Drosophila melanogaster: Genetic Portrait of the Fruit Fly

Reference **D**

A wild-type *Drosophila melanogaster,* or fruit fly, has multifaceted brick red eyes, a tan thorax studded with arched black bristles, a striped abdomen, and a pair of translucent wings (**Fig. D.1**). The fly's rapid life cycle, low chromosome number, small genome size, and giant salivary gland chromosomes are among the experimental advantages that made *D. melanogaster* the central organism in the study of transmission genetics in the first half of the twentieth century. These same features also made *Drosophila* a major model organism in genetic studies of animal development and behavior in the last quarter of the century.

Modern *Drosophila* genetics originated with Thomas Hunt Morgan's discovery of the sex-linked *white* eye mutation in 1910. In the next few decades, *Drosophila* studies by Morgan and his students at Columbia University in New York City laid the experimental foundation for the field of genetics (**Fig. D.2**). Calvin Bridges, for his Ph.D. with Morgan, proved the chromosome theory of inheritance by showing that aberrations in the inheritance of the sex-linked *white* eye gene correlate perfectly with the nondisjunction of the X chromosome (see Chapter 4 of the main textbook). Alfred H. Sturtevant, as an undergraduate in Morgan's lab, generated the first chromosome map by measuring recombination frequencies among the known sex-linked genes (see Chapter 5 of the main textbook). Herman J. Muller demonstrated the mutagenicity of X rays (see Chapter 7 of the main textbook). These and other members of Morgan's "fly room" went on to establish many other basic features of transmission genetics, including the structure and genetic consequences of each major type of chromosome rearrangement (described in Chapter 14 of the main textbook).

From the 1940s through the 1960s, bacteria and their viruses replaced *Drosophila* as the dominant organisms in genetic studies. The simplicity of these microbes and the ease with which they can be grown in vast numbers made them ideally suited for genetic and biochemical work on the basic structure and function of genes. During this time, *Drosophila* research went into almost total eclipse. Then, in the early 1970s, when many geneticists began to investigate how genes control the development and behavior of higher organisms, interest in *Drosophila* increased dramatically as researchers realized the great potential of combining the sophisticated genetic techniques of *Drosophila* with the newly developed methods of recombinant DNA technology. The successful marriage of these approaches depended in part on the fortunate fact that a handful of *Drosophila* geneticists had continued to work and maintain stock collections through the "dark" years.

In this chapter, we see that by using modern molecular methods in conjunction with the rich legacy of genetic techniques originally developed for *Drosophila,* researchers have made stunning progress in understanding how genes control the development and behavior of higher organisms. The impact of this work has been immense, and the accelerating pace of *Drosophila* genetics suggests that the fruit fly will remain a key model organism for the foreseeable future.

A transgenic adult fruit fly expressing green fluorescent protein (GFP) in the posterior portion of each body segment, seen under normal light (top) *or UV light* (bottom).

Figure D.1 **A colorized scanning electron micrograph of a *Drosophila* adult's head and wings.**

Figure D.2 **The fountainhead of *Drosophila* genetics.** Thomas Hunt Morgan and his students in the fly room at Columbia University, at a party in 1919 celebrating the return of Alfred Sturtevant from World War I military service. Individuals whose work is discussed in this book include Morgan (*back row, far left*), Sturtevant (*front row, third from the right*), Calvin Bridges (*back row, third from the right*), and Herman J. Muller (*back row, second from the left*). The "honored guest" between Muller and Bridges has clearly seen better days.

A major theme that emerges from our discussion of the genetic analysis of *Drosophila* is that many of the developmental genes and mechanisms operating in the fly have been conserved in a wide variety of animals. For example, most of the developmentally important genes first found in *Drosophila* have turned out to play similar roles in other animals, including humans. In fact, researchers identified most of the genes known to influence the development of humans as homologs of genes already discovered in *Drosophila.*

As we examine *D. melanogaster*'s use as a model organism, we present:

- The structure and organization of the genome.
- The life cycle.
- Techniques of genetic analysis, including the construction and use of balancer chromosomes, the uses of P-element transposons, the production of mosaics, ectopic gene expression, targeted gene knockouts, and the *Drosophila* Genome Project.
- The genetic analysis of body plan development in *Drosophila,* a comprehensive example.

D.1 Structure and Organization of the *Drosophila* Genome

The *Drosophila* genome is about 5% the size of the human genome and is packaged in far fewer chromosomes. It contains roughly 13,600 genes, about a half of the number found in the human genome. This means that the density of genes on *Drosophila* chromosomes is higher than that of genes in the human genome. Even so, about one-third of the fly genome consists of repetitive sequences that do not encode proteins or that act as transposable elements.

The Chromosomes of *Drosophila*

The haploid genome of *D. melanogaster* contains four chromosomes, designated numerically as 1–4. Chromosome 1 is

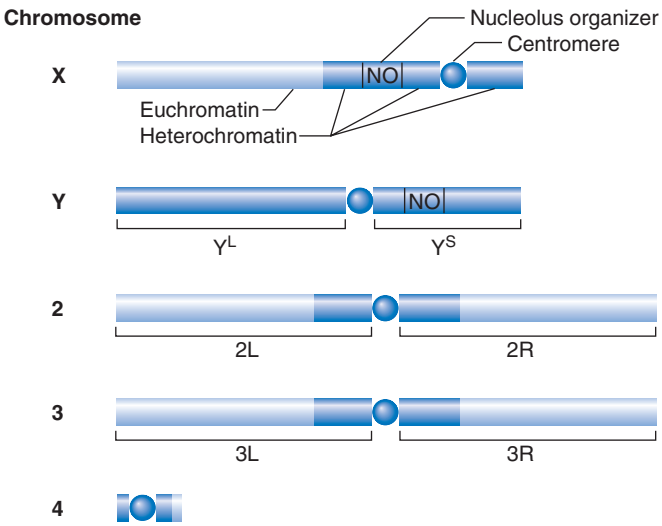


Figure D.3 **The *Drosophila* chromosome complement.** Major blocks of all chromosomes, particularly in regions surrounding the centromeres, are heterochromatic. Key: *dark blue,* heterochromatin; *light blue,* euchromatin; Y^L and Y^S, the long and short arms of the Y chromosome; and 2L, 2R, 3L, and 3R, the left (L) and right (R) arms of chromosomes 2 and 3. See the Fast Forward box in Chapter 4 of the main textbook for a picture of these same chromosomes in mitotic cells.

