicrographs courtesy of James A. Langeland.

FIGURE 13-15 Patterns of toolkit-gene expression correspond to mutant phenotypes. *Drosophila* embryos have been stained with antibodies to the (a) maternally derived Bicoid protein, (b) Krüppel gap protein, (c) Hairy pair-rule protein, and (d) Engrailed segment-polarity protein and visualized by immunoenzymatic (staining is brown) (a) or immunofluorescence (staining is green) (b–d) methods. Each protein is localized to nuclei in regions of the embryo that are affected by mutations in the respective genes.

In addition to what we have learned from the spatial patterns of toolkit-gene expression, the order of toolkit-gene expression over time is logical. The maternal-effect Bicoid protein appears before the zygotic gap proteins, which are expressed before the 7-striped patterns of pair-rule proteins appear, which in turn precede the 14-striped patterns of segment-polarity proteins. The order of gene expression and the progressive refinement of domains within the embryo reveal that the making of the body plan is a step-by-step process, with major subdivisions of the body outlined first and then refined until a fine-grain pattern is established. The order of gene action further suggests that the expression of one set of genes might govern the expression of the succeeding set of genes.

One clue that this progression is indeed the case comes from analyzing the effects of mutations in toolkit genes on the expression of other toolkit genes. For example, in embryos from *Bicoid* mutant mothers, the expression of several gap genes is altered, as well as that of pair-rule and segment-polarity genes. This finding suggests that the Bicoid protein somehow (directly or indirectly) influences the regulation of gap genes.

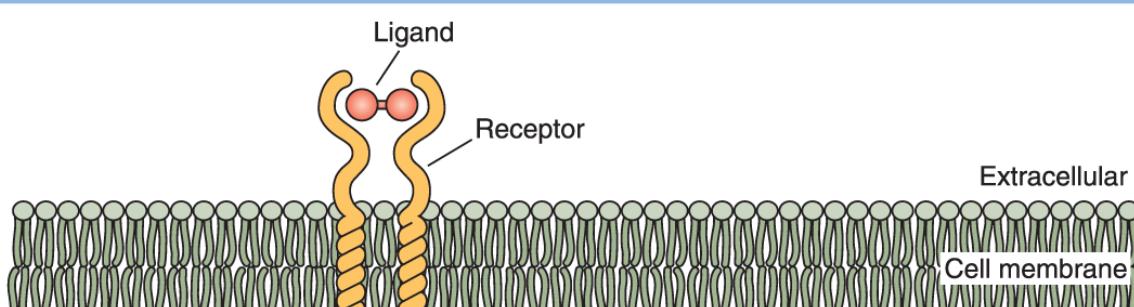
Another clue that the expression of one set of genes might govern the expression of the succeeding set of genes comes from examining the protein products. Inspection of the Bicoid protein sequence reveals that it contains a homeodomain, related to but distinct from those of Hox proteins. Thus, Bicoid has the properties of a DNA-binding transcription factor. Each gap gene also encodes a transcription factor, as does each pair-rule gene, several segment-polarity genes, and, as described earlier, all Hox genes. These transcription factors include representatives of most known families of sequence-specific DNA-binding proteins; so, although there is no

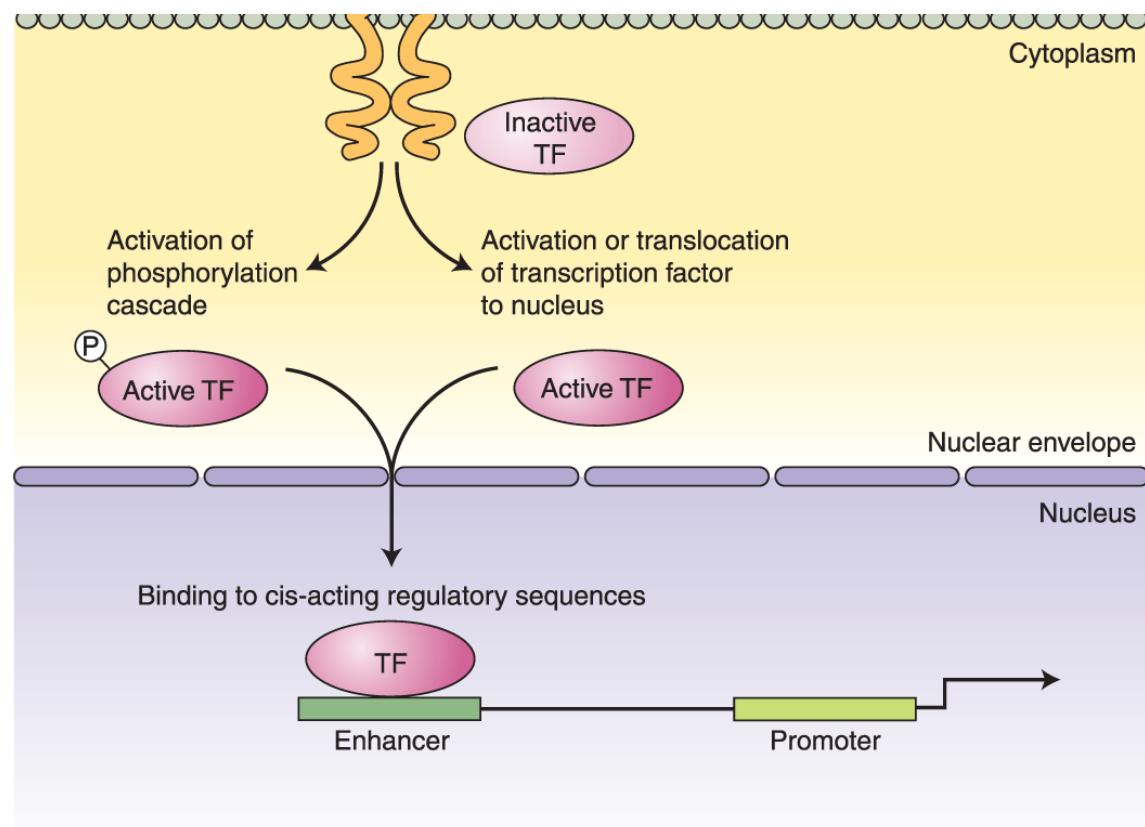
restriction concerning to which family they may belong, many early-acting toolkit proteins are transcription factors. Those that are not transcription factors tend to be components of signaling pathways (**Table 13-1**). These pathways, shown in generic form in **Figure 13-16**, mediate ligand-induced signaling processes between cells, and their output generally leads to gene activation or repression. Thus, most toolkit proteins either directly (as transcription factors) or indirectly (as components of signaling pathways) affect gene regulation.

TABLE 13-1 Examples of *Drosophila* A–P Axis Genes That Contribute to Pattern Formation

Gene symbol	Gene Name	Protein function	Role(s) in early development
<i>bcd</i>	<i>Bicoid</i>	Transcription factor—homeodomain protein	Maternal-effect gene
<i>hb-z</i>	<i>hunchback-zygotic</i>	Transcription factor—zinc-finger protein	Gap gene
<i>Kr</i>	<i>Krüppel</i>	Transcription factor—zinc-finger protein	Gap gene
<i>kni</i>	<i>knirps</i>	Transcription factor—steroid receptor-type protein	Gap gene
<i>eve</i>	<i>even-skipped</i>	Transcription factor—homeodomain protein	Pair-rule gene
<i>ftz</i>	<i>fushi tarazu</i>	Transcription factor—homeodomain protein	Pair-rule gene
<i>opa</i>	<i>odd-paired</i>	Transcription factor—zinc-finger protein	Pair-rule gene
<i>prd</i>	<i>paired</i>	Transcription factor—paired class homeodomain protein	Pair-rule gene
<i>en</i>	<i>engrailed</i>	Transcription factor—homeodomain protein	Segment-polarity gene
<i>wg</i>	<i>wingless</i>	Signaling protein-secreted ligand	Segment-polarity gene
<i>hh</i>	<i>hedgehog</i>	Signaling protein-secreted ligand	Segment-polarity gene
<i>ptc</i>	<i>patched</i>	Signaling protein-transmembrane receptor	Segment-polarity gene
<i>lab</i>	<i>labial</i>	Transcription factor—homeodomain protein	Segment-identity gene
<i>Dfd</i>	<i>Deformed</i>	Transcription factor—homeodomain protein	Segment-identity gene
<i>Antp</i>	<i>Antennapedia</i>	Transcription factor—homeodomain protein	Segment-identity gene
<i>Ubx</i>	<i>Ultrabithorax</i>	Transcription factor—homeodomain protein	Segment-identity gene

A typical signal-transduction pathway





Griffiths et al., *Introduction to Genetic Analysis*, 12e, © 2020 W. H. Freeman and Company

FIGURE 13-16

Most signaling pathways operate through similar logic but have different protein components and signal-transduction mechanisms. Signaling begins when a ligand binds to a membrane-bound receptor, leading to the release or activation of intracellular proteins. Receptor activation often leads to the modification of inactive transcription factors (TF). The modified transcription factors are translocated to the cell nucleus, where they bind to cis-acting regulatory DNA sequences or to DNA-binding proteins and regulate the level of target-gene transcription. [Data from S. B. Carroll, J. K. Grenier, and S. D. Weatherbee, From DNA to Diversity: Molecular Genetics and the Evolution of Animal Design 2e, Blackwell, 2005.]

KEY CONCEPT Most toolkit proteins are transcription factors that regulate the expression of other genes or components of ligand-mediated signal-transduction pathways.

The genetic control of development, then, is fundamentally a matter of gene regulation in space and over time. How does the turning on and off of toolkit genes build animal form? And how is it choreographed during development? To answer these questions, we will examine the interactions among fly toolkit proteins and genes in more detail. The mechanisms that we will see for controlling toolkit-gene expression in the *Drosophila* embryo have emerged as models for the spatial regulation of gene expression in animal and plant development in general.

13.4 SPATIAL REGULATION OF GENE EXPRESSION IN DEVELOPMENT

LO 13.4 Infer how spatially and temporally restricted patterns of gene expression are generated during development from analyses of genetic mutations.

We have seen that toolkit genes are expressed in reference to coordinates in the embryo. But how are the spatial coordinates of the developing embryo conveyed as instructions to genes, to turn them on and off in precise patterns? As described in [Chapters 11](#) and [12](#), the physiological control of gene expression in bacteria and simple eukaryotes is ultimately governed by sequence-specific DNA-binding proteins acting on *cis*-acting regulatory elements (for example, operators and upstream-activation-sequence, or UAS, elements). Similarly, the spatial control of gene expression during development is largely governed by the interaction of transcription factors with *cis*-acting regulatory elements. However, the spatial and temporal control of gene regulation in the development of a three-dimensional multicellular embryo requires the action of more transcription factors on more numerous and more complex *cis*-acting regulatory elements.

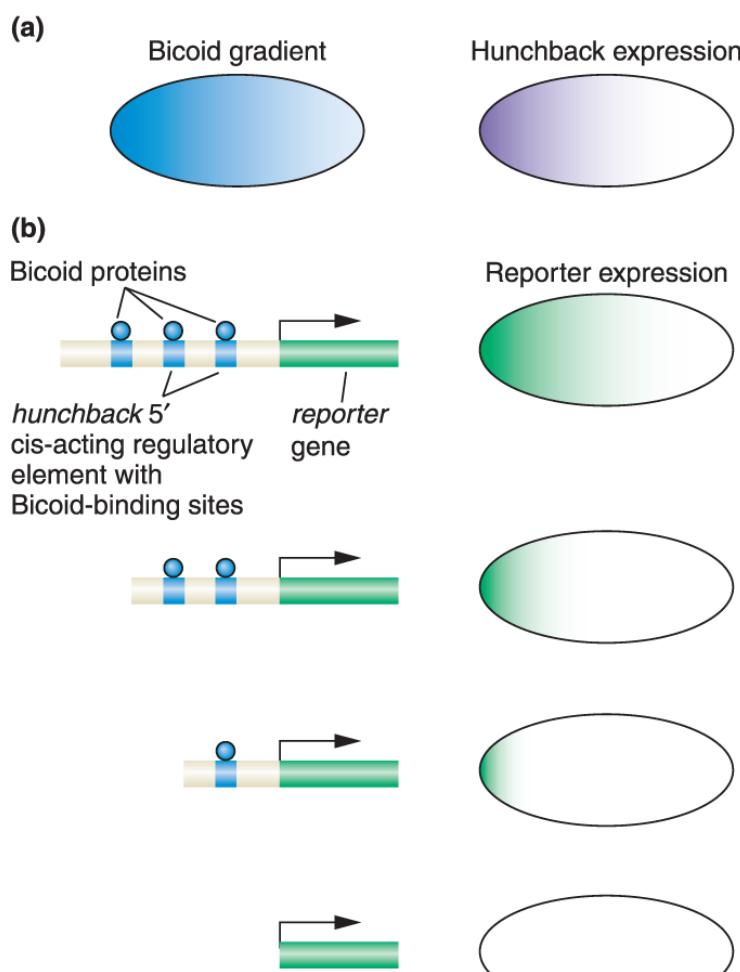
To define a position in an embryo, regulatory information must exist that distinguishes that position from adjacent regions. If we picture a three-dimensional embryo as a globe, then **positional information** must be specified that indicates longitude (location along the anteroposterior axis), latitude (location along the dorsoventral axis), and altitude or depth (position in the germ layers). We will illustrate the general principles of how the positions of gene expression are specified with three examples. These examples should be thought of as just a few snapshots of the vast number of regulatory interactions that govern fly and animal development. Development is a continuum in which every pattern of gene activity has a preceding causal basis. The entire process includes tens of thousands of regulatory interactions and outputs.