the X chromosome; chromosomes 2–4 are autosomes. Sex determination, as we saw in Chapter 4 of the main textbook, is of the XY type, with females being XX and males XY. Unlike the situation in mammals, however, the Y plays no role in sex determination; sex is instead determined solely by the ratio of the number of X chromosomes to the number of copies of each autosome (the X:A ratio). Thus, XXY zygotes, which have an X:A ratio of 1 (two copies of the X : two copies of each autosome), develop as normal females, while XO zygotes, which have an X:A ratio of 0.5, develop as morphologically normal, but sterile, males. (Review Chapters 4 and 18 of the main textbook for more details about sex determination in *Drosophila.*) The Y chromosome, which by convention is not given a number, is required only to confer male fertility. The low chromosome number is one of *Drosophila*'s key advantages for genetic studies, as it simplifies most genetic manipulations.

Figure D.3 shows the structure of *Drosophila* chromosomes prior to S phase DNA replication. The X chromosome is large and acrocentric; chromosomes 2 and 3 are large and metacentric, and chromosome 4 is a tiny acrocentric "dot" chromosome that is only about 2% the size of the major autosomes. Each chromosome contains a large block of heterochromatin near its centromere. Although a few genes are present in these heterochromatic blocks, for the most part, these regions contain highly repetitive DNA sequences and are transcriptionally silent. Centromeric heterochromatin comprises about a quarter of the length of the X chromosome and each of the large autosomes, as well as the majority of the small chromosome 4. The large submetacentric Y chromosome is almost entirely heterochromatic, and it carries only a few (probably seven or eight) genes, which function in the male germ line. Nucleolus organizer regions are located within the short arm of the Y and the centromeric heterochromatin of the X. Each nucleolus organizer region carries 200–250 copies of the 18S and 28S ribosomal RNA genes in tandem array.

The Giant Polytene Chromosomes of the Larval Salivary Gland Are Key Tools for *Drosophila* Genetics

As we saw in Chapter 14 of the main textbook, chromosomes in the fruit fly's salivary glands replicate 10–11 times without cell division. The 1024 or 2048 sister chromatids generated during these rounds of replication stay intimately associated with one another in perfect lateral register, producing a many-stranded polytene chromosome. These chromosomes are truly gigantic: Polytene chromosomes are about as wide as normal metaphase chromosomes are long.

Regions of local condensation along a chromosome become amplified laterally in polytene chromosomes and are visible as distinct crossbands (separated by less condensed interbands) when viewed through a light microscope. A striking feature seen in cells containing polytene chromosomes is that homologous chromosomes are tightly synapsed with one another (review Fig. 14.6 in the main textbook). Such synapsis is found in all somatic cells of *Drosophila*, including diploid cells. Why such somatic synapsis occurs in *Drosophila* is a mystery; in other higher organisms extensive synapsis of homologs occurs only in prophase I of meiosis. In contrast to the euchromatin, regions of centromeric heterochromatin are not replicated during polytenization, and they coalesce into a common region called the *chromocenter* (also shown in Fig. 14.6). As a result, in squash preparations of salivary gland nuclei, five major arms (the left and right arms of chromosomes 2 and 3, and the X chromosome) radiate out from the chromocenter. The much smaller chromosome 4 also associates with the chromocenter, but it is often hard to see. The Y chromosome is completely heterochromatic and therefore is not visible.

The importance of polytene chromosomes in *Drosophila* genetics cannot be overemphasized. As we saw in Chapter 14, they have provided *Drosophila* geneticists with a ready-made detailed physical map of the fly genome. Since the finer visible bands contain only a few kilobases of DNA per chromatid, rearrangement breakpoints can often be mapped with very high precision simply by microscopic examination. Cloned DNA sequences can also be mapped very rapidly and precisely by *in situ* hybridization to polytene chromosomes (see Fig. 14.9 in the main textbook). This technique was of particular utility when the *Drosophila* Genome Project determined the sequence of the fruit fly genome. Researchers could determine the location and extent of the large genomic clones (in YAC, BAC, or P1 vectors), whose sequences they determined in the genome project, by using the clones as probes for *in situ* hybridization to polytene chromosomes.

The *Drosophila* Genome

The haploid genome contains about 170,000 kb of DNA. Roughly 21% of this DNA consists of highly repetitive "satellite" DNA that is concentrated in the centromeric heterochromatin and the Y chromosome. Another 3% consists of repeated genes that encode rRNA, 5S RNA, and the histone proteins. Approximately 9% is composed of around 50 families of transposable elements, which are roughly 2–9 kb in length; the number of copies of each type of element varies from strain to strain but is usually somewhere between 10 and 100. In addition to residing at scattered sites within the euchromatin, many transposable elements are found in the centromeric heterochromatin, while some are concentrated in telomeres. An unusual feature of *Drosophila* chromosomes is that the telomeres do not consist of simple repeats, such as the TTGGGG or TTAGGG repeats found at the ends of chromosomes in other well-characterized organisms. Instead, it appears that the rapid jumping into the telomeric region of certain kinds of transposable elements maintains the telomeres. These insertions into the telomeric region of the transposable elements counter the chromosomal shortening that inevitably occurs during the replication of the ends of linear DNA molecules.