We will focus on a few connections between genes in different levels of the hierarchies that lay out the basic segmental body plan and on *nodal points* where key genes integrate multiple regulatory inputs and respond by producing simpler gene-expression outputs.

Maternal gradients and gene activation

The Bicoid protein is a homeodomain-type transcription factor that is translated from maternally derived mRNA that is deposited in the egg and localized at the anterior pole. Because the early *Drosophila* embryo is a syncytium with all nuclei in one cytoplasm, and lacks any cell membranes that would impede the diffusion of protein molecules, the Bicoid protein can diffuse through the cytoplasm. This diffusion establishes a protein concentration gradient ([Figure 13-17a](#)): the Bicoid protein is highly concentrated at the anterior end, and this concentration gradually decreases as distance from that end increases, until there is very little Bicoid protein beyond the middle of the embryo. This concentration gradient provides positional information about the location along the anteroposterior axis. A high concentration means anterior end, a lower concentration means middle, and so on. Thus, a way to ensure that a gene is activated in only one location along the axis is to link gene expression to the concentration level. A case in point is the gap genes, which must be activated in specific regions along the axis.

Gap genes are activated by specific maternally provided proteins



Griffiths et al., *Introduction to Genetic Analysis*, 12e,
© 2020 W. H. Freeman and Company

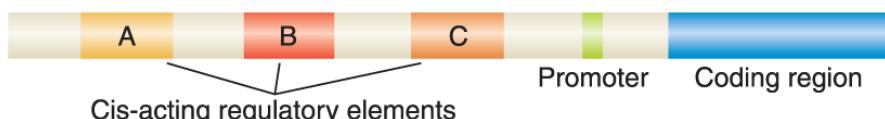
FIGURE 13-17 The Bicoid protein activates zygotic expression of the *hunchback* gene. (a) Bicoid protein expression is graded along the anteroposterior axis. The *hunchback* gap gene is expressed in the anterior half of the zygote. (b) The Bicoid protein (blue) binds to three sites 5' of the *hunchback* gene. When this 5' DNA is placed upstream of a reporter gene, reporter-gene expression recapitulates the pattern of *hunchback* expression (*top right*). However, progressive deletion of one, two, or all three Bicoid-binding sites either leads to more restricted expression of the reporter gene or abolishes it altogether. These observations show that the level and pattern of *hunchback* expression are controlled by Bicoid through its binding to *hunchback* DNA regulatory sequences.

Several zygotic genes, including gap genes, are regulated by different levels of the Bicoid protein. For example, the *hunchback* gene is a gap gene activated in the zygote in the anterior half of the embryo. This activation is through direct binding of the Bicoid protein to three sites 5' of the promoter of the *hunchback* gene. Bicoid binds to these sites *cooperatively*; that is, the binding of one Bicoid protein molecule to one site facilitates the binding of other Bicoid molecules to nearby sites.

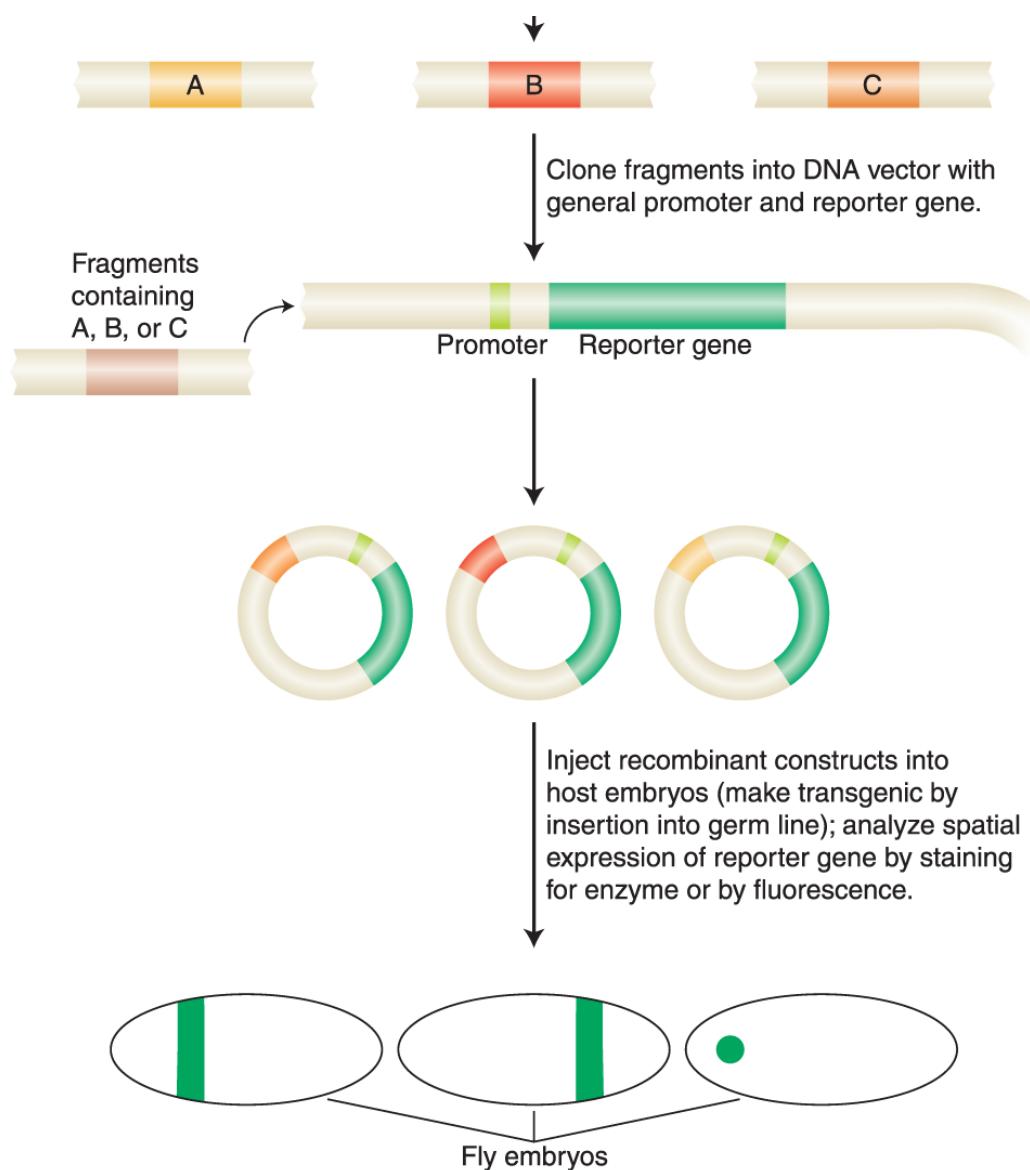
In vivo experiments can demonstrate that the activation of *hunchback* depends on the concentration gradient. These tests require linking gene regulatory sequences to a reporter gene (an enzyme-encoding gene such as the *LacZ* gene or the green fluorescent protein of jellyfish; see [Chapter 10](#)), introducing the DNA construct into the fly germ line, and monitoring reporter expression in the embryo offspring of transgenic flies (a general overview of the method is shown in [Figure 13-18](#)). The wild-type sequences 5' of the *hunchback* gene are sufficient to drive reporter expression in the anterior half of the embryo. Importantly, deletions of Bicoid-binding sites in this *cis*-acting regulatory element reduce or abolish reporter expression ([Figure 13-17b](#)). More than one Bicoid site must be occupied to generate a sharp boundary of reporter expression, which indicates that a threshold concentration of Bicoid protein is required to occupy multiple sites before gene expression is activated. A gap gene with fewer binding sites will not be activated at locations with lower concentration of Bicoid protein.

Analysis of *cis*-acting regulatory elements with reporter genes

Toolkit gene



Isolate *cis*-acting regulatory DNA fragments.



Griffiths et al., *Introduction to Genetic Analysis*, 12e, © 2020 W. H. Freeman and Company

FIGURE 13-18 Toolkit loci (such as *hunchback*, as described in the text) often contain multiple independent cis-acting regulatory elements that control gene expression in different places or at different times during development or both (for example, A, B, C, here). These elements are identified by their ability, when placed in cis to a reporter gene and inserted back into a host genome, to control the pattern, timing, or level, or all three, of reporter-gene expression. In this example, each element drives a different pattern of gene expression in a fly embryo. Most reporter genes encode enzymes or fluorescent proteins that can be easily visualized.

SH Each gap gene contains cis-acting regulatory elements with different arrangements of binding sites, and these binding sites may have different affinities for the Bicoid protein. Consequently, each gap gene is expressed in a unique distinct domain in the embryo, in response to different levels of Bicoid and other transcription-factor gradients. A similar theme is found in the patterning of the dorsoventral axis: cis-acting regulatory elements contain different numbers and

arrangements of binding sites for the maternally supplied Dorsal protein and other zygotic transcription factors. Consequently, genes are activated in discrete domains along the dorsoventral axis.

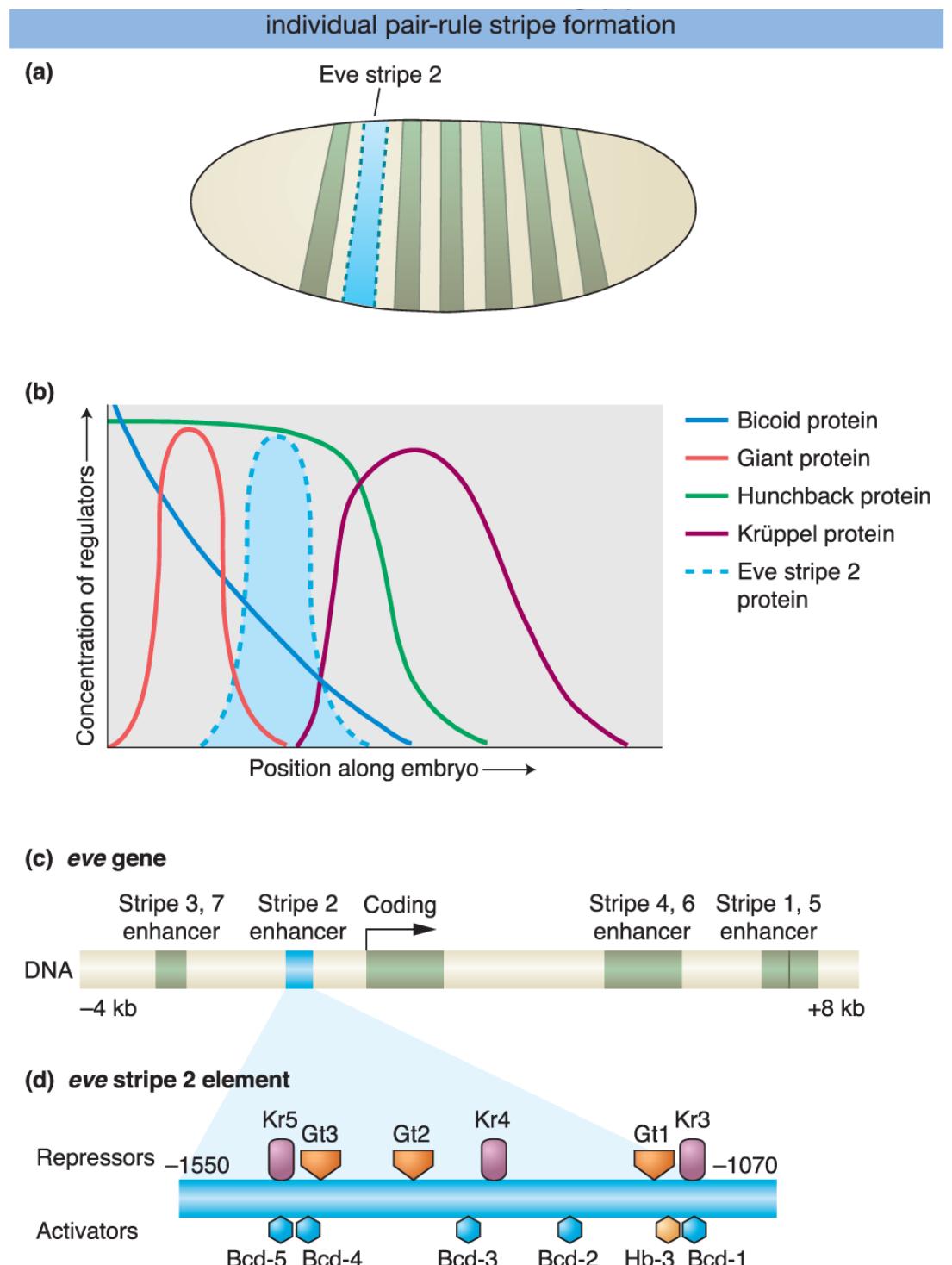
KEY CONCEPT The concentration-dependent response of genes to graded inputs is a crucial feature of gene regulation in the early *Drosophila* embryo. The cis-acting regulatory elements governing distinct responses contain different numbers and arrangements of transcription-factor-binding sites.

Drawing stripes: integration of gap-protein inputs

The expression of each pair-rule gene in seven stripes is the first sign of the periodic organization of the embryo and future animal. How are such periodic patterns generated from prior aperiodic information? Before the molecular analysis of pair-rule-gene regulation, several models were put forth to explain stripe formation. Every one of these ideas viewed all seven stripes as identical outputs in response to identical inputs. However, the actual way in which the patterns of a few key pair-rule genes are encoded and generated is one stripe at a time. The solution to the mystery of stripe generation highlights one of the most important concepts concerning the spatial control of gene regulation in developing animals; namely, the distinct cis-acting regulatory elements of individual genes are controlled independently.

The key discovery was that each of the seven stripes that make up the expression patterns of the *even-skipped* and *hairy* pair-rule genes is controlled independently. Consider the second stripe expressed by the *even-skipped* gene ([Figure 13-19a](#)). This stripe lies within the broad region of *hunchback* expression and on the edges of the regions of expression of two other gap proteins, Giant and Krüppel ([Figure 13-19b](#)). Thus, within the area of the future stripe, there will be large amounts of Hunchback protein and small amounts of Giant protein and Krüppel protein. There will also be a certain concentration of the maternal-effect Bicoid protein. No other stripe of the embryo will contain these proteins in these proportions. The formation of stripe 2 is controlled by a specific cis-acting regulatory element, an [enhancer](#), that contains a number of binding sites for these four proteins ([Figure 13-19c](#)). Detailed analysis of the *eve* stripe 2 cis-acting regulatory element revealed that the position of this “simple” stripe is controlled by the binding of these four aperiodically distributed transcription factors, including one maternal protein and three gap proteins.

Combinations of maternal-effect and gap proteins control



Griffiths et al., *Introduction to Genetic Analysis*, 12e, © 2020 W. H. Freeman and Company

FIGURE 13-19 Regulation of a pair-rule stripe: combinatorial control of an independent cis-acting regulatory element. (a) The regulation of the *eve* stripe 2 cis-acting regulatory element controls the formation of the second stripe of *eve* expression in the early embryo, just one of seven stripes of *eve* expression. (b) The stripe forms within the domains of the Bicoid (Bcd) and Hunchback (Hb) proteins and at the edge of the Giant (Gt) and Krüppel (Kr) gap proteins. Bcd and Hb are activators, Gt and Kr are repressors of the stripe. (c) The *eve* stripe 2 element is just one of several cis-acting regulatory elements of the *eve* gene, each of which controls different parts of *eve* expression. The *eve* stripe 2 element spans from about 1 to 1.7 kb upstream of the *eve* transcription unit. (d) Within the *eve* stripe 2 element, several binding sites exist for each transcription factor (repressors are shown above the element, activators below). The net output of this combination of activators and repressors is expression of the narrow *eve* stripe.

Specifically, the *eve* stripe 2 element contains multiple sites for the maternal Bicoid protein and the Hunchback, Giant, and Krüppel gap proteins ([Figure 13-19d](#)). Mutational analyses of different combinations of binding sites revealed that Bicoid and Hunchback activate the expression of the *eve* stripe 2 element over a broad region. The Giant and Krüppel proteins are repressors that sharpen the boundaries of the stripe to just a few cells wide. The *eve* stripe 2 element acts, then, as a genetic switch, integrating multiple regulatory protein activities to produce one stripe from three to four cells wide in the embryo.

The entire seven-striped periodic pattern of *even-skipped* expression is the sum of different sets of inputs into separate *cis*-acting regulatory elements. The enhancers for other stripes contain different combinations of protein binding sites.

KEY CONCEPT The regulation of *cis*-acting regulatory elements by combinations of activators and repressors is a common theme in the spatial regulation of gene expression. Complex patterns of inputs are often integrated to produce simpler patterns of outputs.

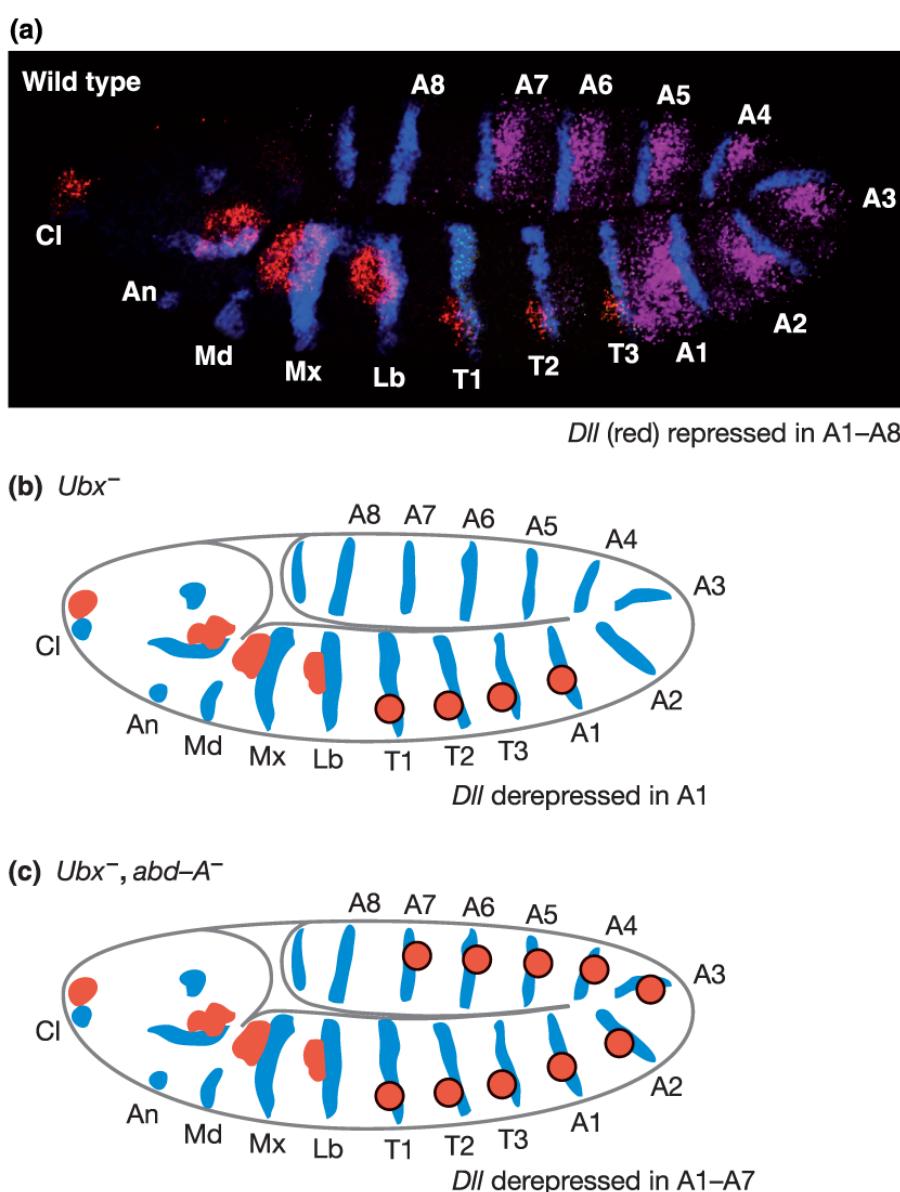
Making segments different: integration of Hox inputs

The combined and sequential activity of the maternal-effect, gap, pair-rule, and segment-polarity proteins establishes the basic segmented body plan of the embryo and larva. How are the different segmental identities established by Hox proteins? This process has two aspects. First, the *Hox* genes are expressed in different domains along the anteroposterior axis. *Hox*-gene expression is largely controlled by segmentation proteins, especially gap proteins, through mechanisms that are similar to those already described herein for *hunchback* and *eve* stripe 2 (as well as some cross-regulation by Hox proteins of other *Hox* genes). The regulation of *Hox* genes will not be considered in depth here. The second aspect of Hox control of segmental identity is the regulation of target genes by Hox proteins. We will examine one example that nicely illustrates how a major feature of the fruit fly's body plan is controlled through the integration of many inputs by a single *cis*-acting regulatory element.

The paired limbs, mouthparts, and antennae of *Drosophila* each develop from initially small populations of about 20 cells in different segments. Different structures develop from the different segments of the head and thorax, whereas the abdomen is limbless. The first sign of the

development of these structures is the activation of regulatory genes within small clusters of cells, which are called the appendage *primordia*. The expression of the *Distal-less* (*Dll*) gene marks the start of the development of the appendages. This gene is one of the key targets of the *Hox* genes, and its function is required for the subsequent development of the distal parts of each of these appendages. The small clusters of cells expressing *Distal-less* arise in several head segments and in each of the three thoracic segments, but not in the abdomen ([Figure 13-20a](#)).

Hox proteins repress appendage formation in the abdomen



Griffiths et al., *Introduction to Genetic Analysis*, 12e, © 2020 W. H.

Freeman and Company

(a) Photomicrograph by Dave Kosman, Ethan Bier, and Bill McGinnis

FIGURE 13-20 The absence of limbs in the abdomen is controlled by *Hox* genes. (a) The expression of the *Distal-less* (*Dll*) gene (red) marks the position of future appendages, expression of the *Hox* gene *Ultrabithorax* (purple) marks the position of the abdominal segments A1 through A7, and expression of the *engrailed* gene (blue) marks the posterior of each segment. (b) Schematic representation of *Ubx*⁻ embryo showing that *Dll* expression (red circles) is derepressed in segment A1. (c)

Schematic representation of $Ubx^- abd-A^-$ embryo showing that Dll expression (red circles) is derepressed in the first seven abdominal segments. [(b and c) Data from B. Gebelein, D. J. McKay, and R. S. Mann, "Direct Integration of Hox and Segmentation Gene Inputs During *Drosophila* Development," *Nature* 431, 2004, 653–659.]

How is *Distal-less* expression restricted to the more anterior segments? Several lines of evidence have revealed that the *Distal-less* gene is repressed in the abdomen by two Hox proteins—the Ultrabithorax and Abdominal-A proteins—working in collaboration with two segmentation proteins. Notice in [Figure 13-6](#) that Ultrabithorax is expressed in abdominal segments one through seven, and Abdominal-A is expressed in abdominal segments two through seven, overlapping with all but the first segment covered by Ultrabithorax. In *Ultrabithorax* mutant embryos, *Distal-less* expression expands to the first abdominal segment ([Figure 13-20b](#)), and in *Ultrabithorax*/Abdominal-A double-mutant embryos, *Distal-less* expression extends through the first seven abdominal segments ([Figure 13-20c](#)), indicating that both proteins are required for the repression of *Distal-less* expression in the abdomen.

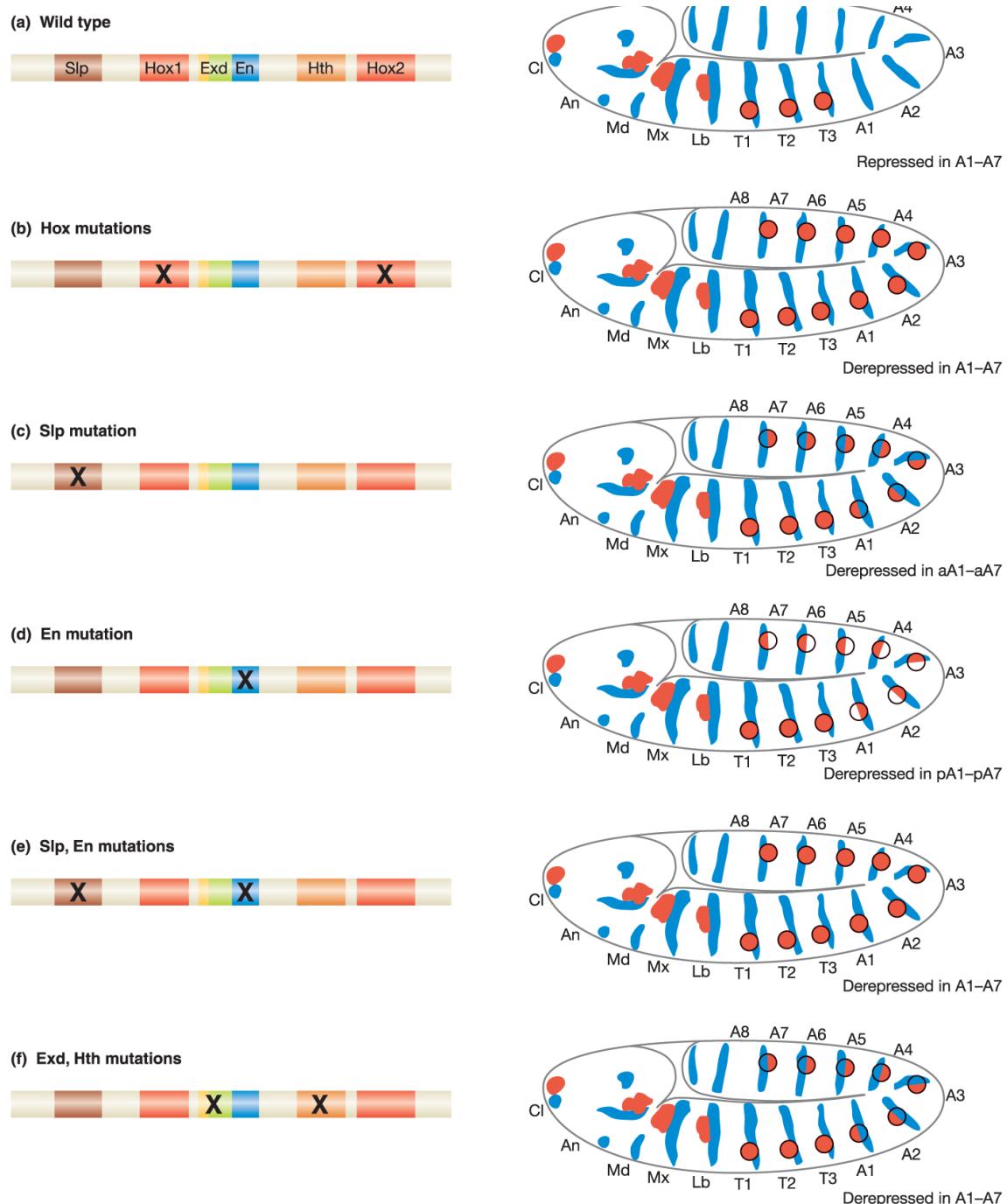
The cis-acting regulatory element responsible for *Distal-less* expression in the embryo has been identified and characterized in detail ([Figure 13-21a](#)). It contains two binding sites for the Hox proteins. If these two binding sites are mutated such that the Hox proteins cannot bind, *Distal-less* expression is derepressed in the abdomen ([Figure 13-21b](#)). Several additional proteins collaborate with the Hox proteins in repressing *Distal-less*. Two are proteins encoded by segment-polarity genes, *Sloppy-paired* (*Slp*) and *engrailed* (*en*). The Sloppy-paired and Engrailed proteins are expressed in stripes that mark the anterior and posterior compartments of each segment, respectively. Each protein also binds to the *Distal-less* cis-acting regulatory element. When the Sloppy-paired-binding site is mutated in the cis-acting regulatory element, reporter-gene expression is derepressed in the anterior compartments of abdominal segments ([Figure 13-21c](#)). When the Engrailed-binding site is mutated, reporter expression is derepressed in the posterior compartments of each abdominal segment ([Figure 13-21d](#)). And when the binding sites for both proteins are mutated, reporter-gene expression is derepressed in both compartments of each abdominal segment, just as when the Hox-binding sites are mutated ([Figure 13-21e](#)). Two other proteins, called Extradenticle and Homothorax, which are broadly expressed in every segment, also bind to the *Distal-less* cis-acting regulatory element and are required for transcriptional repression in the abdomen ([Figure 13-21f](#)).