The remaining 67% of the genome consists of unique DNA sequences that reside mostly in the euchromatic arms. We describe the nature and organization of the approximately 13,600 genes present within these unique DNA sequences later in the chapter, in our discussion of the *Drosophila* Genome Project.

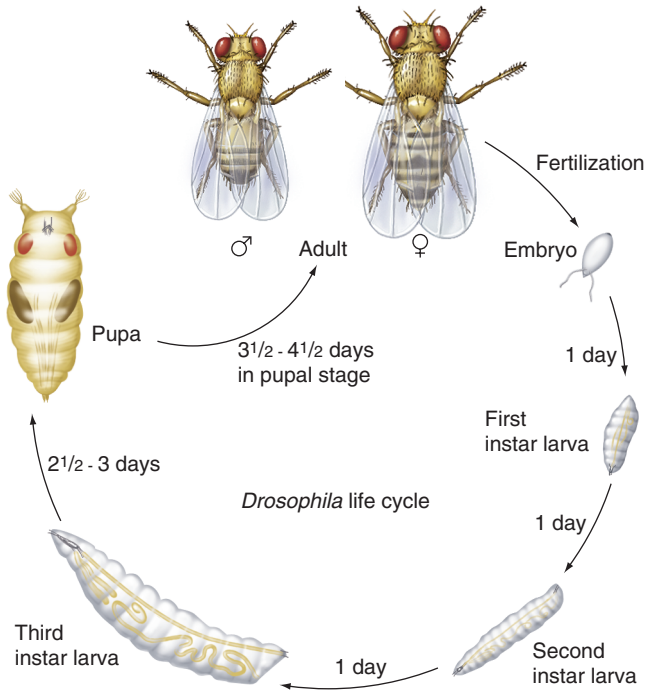


Figure D.4 The *Drosophila* life cycle. The transition from an embryo to a first instar larva is called *hatching.* The transitions between larval instars are *molts.* The process that converts a third instar larva to a pupa is *pupariation.* Emergence of the adult from the pupal case is called *eclosion.*

D.2 Life Cycle

Fruit flies can be raised on a wide variety of media, but the standard formulation is a mixture of corn meal, brewer's yeast, sugar, agar, and a small amount of mold inhibitor. Flies bred on such a medium at 25°C complete their life cycle from fertilization to emergence of the adult fly in 10 days (**Fig. D.4**).

Drosophila adults have a very high reproductive capacity. A female can produce 3000 progeny in her lifetime; a single male can sire well over 10,000 offspring. The female has special sperm storage organs (the spermatheca and seminal receptacle) that allow her to produce several hundred progeny after a single mating. Each *Drosophila* egg is about half a millimeter long and is well supplied with yolk that supports embryonic development. Fertilization normally occurs as the egg is being laid.

The embryo completes its development in just under 24 hours, and some of the early events happen so rapidly they can be watched in real time. For example, in the early stages of embryonic development, the zygotic nuclei divide every 10 minutes, and the morphogenic movements of gastrulation that produce the endoderm, mesoderm, and ectoderm tissue layers take only 20 minutes. Embryogenesis ends with the hatching of a wormlike first instar (or first-stage) larva that is specialized for feeding and grows dramatically. To allow for an increase in body size, the larva molts 24 and 48 hours after hatching to produce second and third instar larvae.

About three days after the second molt, third instar larvae complete their growth, crawl out of the food, and pupate. The timing of each larval molt and pupation is controlled by pulses of a steroid molting hormone called **ecdysone.** How growth is regulated and how ecdysone controls molting are currently areas of very active investigation. Once inside the protective pupal case, the larvae undergo **metamorphosis:** a dramatic reorganization of the fly body plan. This reorganization takes about four days and consists of the disintegration of most larval tissues and their replacement through the proliferation and differentiation of cells that produce adult structures. Most structures specific to adults, such as the wings, legs, eyes, and genitalia, are generated from **imaginal discs:** flattened epithelial sacs that develop from small groups of cells set aside in the early embryo (**Fig. D.5**). Imaginal discs undergo extensive growth and development during the larval and pupal stages. Other adult-forming cells, clustered in nests within each abdominal segment, produce the adult abdominal epidermis. At the end of metamorphosis, when the adult

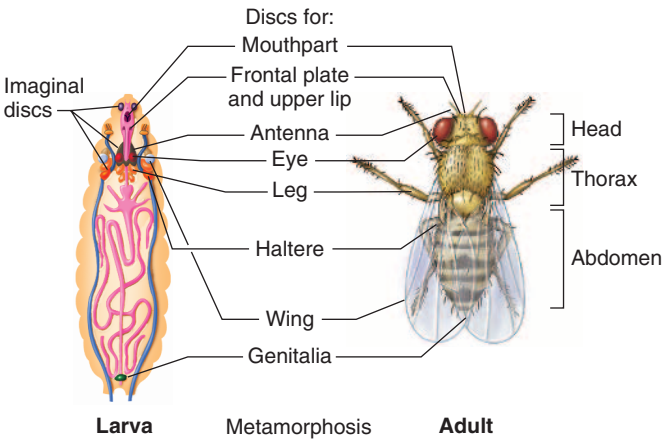


Figure D.5 Many adult structures develop from imaginal discs in larvae and pupae. Lines connect particular imaginal discs in the larva with the structures they generate in the adult.

emerges from the pupal case, its wings expand and the entire exoskeleton hardens and becomes pigmented.

D.3 Techniques of Genetic Analysis

The practice of genetics in *Drosophila* has a unique character that in part reflects the tools created by *Drosophila* researchers over many years and in part reflects certain distinctive features of fruit fly inheritance. For a review of the symbols used by *Drosophila* geneticists, see the *Guidelines for Gene Nomenclature* directly following Chapter 22 of the main textbook. Of particular importance to this chapter, the names or abbreviations of genes are always italicized, while the names or abbreviations of proteins are not italicized and begin with a capital letter.

In *Drosophila,* Crossing-Over Occurs Only in Females

Although the reason for the restriction to females of crossing-over is not clear, in the few other known species in which recombination is restricted to one sex, it is always the *heterogametic sex* (that is, the gender characterized by dissimilar sex chromosomes) that lacks crossing-over. This suggests that the prevention of crossing-over between different sex chromosomes has been of selective importance during evolution.

The Absence of Crossing-Over in the Male Has Considerable Technical Significance

The absence of crossing-over in the male makes it possible to maintain linkage relationships simply by ensuring that a chromosome is inherited strictly through the male parent. This may seem trivial, but it greatly simplifies many genetic manipulations and allows a variety of selective genetic screens.

In Females, Only a Moderate Amount of Crossing-Over Occurs

Typically, only one or two crossovers take place in each major chromosome arm per meiosis, and essentially no crossing-over occurs in the centromeric heterochromatin or anywhere in the fourth chromosome. As in other organisms, however, crossing-over is required for the accurate disjunction of the major chromosomes during meiosis I in females. In the absence of chiasmata, bivalents fall apart at the end of prophase I, causing nondisjunction and chromosome loss. A special mechanism, currently under active study, ensures the segregation of the fourth chromosomes and of the occasional large chromosome bivalents in which crossing-over fails to occur. How the segregation of homologs is directed in males is not well understood, but it appears to depend on special sites along the chromosomes that mediate the association of homologs.

Balancer Chromosomes Help Preserve Linkage

Drosophila geneticists have used X rays as a mutagen to create many special chromosome types useful for genetic manipulations. By far the most important of these is the **balancer chromosome,** which carries multiple, overlapping inversions (**Fig. D.6a**). Most balancers also contain a dominant marker that enables researchers to follow the balancer through crosses, and a recessive lethal mutation that prevents the survival of homozygotes. The utility of balancer chromosomes is that, like all inversion-containing chromosomes, they prevent the recovery of crossovers between normal sequence chromosomes and themselves. As a result, both balancer chromosomes and their homologs are inherited through crosses as intact units, and no recombinant chromosomes are passed on to the next generation.

Balancer Chromosomes Are Useful in Stock Keeping

The term "balancer" derives from the extensive use of these chromosomes in stock keeping. Experimenters establish a typical *Drosophila* stock by crossing males and females that carry a desired mutation heterozygous with a balancer chromosome homolog. If the mutation is a recessive lethal, as is often the case, and the balancer chromosome carries a different recessive lethal, the only surviving progeny will, like their parents, be heterozygous for the chromosome bearing the desired mutation and the balancer (**Fig. D.6b**).

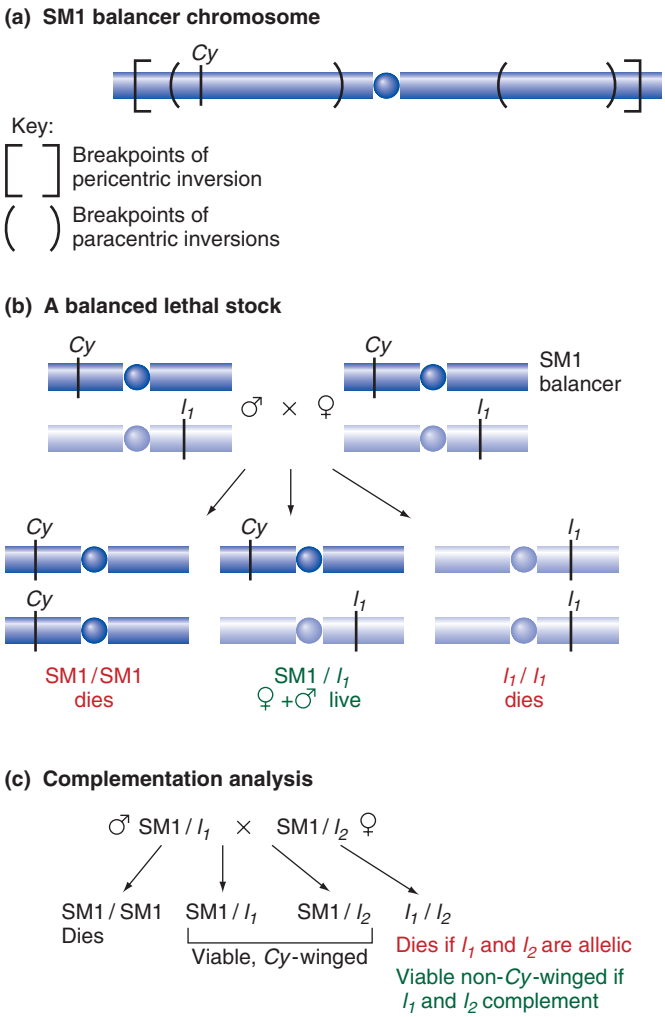


Figure D.6 **Balancer chromosomes in *Drosophila*.** **(a)** SM1, a chromosome 2 balancer, contains overlapping inversions to suppress recombination, and it also carries the dominant marker *Cy* (Curly wings). *Cy* is also a recessive lethal, so that SM1/SM1 homozygotes do not survive. **(b)** When flies heterozygous for a lethal mutation (l_1) and a balancer chromosome are intercrossed, the only adult survivors are l_1/Balancer heterozygotes identical to their parents. **(c)** Are two independently isolated lethal mutations (l_1 and l_2) allelic? The diagrammed cross will produce l_1/l_2 progeny only if the mutations are in different genes.