Hox proteins and segment-polarity proteins control appendage location

CIS-ACTING REGULATORY ELEMENT

REPORTER-GENE EXPRESSION





Griffiths et al., *Introduction to Genetic Analysis*, 12e, © 2020 W. H. Freeman and Company

FIGURE 13-21 Integration of Hox and segmentation-protein inputs by a cis-acting regulatory element. (a) *Left*: A cis-acting regulatory element of the *Dll* gene governs the repression of *Dll* expression in the abdomen by a set of transcription factors. (a) *Right*: *Dll* expression (red) extends to the thorax but not into the abdomen in a wild-type embryo. (b-f) Mutations in the respective binding sites shown derepress *Dll* expression in various patterns in the abdomen. Binding sites are: Slp, Sloppy-paired; Hox1 and Hox2, Ultrabithorax and Abdominal-A; Exd, Extradenticle; En, Engrailed; Hth, Homothorax. [Data from B. Gebelein, D. J. McKay, and R. S. Mann, “Direct Integration of Hox and Segmentation Gene Inputs During Drosophila Development,” *Nature* 431, 2004, 653–659.]

Thus, altogether, two Hox proteins and four other transcription factors bind within a span of 57 base pairs and act together to repress *Distal-less* expression and, hence, appendage formation in

the abdomen. The repression of *Distal-less* expression is a clear demonstration of how Hox proteins regulate segment identity and the number of reiterated body structures. It is also a good illustration of how diverse regulatory inputs act combinatorially on cis-acting regulatory elements. In this instance, the presence of Hox-binding sites is not sufficient for transcriptional repression: collaborative and cooperative interactions are required among several proteins to fully repress gene expression in the abdomen.

KEY CONCEPT Combinatorial and cooperative regulation of gene transcription imposes greater specificity on spatial patterns of gene expression and allows for their greater diversity.

Although evolutionary diversity has not been explicitly addressed in this chapter, the presence of multiple independent cis-acting regulatory elements for each toolkit gene has profound implications for the evolution of form. Specifically, the modularity of these elements allows for changes in one aspect of gene expression independent of other gene functions. The evolution of gene regulation plays a major role in the evolution of development and morphology. We will return to this topic in [Chapter 20](#).

13.5 POST-TRANSCRIPTIONAL REGULATION OF GENE EXPRESSION IN DEVELOPMENT

LO 13.4 Infer how spatially and temporally restricted patterns of gene expression are generated during development from analyses of genetic mutations.

Although transcriptional regulation is a major means of restricting the expression of gene products to defined areas during development, it is not at all the exclusive means of doing so. Alternative RNA splicing also contributes to gene regulation, and so does the regulation of mRNA translation by proteins and microRNAs (miRNAs). In each case, regulatory sequences in RNA are recognized—by splicing factors, mRNA-binding proteins, or miRNAs—and govern the structure of the protein product, its amount, or the location where the protein is produced. We will look at one example of each type of regulatory interaction at the RNA level.

RNA splicing and sex determination in *Drosophila*

A fundamental developmental decision in sexually reproducing organisms is the specification of sex. In animals, the development of many tissues follows different paths, depending on the sex of the individual animal. In *Drosophila*, many genes have been identified that govern *sex determination* through the analysis of mutant phenotypes in which sexual identity is altered or ambiguous.

The *doublesex* (*dsx*) gene plays a central role in governing the sexual identity of somatic (non-germ-line) tissue. Null mutations in *dsx* cause females and males to develop as intermediate *intersexes*, which have lost the distinct differences between male and female tissues. Although *dsx* function is required in both sexes, different gene products are produced from the locus in different sexes. In males, the product is a specific, longer isoform, Dsx^M , that contains a unique C-terminal region of 150 amino acids not found in the female-specific isoform Dsx^F , which instead contains a unique 30-amino-acid sequence at its carboxyl terminus. Each form of the *Dsx* protein is a DNA-

binding transcription factor that apparently binds the same DNA sequences. However, the activities of the two isoforms differ: Dsx^F activates certain target genes in females that Dsx^M represses in males.

The alternative forms of the Dsx protein are generated by alternative splicing of the primary dsx RNA transcript. Thus, in this case, the choice of splice sites must be regulated to produce mature mRNAs that encode different proteins. The various genetic factors that influence Dsx expression and sex determination have been identified by mutations that affect the sexual phenotype.

One key regulator is the product of the *transformer* (*tra*) gene. Whereas null mutations in *tra* have no effect on males, XX female flies bearing *tra* mutations are transformed into the male phenotype. The Tra protein is an alternative splicing factor that affects the splice choices in the dsx RNA transcript. In the presence of Tra (and a related protein Tra2), a splice occurs that incorporates exon 4 of the dsx gene into the mature dsx^F transcript (Figure 13-22), but not exons 5 and 6. Males lack the Tra protein; so this splice does not occur, and exons 5 and 6 are incorporated into the dsx^M transcript, but not exon 4.

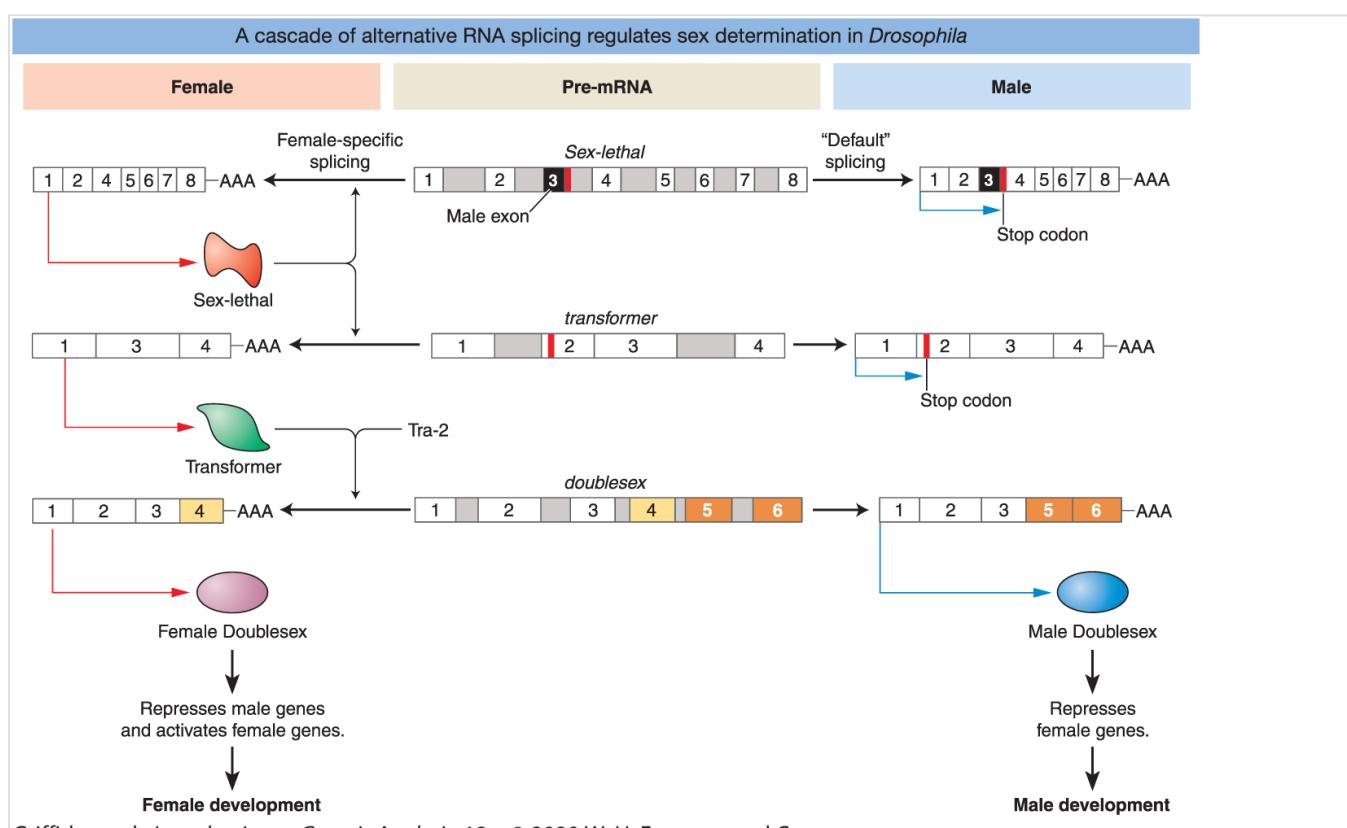


FIGURE 13-22 Three pre-mRNAs of major *Drosophila* sex-determining genes are alternatively spliced. The female-specific pathway is shown on the left and the male-specific pathway shown on the right. The pre-mRNAs are identical in both sexes and shown in the middle. In the male *Sex-lethal* and *transformer* mRNAs, there are stop codons that terminate translation. These sequences are removed by splicing to produce functional proteins in the female. The Transformer and Tra-2 proteins

then splice the female *doublesex* pre-mRNA to produce the female-specific isoform of the Dsx protein, which differs from the male-specific isoform by the alternative splicing of several exons.

ANIMATED ART Sapling Plus

Sex determination in *Drosophila*

The Tra protein explains how alternative forms of Dsx are expressed, but how is Tra expression itself regulated to differ in females and males? The *tra* RNA itself is alternatively spliced. In females, a splicing factor encoded by the *Sex-lethal (Sxl)* gene is present. This splicing factor binds to the *tra* RNA and prevents a splicing event that would otherwise incorporate an exon that contains a stop codon. In males, no Tra protein is made because this stop codon is present.

The production of the Sex-lethal protein is, in turn, regulated both by RNA splicing and by factors that alter the level of transcription. The level of *Sxl* transcription is initially governed by activators on the X chromosome and repressors on the autosomes. In females, which have two X chromosomes and therefore a double dose of activators, *Sxl* activation prevails and the *Sxl* protein is produced, which regulates *tra* RNA splicing and feeds back to regulate the splicing of *Sxl* RNA itself. In females, a stop codon is spliced out so that *Sxl* protein production can continue. However, in males, which have only one X chromosome and therefore only half the dose of X-linked activators, transcription of *Sxl* is initially repressed. Later, *Sxl* transcription is activated in males, but the absence of *Sxl* protein means that the stop codon is still present in unspliced *Sxl* RNA transcript and no *Sxl* protein can be produced.

This cascade of sex-specific RNA splicing in *D. melanogaster* illustrates one way that the sex-chromosome genotype leads to different forms of regulatory proteins being expressed in one sex and not the other. Interestingly, the genetic regulation of sex determination differs greatly between animal species, in that sexual genotype can lead to differential expression of regulatory genes through distinctly different paths. However, proteins related to Dsx do play roles in sexual differentiation in a wide variety of animals, including humans. Thus, although there are many ways to generate differential expression of transcription factors, a family of similar proteins plays conserved roles in sexual differentiation across a diversity of species.

KEY CONCEPT The sex determination pathway in *D. melanogaster* is an example of how the spatial and temporal expression of genes involved in developmental pathways can be regulated by differential splicing.

Regulation of mRNA translation and cell lineage in *C. elegans*

In many animal species, the early development of the embryo entails the partitioning of cells or groups of cells into discrete lineages that will give rise to distinct tissues in the adult. This process is best understood in the nematode worm *C. elegans*, in which the adult animal is composed of just about 1000 somatic cells (a third of which are nerve cells) and a similar number of germ cells in the gonad. The simple construction, rapid life cycle, and transparency of *C. elegans* has made it a powerful model for developmental analysis (see [the Model Organism box on *C. elegans* on page 451](#)). All of this animal's cell lineages were mapped out in a series of elegant studies led by John Sulston at the Medical Research Council (MRC) Laboratory of Molecular Biology in Cambridge, England. Systematic genetic screens for mutations that disrupt or extend cell lineages have provided a bounty of information about the genetic control of lineage decisions. *C. elegans* genetics has been especially important in understanding the role of post-transcriptional regulation at the RNA level, and we will examine two mechanisms here: (1) control of translation by mRNA-binding proteins, and (2) miRNA control of gene expression.