The only requirement for maintaining the mutation is to transfer progeny to fresh medium periodically. Stocks of this type are said to be "balanced lethals," hence the name "balancer."

Balancer Chromosomes Are Also Useful in Many Genetic Experiments

Balancer chromosomes play a dominant role in *Drosophila* genetics, and are utilized in almost every genetic manipulation. As just one example, consider a complementation test between two independently isolated recessive lethal mutations (*lethal*$_1$ [l_1] and *lethal*$_2$ [l_2]) located on the second chromosome (**Figure D.6c**). The cross involves a second chromosome balancer, designated SM1, which carries the dominant wing mutation *Curly* (*Cy*). Males of genotype l_1/SM1 are crossed to l_2/SM1 females. If the mutations are allelic, only progeny with the Curly wing phenotype will emerge, whereas if the mutations complement each other, non-Curly (l_1/l_2) progeny will also appear.

P-Element Transposons Are Critical Tools in *Drosophila* Molecular Genetics

The discovery of P elements arose from research on a phenomenon called **hybrid dysgenesis.** P elements (**Fig. D.7a**) are the causative agent of hybrid dysgenesis, which occurs when males from *Drosophila* strains carrying P elements (referred to as P strain males) are crossed with females that lack P elements (known as M strain females). The P elements become highly mobile in the germ line of the F_1 hybrids produced from such crosses (**Fig. D.7b and c**). Chromosome breakage resulting from this high transposon mobility causes reduced fertility in the hybrids; the few progeny of the F_1 flies that do survive carry many new mutations induced by the insertion of P elements at new sites in the genome. Hybrid dysgenesis does not occur in any other combination of M or P males and females, such as in crosses of P males and P females (**Fig. D.7b**).

Molecular details of P-element biology help clarify the hybrid dysgenesis phenomenon. The P-element primary transcript encodes a transposase enzyme that acts on the 31 bp inverted repeats at P-element ends to catalyze transposition. Hybrid dysgenesis obviously does not occur in M × M matings without P elements. The reason hybrid dysgenesis does not take place in crosses involving P females is that the eggs produced by P females contain a repressor that prevents transcription of the transposase gene. The nature of this repressor is still not entirely clear; a protein translated from an alternately spliced P-element mRNA appears to be involved, but there is evidence that an RNAi-like mechanism is also important. In any case, eggs from M strain mothers lack a cytoplasmic repressor of P-element transposition. As a result, P elements in sperm from P fathers transpose rapidly when they fertilize eggs from M mothers, explaining the hybrid dysgenesis seen in crosses of P strain males to M strain females (**Fig. D.7b and c**). In the embryos from dysgenic crosses, transposition occurs only in germ-line cells because a splicing event needed to make transposase mRNA does not take place in somatic cells. The adult progeny of a dysgenic mating thus appear normal, but they are semisterile because their gametes carry many chromosomal aberrations and mutations caused by P-element movement.

The high mobility of P elements during hybrid dysgenesis makes it possible to exploit these elements for two purposes: as transformation vectors and as genetic tags.

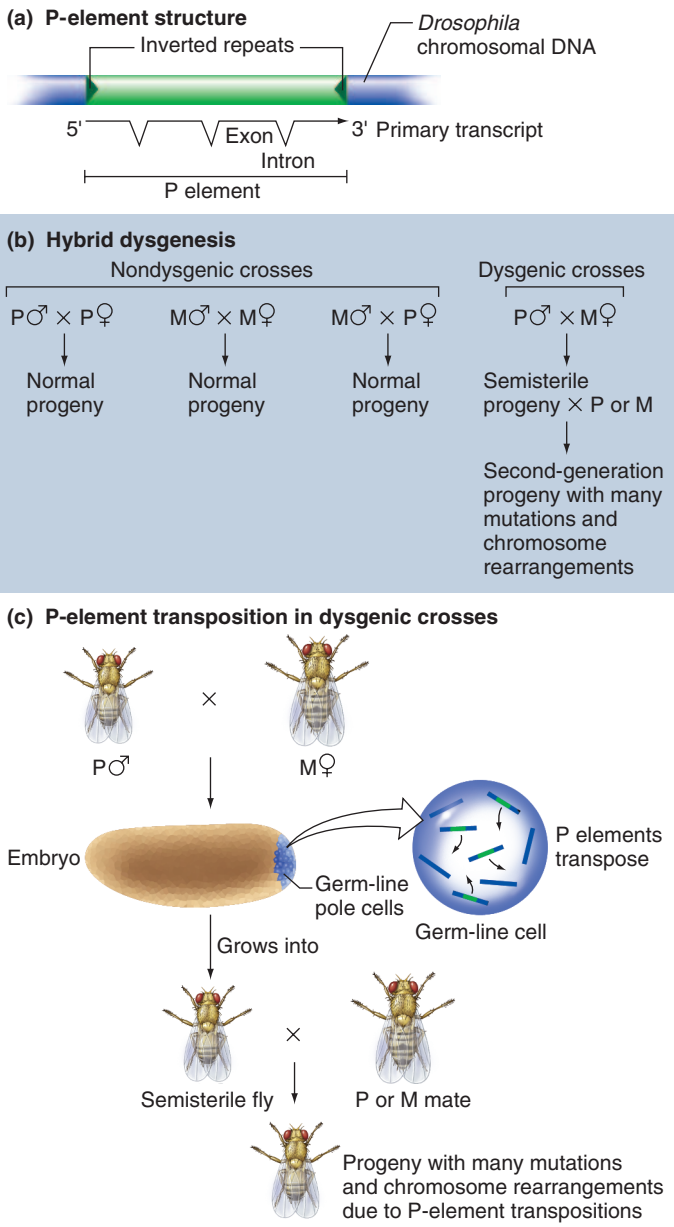


Figure D.7 **P elements and hybrid dysgenesis.** **(a)** Transposition requires the action of the P-element transposase enzyme on the inverted repeats at P-element ends. **(b)** Hybrid dysgenesis occurs only when P males mate with M females. The progeny of such a mating appear wild type but are semisterile. **(c)** P elements on sperm brought into embryos derived from M females transpose at high rates, but only in germ-line cells.

Transformation: The Introduction of Cloned DNA into Fruit Flies

Drosophila geneticists realized from the following experiment that P elements can serve as vectors to introduce cloned DNA into the organism. Researchers first constructed a recombinant plasmid that contained a fragment of *Drosophila* genomic DNA including an intact P element. When they injected this plasmid into early syncytial embryos from M strain mothers, the P element began to transpose and insert itself into the chromosomes of host germ-line cells at a high frequency. This artificial situation mimicked the phenomenon of hybrid dysgenesis. Additional experiments demonstrated that this transposition is catalyzed by the P-element-encoded transposase and requires the presence of the inverted terminal repeats normally found at P-element ends.

For germ-line transformation, *Drosophila* geneticists use vectors that contain the P-element ends but not the P-element gene that encodes the P transposase. They clone novel DNA transgene sequences between the P-element ends and inject the resulting DNA construct into eggs from M strain mothers. A helper P element is coinjected to provide a source of the P-element transposase; this helper cannot itself integrate into a chromosome because it lacks both of the P-element ends. **Figure D.8** illustrates one protocol for using P elements to introduce cloned genes into fruit flies. Note that pCaSpeR, one of the most frequently used transformation vectors, contains a *white*$^+$ marker gene that makes it easy to trace the presence of the vector in the genome by the red eye color it confers.

P-Element Tagging Has Enabled the Cloning of Many Genes of Developmental Interest

Many genes have been cloned by first recovering a mutant allele caused by insertion of a P element, and then isolating DNA sequences flanking the P element. P insertion alleles are usually generated using two special strains of flies. One carries a chromosomal copy of a defective P element that provides a large amount of transposase, but is not itself able to transpose because it lacks normal P-element ends. The other strain carries one or more specially engineered P elements that lack the transposase gene and contain a marker (usually the w^+ eye color gene) that allows the element to be followed in crosses. When the two strains are crossed, transposase produced in the germ cells of the F1 flies catalyzes transposition of the w^+ P element. Depending on the mating scheme used, movements of the marked P element to new sites can be identified by following inheritance of the w^+ marker in the F2 progeny. For example, a common approach is to start with an insertion of a w^+ P element in the X chromosome. When F1 males carrying this insertion and the defective P element that produces transposase are crossed to *w*/*w* females, new autosomal insertions are identified as F2 male progeny having red eyes.

When a P insertion in an interesting gene is found, researchers can rapidly clone DNA from the gene by a technique called *plasmid rescue.* As illustrated in **Fig. D.9a**, this technique requires that the P element contain bacterial plasmid sequences with an antibiotic resistance gene. Scientists cut genomic DNA from the mutant insertion strain with a restriction enzyme, circularize the restriction fragments with DNA ligase, and use the circularized DNA to transform *Escherichia coli.* The only bacterial colonies

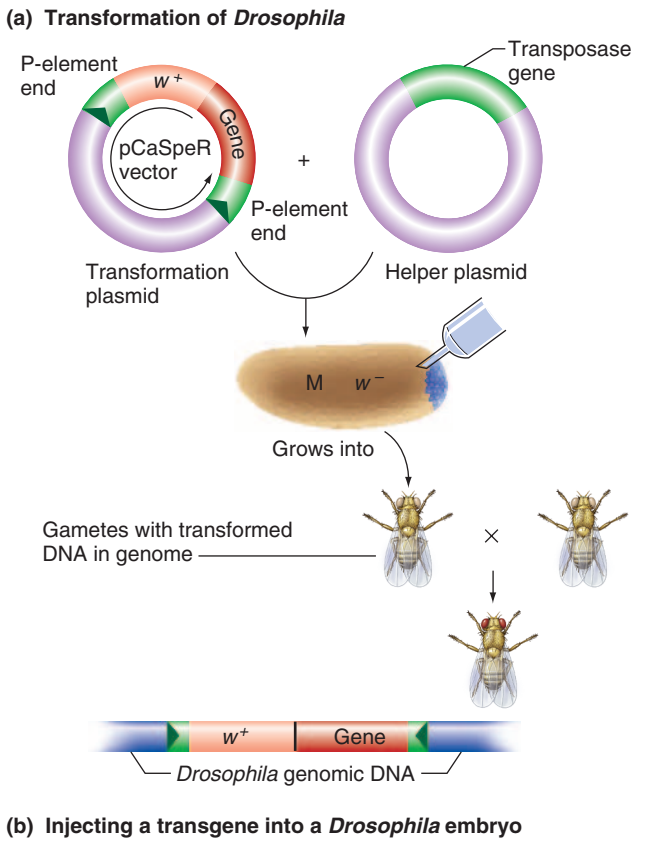


Figure D.8 **P-element-mediated transformation of *Drosophila*.** **(a)** A transgene (Gene) is cloned into a transformation vector (pCaSpeR) containing the *white*$^+$ (w^+) marker gene between the P-element inverted ends. Researchers coinject this construct, along with a helper P element that encodes transposase but has no P-element ends, into M-type w^- fertilized embryos. Transposition of the w^+ gene and the transgene occurs in some germ-line cells. If adults with these germ cells are mated with w^- flies, some progeny will have red (w^+) eyes and an integrated transgene. **(b)** Injection of DNA into the posterior end of a *Drosophila* embryo.