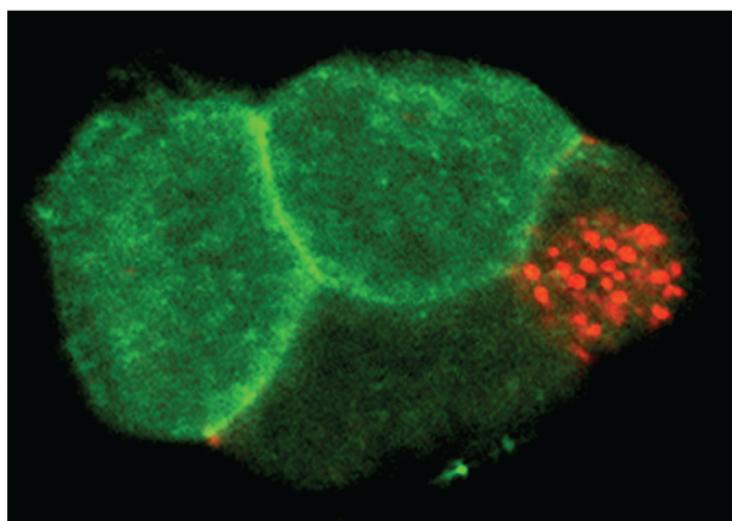
Translational control in the early embryo

We first look at how a cell lineage begins. After two cell divisions, the *C. elegans* embryo contains four cells, called blastomeres. Each cell will begin a distinct lineage, and the descendants of the separate lineages will have different fates. Already at this stage, differences are observed in the proteins present in the four blastomeres. However, the mRNAs encoding some worm toolkit proteins are present in *all* cells of the early embryo, and post-transcriptional regulation determines which of these mRNAs will be translated into proteins. Thus, in the *C. elegans* embryo, post-transcriptional regulation is critical for the proper specification of early cell fates. During the very first cell division, polarity within the zygote leads to the partitioning of regulatory molecules to specific embryonic cells. For example, the *glp-1* gene encodes a transmembrane receptor protein (related to the Notch receptor of flies and other animals). Although the *glp-1* mRNA is present in all cells at the four-cell stage, the GLP-1 protein is translated only in the two anterior cells ABa

and ABp (**Figure 13-23a**). This localized expression of GLP-1 is critical for establishing distinct fates. Mutations that abolish *glp-1* function at the four-cell stage alter the fates of ABp and ABa descendants.

mRNA-binding proteins repress mRNA translation to determine cell lineages

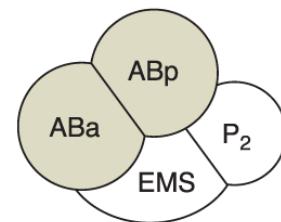
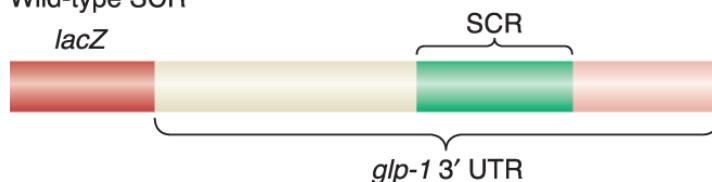
(a)



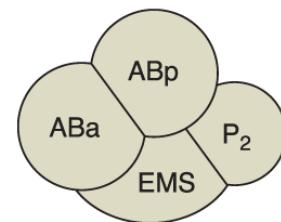
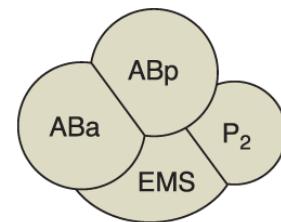
mRNA–reporter-gene construct

Reporter expression

(b) Wild-type SCR



(c) Mutated SCR

(d) Wild-type SCR in *gld⁻* embryo

Griffiths et al., *Introduction to Genetic Analysis*, 12e, © 2020 W. H. Freeman and Company
 (a) Courtesy of Thomas C. Evans, University of Colorado Anschutz Medical Campus.

FIGURE 13-23 Translational regulation and cell-lineage decisions in the early *C. elegans* embryo. (a) At the four-cell stage of the *C. elegans* embryo, the GLP-1 protein is expressed in two anterior cells ABa and ABp (bright green), but not in the EMS or P₂ cell (red). Translation of the *glp-1* mRNA is regulated by the GLD-1 protein in posterior cells. (b) Fusion of the *glp-1* 3' UTR to the *lacZ* reporter gene leads to reporter expression in the ABa and ABp cells of the four-cell stage of the *C.*

elegans embryo (shaded, right). (c) Mutations in GLD-1-binding sites in the spatial control region (SCR) cause derepression of translation in the EMS and P₂ lineages, as does (d) loss of *gld* function.

GLP-1 is localized to the anterior cells by repressing its translation in the posterior cells. The repression of GLP-1 translation requires sequences in the 3' UTR of the *glp-1* mRNA—specifically, a 61-nucleotide region called the spatial control region (SCR). The importance of the SCR has been demonstrated by linking mRNA transcribed from reporter genes to different variants of the SCR. Deletion of this region or mutation of key sites within it causes the reporter gene to be expressed in all four blastomeres of the early embryo ([Figure 13-23c](#)).

On the basis of how we have seen transcription controlled, we might guess that one or more proteins bind(s) to the SCR to repress translation of the *glp-1* mRNA. To identify these repressor proteins, researchers isolated proteins that bind to the SCR. One protein, GLD-1, binds specifically to a region of the SCR. Furthermore, the GLD-1 protein is enriched in posterior blastomeres, just where the expression of *glp-1* is repressed. Finally, when GLD-1 expression is inhibited by using RNA interference, the GLP-1 protein is expressed in posterior blastomeres ([Figure 13-23d](#)). This evidence suggests that GLD-1 is a translational repressor protein controlling the expression of *glp-1*.

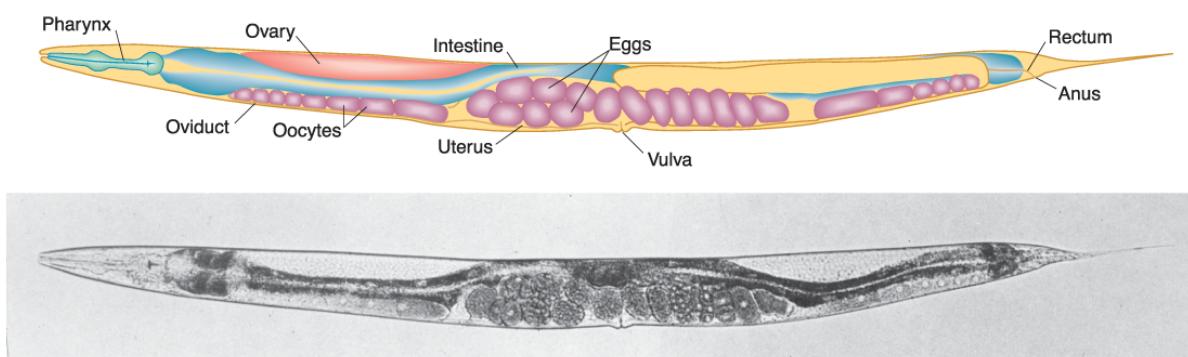
The spatial regulation of GLP-1 translation is but one example of translational control in development. Translational control is also important in the establishment of the anteroposterior axis in *Drosophila* and the development of sperm in mammals. Again, we see that genetic analysis in model organisms can reveal deeply conserved mechanisms for the regulation of gene expression.

MODEL ORGANISM

Caenorhabditis elegans

The Nematode *Caenorhabditis elegans* as a Model for Cell-Lineage-Fate Decisions

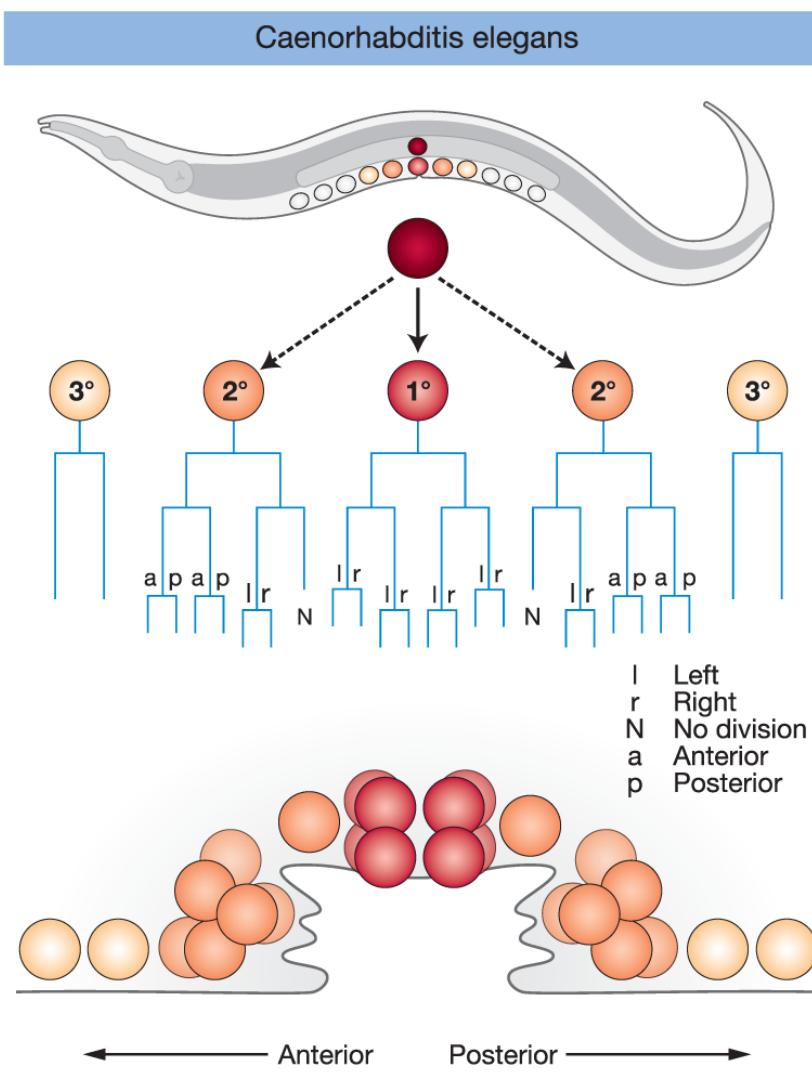
In the past 20 years, studies of the nematode worm *Caenorhabditis elegans* (see [the upper figure](#)) have greatly advanced our understanding of the genetic control of cell-lineage decisions. The transparency and simple construction of this animal led Sydney Brenner to establish its use as a model organism. The adult worm contains about 1000 somatic cells, and researchers, led by John Sulston, have carefully mapped out the entire series of somatic-cell decisions that produce the adult animal.



Republished with permission of Elsevier, from J. E. Sulston and H. R. Horvitz, "Post-embryonic Cell Lineages of the Nematode, *Caenorhabditis elegans*," *Developmental Biology*, 1977, March; 56(1):110–56, Figure 1. Permission conveyed through Copyright Clearance Center, Inc.

An adult hermaphrodite *Caenorhabditis elegans*, showing various organs.

Some of the lineage decisions, such as the formation of the vulva (the opening from which eggs are laid), have been key models of so-called *inductive interactions* in development, where signaling between cells induces cell-fate changes and organ formation (see [the lower figure](#)). Exhaustive genetic screens have identified many components participating in signaling and signal transduction involved in the specification of the different cell types that form the vulva.





Republished with permission of Jennifer L. Green, Takao Inoue and Paul W. Sternberg, *Development and The Company of Biologists*, The *C. elegans* ROR receptor tyrosine kinase, CAM-1, nonautonomously inhibits the Wnt pathway, of Jennifer L. Green, Takao Inoue and Paul W. Sternberg, *Development* 134, 4053–4062 (2007), and The Company of Biologists; permission conveyed through Copyright Clearance Center, Inc.

Production of the vulval-cell lineages. Parts of the vulval anatomy are occupied by so-called primary (1°), secondary (2°), and tertiary (3°) cells. The lineages or pedigrees of the primary, secondary, and tertiary cells are distinguished by their cell-division patterns and will give rise to different parts of the vulva in the adult worm, as shown in the bottom image.