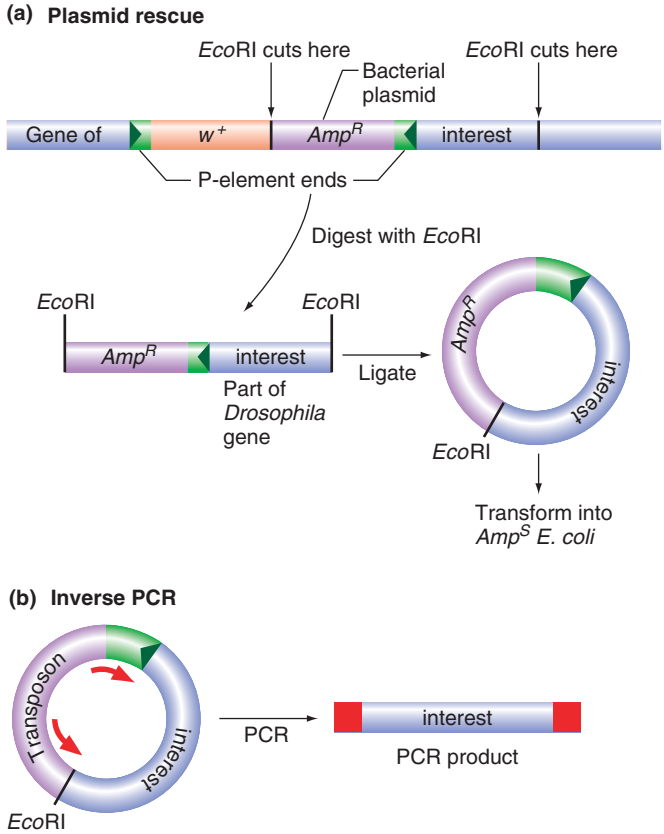


Figure D.9 **Isolation of DNA sequences from a transposon-tagged gene.** **(a)** Plasmid rescue of the gene of interest requires an engineered P element that contains bacterial plasmid DNA including an ampicillin resistance gene (*Amp*^{R}). To identify genes mutated by P-element insertion, researchers digest genomic DNA from mutant flies with an enzyme like *Eco*RI that cuts on one side of the bacterial plasmid but nowhere else in the P element. When the *Eco*RI-cut genomic DNA is circularized by DNA ligase and transformed into *Amp*^{S} bacteria, all ampicillin-resistant colonies should harbor a plasmid that includes *Drosophila* DNA adjacent to the P-element insertion. **(b)** Inverse PCR involves the same digestion and ligation steps to make a circular DNA molecule. PCR using the two primers shown as *red arrows* derived from the transposon DNA sequence amplifies part of the gene of interest.

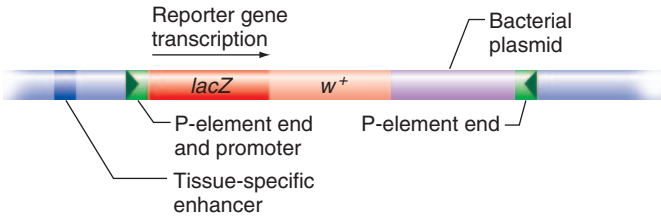
able to grow on antibiotic-containing medium will contain the bacterial plasmid sequences plus the gene-containing *Drosophila* DNA flanking the P-element insertion.

A more recent alternative to plasmid rescue for identifying tagged genes called *inverse PCR* can be used with any marked P element, even if it does not contain a bacterial gene for antibiotic resistance (**Fig. D.9b**). Circularization of a restriction fragment that includes the transposon and adjacent *Drosophila* genomic sequences creates a template for PCR from which DNA corresponding to part of the gene of interest can be amplified.

Enhancer Trapping: Using P Elements to Identify Genes by Their Expression Patterns

Drosophila developmental geneticists have constructed P elements in which the *E. coli lacZ* gene is placed downstream of the P-element promoter (**Fig. D.10a**). Remarkably, when such an element is mobilized, over 65% of the new insertions isolated show patterned expression of the *lacZ* reporter during development, with the *lacZ* gene on in some tissues but off in others. The patterned expression results from the fact that the P-element promoter by itself is

(a) An enhancer trap construct



(b) Detecting tissue-specific enhancers

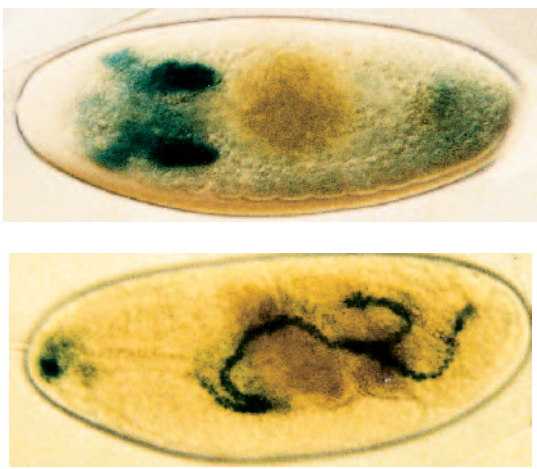


Figure D.10 **Enhancer trapping.** **(a)** This "enhancer trap" contains a *lacZ* reporter gene encoding β-galactosidase downstream of a P element's promoter. Expression of β-galactosidase in transformed animals reflects enhancers near the P-element insertion site. **(b)** Late embryos from two lines of flies with the P-element "enhancer trap" inserted into different genomic locations were stained for β-galactosidase (*blue*). Expression of the reporter is seen in the salivary glands (*top*) or in the Malpighian tubules (*bottom*).