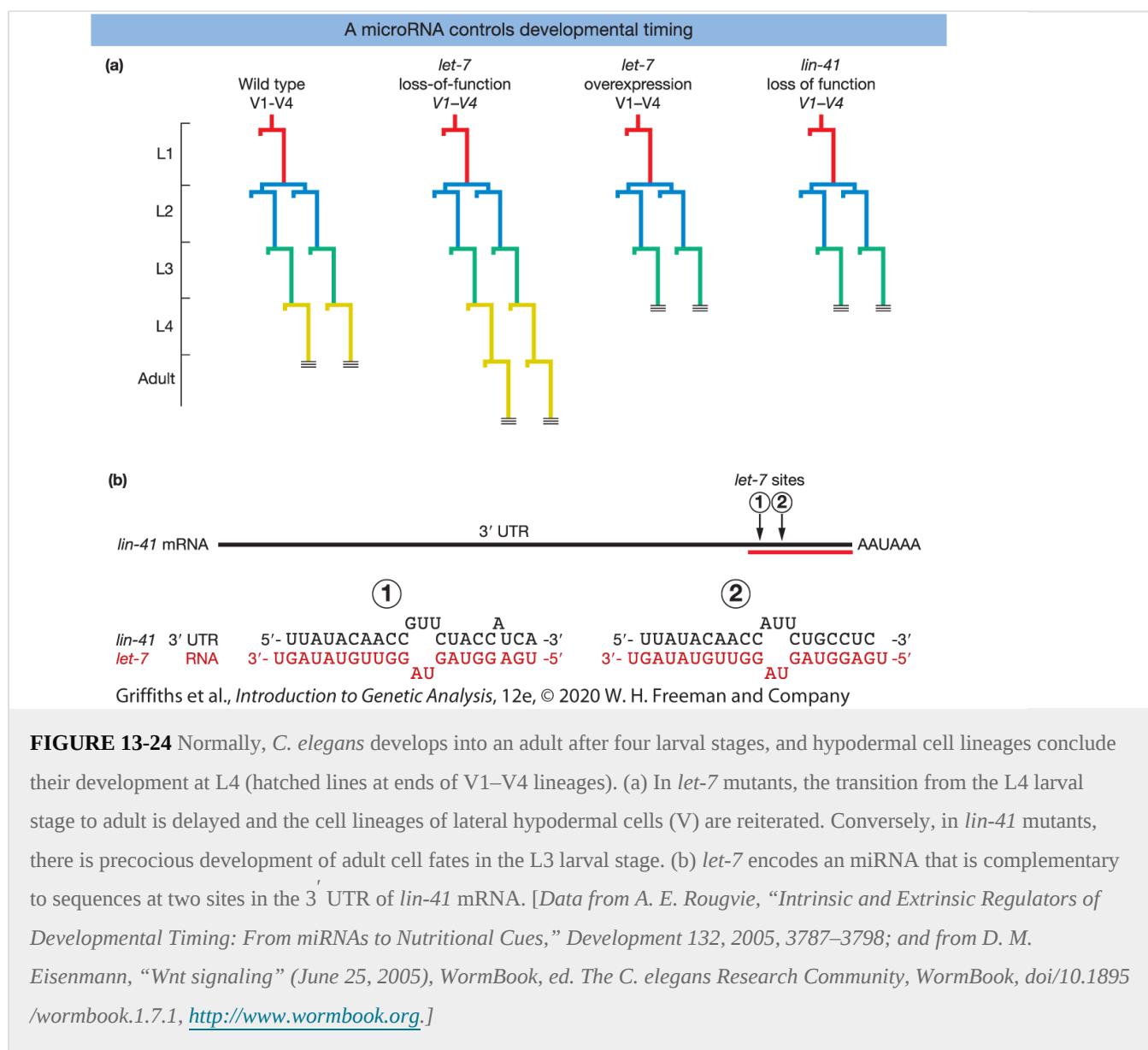
For some of the embryonic and larval cell divisions, particularly those that will contribute to a worm's nervous system, a progenitor cell gives rise to two progeny cells, one of which then undergoes programmed cell death. Analysis of mutants in which programmed cell death is aberrant, led by Robert Horvitz, has revealed many components of programmed-cell-death pathways common to most animals. Sydney Brenner, John Sulston, and Robert Horvitz shared the 2002 Nobel Prize in Physiology or Medicine for their pioneering work based on *C. elegans*.

miRNA control of developmental timing in *C. elegans* and other species

Development is a temporally as well as spatially ordered process. When events take place is just as important as where. Mutations in the **heterochronic genes** of *C. elegans* have been sources of insight into the control of developmental timing. Mutations in these genes alter the timing of events in cell-fate specification, causing such events to be either reiterated or omitted. Detailed investigation into the products of heterochronic genes led to the discovery of an entirely unexpected mechanism for regulating gene expression, through microRNAs (see [Chapter 9](#)).

Among the first members of this class of regulatory molecules to be discovered in *C. elegans* is RNA produced by the *let-7* gene. The *let-7* gene regulates the transition from late-larval to adult cell fates. In *let-7* mutants, for example, larval cell fates are reiterated in the adult stage ([Figure](#)

13-24a). Conversely, increased *let-7* gene dosage causes the precocious specification of adult fates in larval stages.

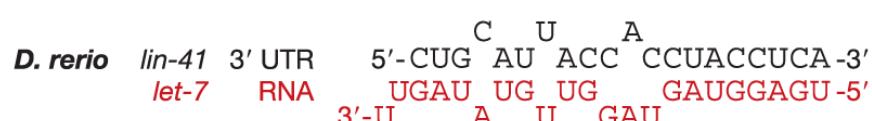


The *let-7* gene does not encode a protein. Instead, it encodes a temporally regulated mature 22-nucleotide RNA that is processed from an approximately 70-nucleotide precursor. The mature RNA is complementary to sequences in 3' untranslated regions of a variety of developmentally regulated genes, and the binding of the miRNA to these sequences hinders translation of these gene transcripts. One of these target genes, *lin-41*, also affects the larval-to-adult transition. The *lin-41* mutants cause precocious specification of adult cell fates, suggesting that the effect of *let-7* overexpression is due at least in part to an effect on *lin-41* expression. The *let-7* mRNA binds to *lin-41* RNA in vitro at several imperfect complementary sites (Figure 13-24b).

The role of miRNAs in *C. elegans* development extends far beyond *let-7*. Several hundred miRNAs have been identified, and many target genes have been shown to be miRNA regulated. Moreover, the discovery of this class of regulatory RNAs prompted the search for such genes in other genomes, and, in general, hundreds of candidate miRNA genes have been detected in both plant and animal genomes, including those of humans.

Quite surprisingly, the *let-7* miRNA gene is widely conserved and found in *Drosophila*, ascidian, mollusc, annelid, and vertebrate (including human) genomes. The *lin-41* gene also is conserved, and evidence suggests that the *let-7–lin-41* regulatory interaction also controls the timing of events in the development of other species, such as mouse and zebrafish (**Figure 13–25**).

The *lin-41*/*let-7* interaction is conserved across phyla



Griffiths et al., *Introduction to Genetic Analysis*, 12e, © 2020 W. H. Freeman and Company

FIGURE 13-25 The sequences of both the *let-7* miRNA and its binding site in the 3' UTR of the *lin-41* mRNA are conserved across *C. elegans*, *D. rerio* (zebrafish), and *D. melanogaster*. [Data from A. E. Pasquinelli et al., “Conservation of the Sequence and Temporal Expression of *let-7* Heterochronic Regulatory RNA,” *Nature* 408, 2000, 86–89.]

The discoveries of miRNA regulation of developmental genes and of the scope of the miRNA repertoire are fairly recent. Geneticists and other biologists are quite excited about the roles of this class of regulatory molecules in normal development, as well as in the pathology and treatment of disease, leading to a very vigorous, fast-paced area of new research.

KEY CONCEPT Sequence-specific RNA-binding proteins and micro RNAs act through cis-acting sequences in the 3' untranslated regions of mRNAs to regulate the spatial and temporal pattern of protein translation.

13.6 FROM FLIES TO FINGERS, FEATHERS, AND FLOOR PLATES: THE MANY ROLES OF INDIVIDUAL TOOLKIT GENES

LO 13.1 Outline experimental approaches to identify and characterize members of the genetic toolkit for development in different animal phyla.

LO 13.5 Summarize the evidence that the genetic toolkit for development is conserved across animal phyla.

We have seen that toolkit proteins and regulatory RNAs have multiple roles in development. For example, recall that the Ultrabithorax protein represses limb formation in the fly abdomen and promotes hind-wing development in the fly thorax. Similarly, Sloppy-paired and Engrailed participate in the generation of the basic segmental organization of the embryo and collaborate with Hox proteins to suppress limb formation. These roles are just a few of the many roles played by these toolkit genes in the entire course of fly development. Most toolkit genes function at more than one time and place, and most may influence the formation or patterning of many different structures that are formed in different parts of the larval or adult body. Those that regulate gene expression may directly regulate scores to hundreds of different genes. The function of an individual toolkit protein (or RNA) is almost always context dependent, which is why the toolkit analogy is perhaps so fitting. As with a carpenter's toolkit, a common set of tools can be used to fashion many structures.

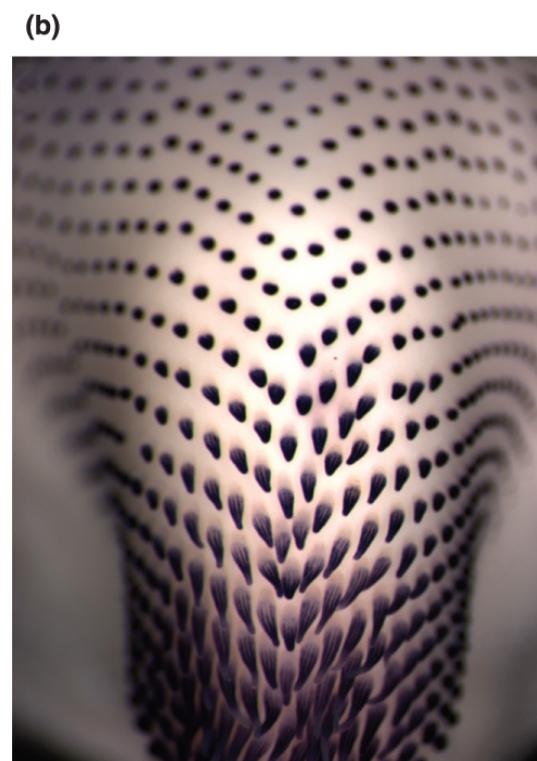
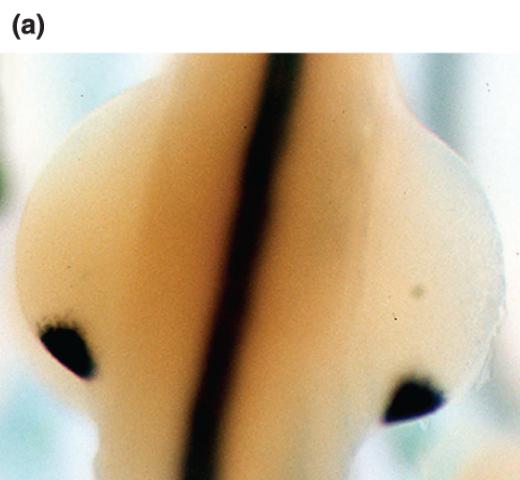
To illustrate this principle more vividly, we will look at the role of one toolkit protein in the development of many vertebrate features, including features present in humans. This toolkit protein is the vertebrate homolog of the *Drosophila hedgehog* gene. The *hedgehog* gene was first identified by Nüsslein-Volhard and Wieschaus as a segment-polarity gene. It has been characterized as encoding a signaling protein secreted from cells in *Drosophila*.

As the evidence grew that toolkit genes are common to different animal phyla, the discovery and characterization of fly toolkit genes such as *hedgehog* became a common springboard to the characterization of genes in other taxa, particularly vertebrates. The identification of homologous genes based on sequence similarity was a fast track to the identification of vertebrate toolkit genes. The application of this strategy to the *hedgehog* gene illustrates the power and payoffs of using homology to discover important genes. Several distinct homologs of *hedgehog* were isolated

from vertebrates including zebrafish, mice, chickens, and humans. In the whimsical spirit of the *Drosophila* gene nomenclature, the three vertebrate homologs were named *Sonic hedgehog* (after the video-game character), *Indian hedgehog*, and *Desert hedgehog*.

One of the first means of characterizing the potential roles of these genes in development was to examine where they are expressed. *Sonic hedgehog* (*Shh*) was found to be expressed in several parts of the developing chicken, with similar expression patterns in other vertebrates. Most intriguing was its expression in the posterior part of the developing limb bud ([Figure 13-26a](#)). This part of the limb bud was known for decades to be the *zone of polarizing activity* (ZPA) because it is an organizer responsible for establishing the anteroposterior polarity of the limb and its digits (see [Figure 13-2](#)). To test whether *Shh* might play a role in ZPA function, Cliff Tabin and his colleagues at Harvard Medical School caused the *Shh* protein to be expressed in the *anterior* region of developing chick limb buds. They observed the same effect as transplantation of the ZPA—the induction of extra digits with reversed polarity. Their results were stunning evidence that *Shh* was the long-sought morphogen produced by the ZPA ([Figure 13-27](#)).

The *Sonic hedgehog* toolkit gene has multiple roles



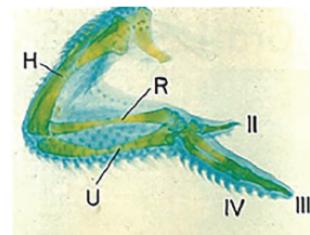
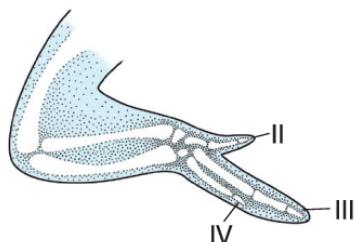
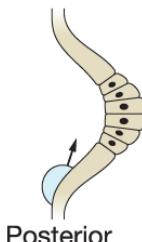
Photomicrographs courtesy of (a) Cliff Tabin and (b) Photomicrographs courtesy Dr. John Fallon, University of Wisconsin / Matthew Harris, Harvard Medical School, Department of Genetics.

FIGURE 13-26 The *Shh* gene is expressed in many different parts of the developing chick embryo (indicated by dark staining), including (a) the zone of polarizing activity in each of the two developing limb buds and the long neural tube, and (b) the developing feather buds. *Shh* mRNA is visualized by *in situ* hybridization.

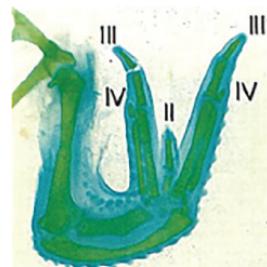
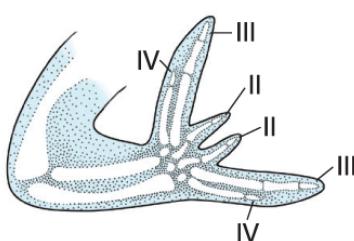
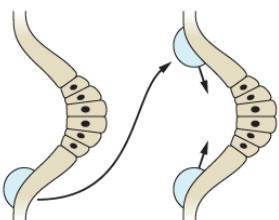
The Sonic hedgehog gene is the morphogen produced by the limb organizer

(a) Normal chicken limb

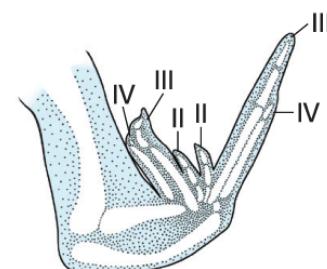
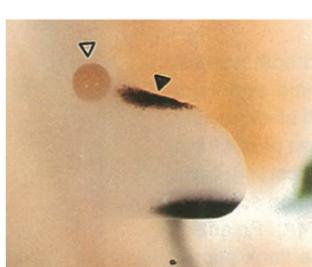
Anterior



(b) ZPA graft



(c) *Shh* ectopic expression



Figures 6 and 9 republished with permission of Elsevier, from Robert D. Riddle, Randy L. Johnson, Ed Laufer, Cliff Tabin, "Sonic hedgehog mediates the polarizing activity of the ZPA," Cell, 1993, 31 December; 75 (7): 1401–1416. [https://doi.org/10.1016/0092-8674\(93\)90626-2](https://doi.org/10.1016/0092-8674(93)90626-2). Permission conveyed through Copyright Clearance Center, Inc.

FIGURE 13-27 (a) A normal chicken limb with a single organizer, the zone of polarizing activity (ZPA), has three digits (II-III-IV). (b) Transplantation of the ZPA from a donor limb bud to the anterior position in a recipient limb bud induces extra digits with reverse polarity (IV-III-II-II-III-IV). (c) Similarly, ectopic expression of the *Shh* gene (dark staining) in the anterior limb bud results in a mirror image duplication of the digits (IV-III-II-II-III-IV).

Shh is also expressed in other intriguing patterns in the chicken and other vertebrates. For example, Shh is expressed in developing feather buds, where it plays a role in establishing the pattern and polarity of feather formation ([Figure 13-26b](#)). Shh is also expressed in the developing

neural tube of vertebrate embryos, in a region called the *floor plate* ([Figure 13-26a](#)). Subsequent experiments have shown that Shh signaling from these floor-plate cells is critical for the subdivision of the brain hemispheres and the subdivision of the developing eye into the left and right sides. When the function of the *Shh* gene is eliminated by mutation in the mouse using techniques described in [Chapters 10](#) and [14](#), these hemispheres and eye regions do not separate, and the resulting embryo is *cyclopic*, with one central eye and a single forebrain (it also lacks limb structures).

Shh is just one striking example of the dramatic and diverse roles played by toolkit genes at different places and times in development. The outcomes of Shh signaling are different in each case: the Shh signaling pathway will induce the expression of one set of genes in the developing limb, a different set in the feather bud, and yet another set in the floor plate. How are different cell types and tissues able to respond differently to the same signaling molecule? Just as we learned from investigating the genetic control of patterning in the *Drosophila* embryo, the outcome of Shh signaling depends on the integration with the signals provided by other toolkit genes that are acting at the same time and in the same place.

KEY CONCEPT Most toolkit genes have multiple roles in different tissues and cell types. The specificity of their action is determined by the context provided by the other toolkit genes that act in combination with them.

13.7 DEVELOPMENT AND DISEASE

LO 13.5 Summarize the evidence that the genetic toolkit for development is conserved across animal phyla.

The discovery that the fly genetic toolkit for development is largely conserved in vertebrates has also had a profound effect on the study of the genetic basis of human diseases, particularly of birth defects and cancer. A large number of toolkit-gene mutations have been identified that affect human development and health. We will focus here on just a few examples that illustrate how understanding gene function and regulation in model animals has translated into better understanding of human biology.

Polydactyly

A fairly common syndrome in humans is the development of extra partial or complete digits on the hands and feet. This condition, called *polydactyly*, arises in about 5 to 17 of every 10,000 live births. In the most dramatic cases, the condition is present on both hands and feet ([Figure 13-28](#)). Polydactyly occurs widely throughout vertebrates—in cats, chickens, mice, and other species.

Polydactyly in humans



Courtesy of Dr. Robert Hill, MRC Human Genetics Unit, Edinburgh, Scotland; from L. A. Lettice et al., "Disruption of a Long-Range Cis-Acting Regulator for Shh Causes Preaxial Polydactyly," Proc. Natl. Acad. Sci. USA 99, 7548. Copyright (2002) National Academy of Sciences, U.S.A.

FIGURE 13-28

This person has six fingers on each hand and seven toes on each foot owing to a regulatory mutation in the *Sonic hedgehog* gene.

The discovery of the role of *Shh* in digit patterning led geneticists to investigate whether the *Shh* gene was altered in polydactylous humans and other species. In fact, some cases of polydactyly in humans (and also in cats) result from mutations of the *Shh* gene. Importantly, the mutations are not in the coding region of the *Shh* gene; rather, they lie in a *cis*-acting regulatory element, far from the coding region, that controls *Shh* expression in the developing limb bud. The extra digits are induced by the expression of *Shh* in a part of the limb where the gene is not normally expressed. Mutations in *cis*-acting regulatory elements have two important properties that are distinct from mutations in coding regions. First, because they affect regulation in *cis*, the phenotypes are often dominant. Second, because only one of several *cis*-acting regulatory elements may be affected, other gene functions may be completely normal. Polydactyly can occur without any collateral developmental problems that would be expected given the multiple roles of *Shh* in development. For similar reasons, we will see in [Chapter 20](#) that mutations in *cis*-acting regulatory elements of toolkit genes also play key roles during the evolution of morphological differences among species. Coding mutations in *Shh*, however, tell a different story, as we will see in the next section.

Holoprosencephaly

Mutations in the human *Shh* coding region also have been identified. The consequent alterations in the *Shh* protein are associated with a syndrome termed *holoprosencephaly*, in which abnormalities occur in brain size, in the formation of the nose, and in other midline structures. These abnormalities appear to be less severe counterparts of the developmental defects observed in homozygous *Shh* mutant mice. Indeed, the affected children seen in clinics are heterozygous. One copy of a normal *Shh* gene appears to be insufficient for normal midline development (the gene is *haploinsufficient*). Human fetuses homozygous for loss-of-function *Shh* mutations very likely die in gestation with more severe defects.

Holoprosencephaly is not caused exclusively by *Shh* mutations. *Shh* is a ligand in a signal-transduction pathway. As might be expected, mutations in genes encoding other components of the pathway affect the efficiency of *Shh* signaling and are also associated with holoprosencephaly. Several components of the human *Shh* pathway were first identified as homologs of members of the fly pathway, demonstrating once again both the conservation of the genetic toolkit and the power of model systems for biomedical discovery.

Cancer as a developmental disease

In long-lived animals, such as ourselves and other mammals, development does not cease at birth or at the end of adolescence. Tissues and various cell types are constantly being replenished. The maintenance of many organ functions depends on the controlled growth and differentiation of cells that replace those that are sloughed off or otherwise die. Tissue and organ maintenance is generally controlled by signaling pathways. Inherited or spontaneous mutations in genes encoding components of these pathways can disrupt tissue organization and contribute to the loss of control of cell proliferation. Because unchecked cell proliferation is a characteristic of cancer, the formation of cancers may be a consequence. Cancer, then, is a developmental disease, a product of normal developmental processes gone awry.

Some of the genes associated with types of human cancers are shared members of the animal toolkit. For example, the *patched* gene encodes a receptor for the Hedgehog signaling proteins. In addition to causing inherited developmental disorders such as polydactyly and holoprosencephaly, mutations in the human *patched* gene are associated with the formation of a variety of cancers. About 30 to 40 percent of patients with a dominant genetic disorder called *basal cell nevus syndrome* (BCNS) carry *patched* mutations. These persons are strongly disposed to develop a type of skin cancer called basal-cell carcinoma. They also have a greatly increased incidence of medulloblastoma, a very deadly form of brain tumor. A growing list of cancers are now associated with disruptions of signal-transduction pathways—pathways that were first elucidated by these early systematic genetic screens for patterning mutants in fruit flies ([Table 13-2](#)).

TABLE 13-2 Some Toolkit Genes Having Roles in Cancer

	Fly gene	Mammalian gene	Cancer type
Signaling-Pathway Components			
Wingless	<i>armadillo</i>	β -catenin (<i>CTNNB</i>)	Colon and skin
	<i>TCF/pangolin</i>	<i>TCF/LEF</i>	Colon
Hedgehog	<i>cubitus interruptus</i>	<i>GLI1</i>	Basal-cell carcinoma
	<i>patched</i>	<i>PTCH</i>	Basal-cell carcinoma, medulloblastoma
	<i>smoothened</i>	<i>SMO</i>	Basal-cell carcinoma
Notch	<i>Notch</i>	<i>NOTCH1</i>	T-cell leukemia, lymphoma, breast
EGF receptor	<i>torpedo</i>	<i>ERBB2</i>	Breast and colon
Decapentaplegic/TGF- β	<i>Medea</i>	<i>SMAD4</i>	Pancreatic and colon
Toll	<i>dorsal</i>	<i>NF-κB</i>	Lymphoma

Homeobox

*extradenticle**PBX1*

Acute pre-B-cell leukemia

The discoveries of links between mutations of signal-transduction-pathway genes and human cancer have greatly facilitated the study of the biology of cancer and the development of new therapies. For example, about 30 percent of mice heterozygous for a targeted mutation in the *patched* gene develop medulloblastoma. These mice therefore serve as an excellent model for the biology of human disease and a testing platform for therapy.

One promising avenue for the development of new cancer therapies is to identify drugs that can specifically target and kill cancer cells without affecting normal cells. These so-called targeted therapies are already employed today for the treatment of some cancers. For example, Herceptin is a drug used to treat breast cancers with overexpression of the human *epidermal growth factor receptor 2* (*HER2* or *ERBB2*), a homolog of the *Drosophila torpedo* gene ([Table 13-2](#)). Much current research is focused on identifying additional drugs to specifically target the signal-transduction pathways that are disrupted in different types of tumors and that were often first identified in genetic screens in flies and worms.

It is fair to say that even the most optimistic and farsighted researchers did not expect that the discovery of the genetic toolkit for building a fly would have such far-ranging effects on understanding human development and disease. But such huge unforeseen dividends are familiar in the recent history of basic genetic research. The advent of genetically engineered medicines, monoclonal antibodies for diagnosis and therapy, and forensic DNA testing all had similar origins in seemingly unrelated investigations.

KEY CONCEPT Investigation into the genetic control of development in model organisms such as *Drosophila* and *C. elegans* has led to unexpected and far-reaching consequences for the understanding and treatment of human disease.

SUMMARY

In [Chapter 11](#), we mentioned the quip from Jacques Monod and François Jacob that “anything found to be true of *E. coli* must also be true of Elephants.”² Now that we have seen the regulatory processes that build worms, flies, chickens, humans, and elephants, would we say that they were right? If Monod and Jacob were referring to the principle that gene transcription is controlled by sequence-specific regulatory proteins, we have seen that the bacterial Lac repressor and the fly Hox proteins do indeed act similarly. Moreover, their DNA-binding proteins have the same type of motif. The fundamental insights that Jacob and Monod had concerning the central role of the control of gene transcription in bacterial physiology and that they expected would apply to cell differentiation and development in complex multicellular organisms have been borne out in many respects in the genetic control of animal development.

Many features in single-celled and multicellular eukaryotes, however, are not found in bacteria and their viruses. Geneticists and molecular biologists have discovered the functions of introns, RNA splicing, distant and multiple cis-acting regulatory elements, chromatin, alternative splicing, and, more recently, miRNAs. Still, central to the genetic control of development is the control of differential gene expression.

This chapter has presented an overview of the logic and mechanisms for the control of gene expression and development in fruit flies and a few other model species. We have concentrated on the toolkit of animal genes for developmental processes and the mechanisms that control the organization of major features of the body plan—the establishment of body axes, segmentation, and segment identity. Although we explored only a modest number of regulatory mechanisms in depth, and just a few species, similarities in regulatory logic and mechanisms allow us to identify some general themes concerning the genetic control of development.

1. *Despite vast differences in appearance and anatomy, animals have in common a conserved toolkit of genes that govern development.* This toolkit is a small fraction of all genes in the genome, and most of these toolkit genes control transcription factors and components of signal-transduction pathways. Individual toolkit genes typically have multiple functions and affect the development of different structures at different stages.
2. *The development of the growing embryo and its body parts takes place in a spatially and temporally ordered progression.* Domains within the embryo are established by the expression of toolkit genes that mark out progressively finer subdivisions along both embryonic axes.

3. *Spatially restricted patterns of gene expression are products of combinatorial regulation.*

Each pattern of gene expression has a preceding causal basis. New patterns are generated by the combined inputs of preceding patterns. In the examples presented in this chapter, the positioning of pair-rule stripes and the restriction of appendage-regulatory-gene expression to individual segments requires the integration of numerous positive and negative regulatory inputs by *cis*-acting regulatory elements.

Post-transcriptional regulation at the RNA level adds another layer of specificity to the control of gene expression. Alternative RNA splicing and translational control by proteins and miRNAs also contribute to the spatial and temporal control of toolkit-gene expression.

Combinatorial control is key to both the *specificity* and the *diversity* of gene expression and toolkit-gene function. In regard to specificity, combinatorial mechanisms provide the means to localize gene expression to discrete cell populations by using inputs that are not specific to cell type or tissue type. The actions of toolkit proteins can thus be quite specific in different contexts. In regard to diversity, combinatorial mechanisms provide the means to generate a virtually limitless variety of gene-expression patterns.

4. *The modularity of cis-acting regulatory elements allows for independent spatial and temporal control of toolkit-gene expression and function.*

Just as the operators and UAS elements of bacteria and simple eukaryotes act as switches in the physiological control of gene expression, the *cis*-acting regulatory elements of toolkit genes act as switches in the developmental control of gene expression. The distinguishing feature of toolkit genes is the typical presence of numerous independent *cis*-acting regulatory elements that govern gene expression in different spatial domains and at different stages of development. The independent spatial and temporal regulation of gene expression enables individual toolkit genes to have different but specific functions in different contexts. In this light, it is not adequate or accurate to describe a given toolkit-gene function solely in relation to the protein (or miRNA) that it encodes because the function of the gene product almost always depends on the context in which it is expressed.

KEY TERMS

enhancer

gain-of-function mutation

gap gene

gene complex

genetic screen

genetic toolkitheterochronic genehomeoboxhomeodomainhomeotic transformationhousekeeping geneHox geneloss-of-function mutationmaternal-effect genemorphogenorganizerpair-rule genepositional informationsegment-polarity geneserially reiterated structurezygote

SOLVED PROBLEMS

SOLVED PROBLEM 1

The *Bicoid* gene (*bcd*) is a maternal-effect gene required for the development of the *Drosophila* anterior region. A mother heterozygous for a *bcd* deletion has only one copy of the *bcd* gene. With the use of *P* elements to insert copies of the cloned *bcd*⁺ gene into the genome by transformation, it is possible to produce mothers with extra copies of the gene. The early *Drosophila* embryo develops an indentation called the cephalic furrow that is more or less perpendicular to the longitudinal, anteroposterior (A–P) body axis. In the progeny of mothers with only a single copy of *bcd*⁺, this furrow is very close to the anterior tip, lying at a position one-sixth of the distance from the anterior to the posterior tip. In the progeny of standard wild-type diploids (having two copies of *bcd*⁺), the cephalic furrow arises more posteriorly, at a position one-fifth of the distance from the anterior to the posterior tip of the embryo. In the progeny of mothers with three copies of *bcd*⁺, it is even more posterior. As additional gene doses are added, the cephalic furrow moves more and more posteriorly, until, in the progeny of mothers with six copies of *bcd*⁺, it is midway along the A–P axis of the embryo. Explain the gene-dosage effect of *bcd*⁺ on the formation of the cephalic furrow in light of the contribution that *bcd*⁺ makes to A–P pattern formation.

SOLUTION

The determination of anterior–posterior parts of the embryo is governed by a concentration gradient of Bicoid protein, which is therefore a morphogen. The furrow develops at a critical concentration of *bcd*. As *bcd*⁺ gene dosage (and, therefore, Bicoid protein concentration) decreases, the furrow shifts anteriorly; as the gene dosage increases, the furrow shifts posteriorly.

PROBLEMS

Visit SaplingPlus for supplemental content. Problems with the  icon are available for review/grading.

WORKING WITH THE FIGURES

(The first 16 questions require inspection of text figures.)

1. In [Figure 13-2](#), the transplantation of certain regions of embryonic tissue induces the development of structures in new places. What are these special regions called, and what are the substances they are proposed to produce?
2. In [Figure 13-5](#), two different methods are illustrated for visualizing gene expression in developing animals. Which method would allow one to detect where within a cell a protein is localized?
3. [Figure 13-7](#) illustrates the expression of the Ultrabithorax (*Ubx*) Hox protein in developing flight appendages. What is the relationship between where the protein is expressed and the phenotype resulting from the loss of its expression (shown in [Figure 13-1](#))?
4. Why might there be more differences among the sequences of all the Hox proteins within *Drosophila* (shown in [Figure 13-8](#)) than there are among the sequences of the Hox group 4 proteins in *Drosophila* and different vertebrate species (shown in [Figure 13-9](#))?
5. In [Figure 13-11](#), what is the evidence that vertebrate *Hox* genes govern the identity of serially repeated structures?
6. As shown in [Figure 13-14](#), what is the fundamental distinction between a pair-rule gene and a segment-polarity gene?
7. In [Table 13-1](#), what is the most common function of proteins that contribute to pattern formation? Why is this the case?

8. Based on the information provided in [Figure 13-17](#) and [Figure 13-19](#), do you predict that there are many or few Bicoid-binding sites in the regulatory elements that control expression of the *Giant* gene?
9. In [Figure 13-19](#), which gap protein regulates the posterior boundary of *eve* stripe 2? Describe how it does so in molecular terms.
10. In [Figure 13-20](#), the *Ultrabithorax* (*Ubx*) gene is expressed in abdominal segments one through seven, and the *Distal-less* (*Dll*) gene is expressed in the head and thoracic segments. What do you predict would happen to *Dll* expression if *Ubx* were expressed in thoracic segments one through three?
11. [Figure 13-21](#) shows a cis-acting regulatory element of the *Distal-less* (*Dll*) gene.
 - a. How many different transcription factors govern where the *Dll* gene will be expressed?
 - b. Are there any combinations of mutations that would lead to expression of the *Dll* gene in abdominal segment 8?
12. Examine the *Drosophila* sex determination cascade shown in [Figure 13-22](#).
 - a. Which isoform of the *doublesex* transcript would be found in males that express the Sex-lethal protein?
 - b. Which isoform of the *doublesex* transcript would be found in males that have a loss-of-function mutation in the *Sex-lethal* gene?
 - c. Which isoform of the *doublesex* transcript would be found in females that have a loss-of-function mutation in the *Sex-lethal* gene?
13. What do you predict would happen to expression of the *lacZ* reporter gene in [Figure 13-23](#) if the GLD-1 protein was expressed in all four cells of the early *C. elegans* embryo?
14. In [Figure 13-24](#), we see that overexpression of the *let-7* gene has the same phenotype as a loss-of-function mutation in the *lin-41* gene. Explain this result based on the molecular function of the *let-7* gene.
15. As shown in [Figure 13-26](#), the *Sonic hedgehog* gene is expressed in many places in a developing chicken. Is the identical Sonic hedgehog protein expressed in each tissue? If so, how do the tissues develop into different structures? If not, how are different Sonic hedgehog proteins produced?
16. Mutations in a cis-acting regulatory element of the *Sonic hedgehog* gene lead to polydactyly in humans, as shown in [Figure 13-28](#). Based on [Figure 13-27](#), where do you think the *Sonic hedgehog* gene is expressed in humans with this mutation during limb development?

BASIC PROBLEMS

17. *Engrailed*, *even-skipped*, *hunchback*, and *Antennapedia*. To a *Drosophila* geneticist, what are they? How do they differ? 
18. Describe the expression pattern of the *Drosophila* gene *eve* in the early embryo, and the phenotypic effects of mutations in the *eve* gene. 
19. Contrast the function of homeotic genes with that of pair-rule genes. 
20. When an embryo is homozygous mutant for the gap gene *Kr*, the fourth and fifth stripes of the pair-rule gene *ftz* (counting from the anterior end) do not form normally. When the gap gene *kni* is mutant, the fifth and sixth *ftz* stripes do not form normally. Explain these results in regard to how segment number is established in the embryo.
21. Some of the mammalian *Hox* genes have been shown to be more similar to one of the insect *Hox* genes than to the others. Design an experimental approach that would enable you to demonstrate this finding in a functional test in living flies.
22. The three homeodomain proteins *Abd-B*, *Abd-A*, and *Ubx* are encoded by genes within the *Bithorax* complex of *Drosophila*. In wild-type embryos, the *Abd-B* gene is expressed in the posterior abdominal segments, *Abd-A* in the middle abdominal segments, and *Ubx* in the anterior abdominal and posterior thoracic segments. When the *Abd-B* gene is deleted, *Abd-A* is expressed in both the middle and the posterior abdominal segments. When *Abd-A* is deleted, *Ubx* is expressed in the posterior thorax and in the anterior and middle abdominal segments. When *Ubx* is deleted, the patterns of *Abd-A* and *Abd-B* expression are unchanged from wild type. When both *Abd-A* and *Abd-B* are deleted, *Ubx* is expressed in all segments from the posterior thorax to the posterior end of the embryo. Explain these observations, taking into consideration the fact that the gap genes control the initial expression patterns of the homeotic genes.
23. What genetic tests allow you to tell if a gene is required zygotically or if it has a maternal effect? 
24. In considering the formation of the A–P and D–V axes in *Drosophila*, we noted that, for mutations such as *bcd*, homozygous mutant mothers uniformly produce mutant offspring with segmentation defects. This outcome is always true regardless of whether the offspring themselves are *bcd*⁺/*bcd* or *bcd/bcd*. Some other maternal-effect lethal mutations are different, in that the mutant phenotype can be “rescued” by introducing a wild-type allele of the gene from the father. In other words, for such rescuable maternal-effect lethals, *mut*⁺/*mut* animals are normal, whereas *mut/mut* animals have the mutant defect. Explain the difference between rescuable and nonrescuable maternal-effect lethal mutations.

25. Suppose you isolate a mutation affecting A–P patterning of the *Drosophila* embryo in which every other segment of the developing mutant larva is missing.
- Would you consider this mutation to be a mutation in a gap gene, a pair-rule gene, a segment-polarity gene, or a segment-identity gene?
 - You have cloned a piece of DNA that contains four genes. How could you use the spatial-expression pattern of their mRNA in a wild-type embryo to identify which represents a candidate gene for the mutation described?
 - Assume that you have identified the candidate gene. If you now examine the spatial-expression pattern of its mRNA in an embryo that is homozygous mutant for the gap gene *Krüppel*, would you expect to see a normal expression pattern? Explain.
26. In an embryo from a homozygous *Bicoid* mutant female, which class(es) of gene expression is (are) abnormal?
- Gap genes
 - Pair-rule genes
 - Segment-polarity genes
 - Hox* genes
 - All answer options are correct.
27. The Hunchback protein is normally expressed in the anterior half of the *Drosophila* embryo. You find a mutation in the 3' untranslated region of the *hunchback* gene that results in expression of Hunchback protein throughout the entire embryo. Provide a molecular explanation for this result.
28. During mouse development, a homolog of the *Drosophila wingless* gene called *Wnt7a* is expressed in the developing limbs and the female reproductive tract. What phenotypes would be predicted to occur in mice with a mutation in the coding region of *Wnt7a*?
29. Mutations in the *Wnt7a* gene have been associated with a human syndrome in which there are abnormalities of both the limbs and genitalia. Do you predict that these mutations are in the coding sequence or in a cis-acting regulatory element of the *Wnt7a* gene?

CHALLENGING PROBLEMS

30. Which of the proteins involved in *Drosophila* development can be classified as a morphogen?
31. You are interested in the genes that control the development of the eyes in *Drosophila*.

- a. Outline the steps you would take to identify and characterize these genes.
 - b. Outline the steps you would take to determine whether the genes you find in *Drosophila* are found in other species.
 - c. Summarize the advantages of your experimental approaches.
32. The *eyeless* gene is required for eye formation in *Drosophila*. It encodes a homeodomain.
- a. What would you predict about the biochemical function of the Eyeless protein?
 - b. Where would you predict that the *eyeless* gene is expressed in development? How would you test your prediction?
 - c. The *Small eye* and *Aniridia* genes of mice and humans, respectively, encode proteins with very strong sequence similarity to the fly Eyeless protein, and they are named for their effects on eye development. Devise one test to examine whether the mouse and human genes are functionally equivalent to the fly *eyeless* gene.
33. Gene *X* is expressed in the developing brain, heart, and lungs of mice. Mutations that selectively affect gene *X* function in these three tissues map to three different regions (A, B, and C, respectively) 5' of the *X* coding region.
- a. Explain the nature of these mutations.
 - b. Draw a map of the *X* locus consistent with the preceding information.
 - c. How would you test the function of the A, B, and C regions?
34. Why are regulatory mutations at the mouse *Sonic hedgehog* gene dominant and viable? Why do coding mutations cause more widespread defects?
35. A mutation occurs in the *Drosophila doublesex* gene that prevents Tra from binding to the *dsx* RNA transcript. What would be the consequences of this mutation for Dsx protein expression in males? In females?
36. You isolate a *glp-1* mutation of *C. elegans* and discover that the DNA region encoding the spatial control region (SCR) has been deleted. What will the GLP-1 protein expression pattern be in a four-cell embryo in mutant heterozygotes? In mutant homozygotes? 
37. Assess the validity of Monod and Jacob's remark that "anything found to be true of *E. coli* must also be true of Elephants."
- a. Compare the structures and mechanisms of action of animal Hox proteins and the Lac repressor. In what ways are they similar? In what ways are they different?

- b. Compare the structure and function of the *lac* operator with the even-skipped stripe 2 enhancer (*eve* stripe 2). How is the control of these “genetic switches” similar or different? 

GENETICS AND SOCIETY

Justify the genetic study of development in model organisms such as *Drosophila* and *C. elegans* to understand human development and disease.