very weak. The promoter can activate *lacZ* transcription strongly only when it comes under the control of regulatory sequences (enhancers) of genes near its new insertion site. (See Chapter 18 of the main textbook for a definition and discussion of enhancers.) Many such insertions into known genes have been recovered, and in most of these cases, the expression pattern of β-galactosidase (the protein product of *lacZ*) accurately reflects the normal expression of the gene into which the element has been inserted (**Fig. D.10b**). The identification of P-element insertion lines with particular β-galactosidase patterns is known as **enhancer trapping.** It is a powerful way to detect genes that are turned on in specific tissues.

A particularly good example of the utility of P-element tagging and enhancer trapping is seen in the work of Thomas Kornberg and his colleagues, who have studied the function of a gene called *engrailed* (*en*), which plays a key role in establishing and patterning body segments. As described in a later section of this chapter, *en* encodes a transcription factor that is expressed in a stripe in the posterior portion of each body segment. To identify target genes regulated by *en,* Kornberg's group screened for enhancer trap insertions that are expressed in the same or a complementary striped pattern. In this way, they identified several key targets of *en.* Among them is a crucially important gene called *hedgehog* (*hh*), which, as explained later, encodes a secreted protein responsible for patterning not only body segments but many other structures as well.

The Production of Genetic Mosaics

Genetic mosaics are individuals composed of cells of more than one genotype. Studies of mosaics allow developmental geneticists to answer many types of questions about the developmental fate of various cells and the functions of particular genes.

The most important method for generating mosaics in flies is mitotic recombination, previously introduced in Chapter 5 of the main textbook. Although crossing over is normally restricted in *Drosophila* to the prophase of meiosis I in females, after X-irradiation of embryos or larvae, crossing over between homologs occurs quite frequently in somatic cells. It is thought that X-ray induced chromosome breakage initiates mitotic recombination, and that the synapsis of homologs in somatic cells of *Drosophila* greatly facilitates this recombination. If mitotic recombination occurs in the G2 portion of the cell cycle, and is followed by appropriate segregation of chromatids in the subsequent mitotic division, the resulting daughter cells will each be homozygous for all loci distal to the crossover (that is, for all genes farther away from the centromere than the site of recombination). The mitotic descendants of these cells will then produce genetically distinct patches in the emerging adults (review Fig. 5.24 of the main textbook).

Although X-rays were used for many years to induce mitotic recombination, this method has now been supplanted by use of a site-specific recombination enzyme from yeast called FLP (**Fig. D.11**). The normal function of FLP is to catalyze reciprocal crossing-over at specific recombination targets (*FRT*s) contained within inverted repeats of a yeast DNA plasmid called the 2 micron circle. Investigators cloned the FLP gene downstream of, and therefore under the control of, a heat-shock promoter whose expression is turned on at elevated temperatures, and they introduced this construct into *Drosophila* by P-element-mediated transformation. To insert the *FRT* target into *Drosophila* chromosomes, they cloned *FRT* into a P element and introduced this second construct into the fly genome by transformation. In this way, they have obtained *FRT* insertions close to the centromere on each chromosome arm of *Drosophila.* If *FRT*s are present at the same location on both homologs, FLP recombinase (whose synthesis is induced by heat shock) catalyzes recombination between them. The resulting frequency of recombinant mitotic clones is extremely high, with several clones on each fly. This technology is so efficient that researchers can now conduct mutant screens in which new mutations are identified by their effects in homozygous patches.

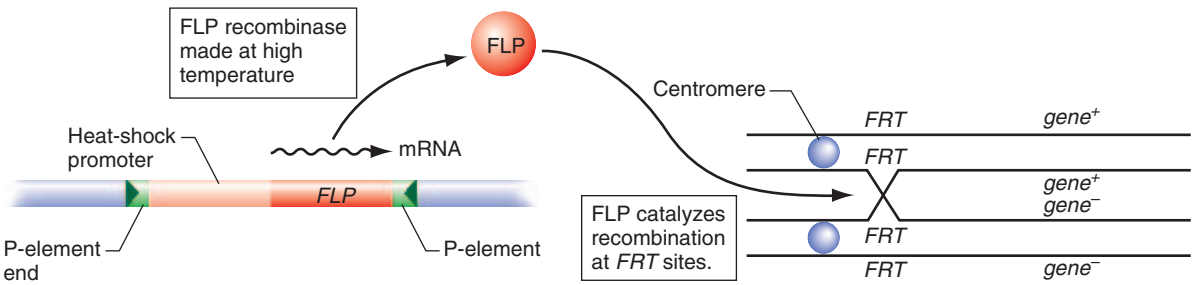


Figure D.11 **Efficient induction of mitotic recombination in *Drosophila*.** When a fly of the genotype shown is subjected to elevated temperature, FLP recombinase is produced. This enzyme catalyzes recombination between *FRT* sites in mitotic cells. Note that the *FRT* sites must be closer to the centromere than the gene for which one wishes to obtain homozygous mutant clones. In this and subsequent figures in this chapter, certain features of modified P elements such as the w^+ gene (needed so investigators can identify transgene-containing flies) are not shown for simplicity.

To recognize clones, several different markers can be used. For the adult cuticle, markers affecting pigmentation (e.g., *yellow*), or bristle morphology (*forked, singed, Kinked, Stubble*) are frequently employed. Much of the adult body is covered with fine hairs called trichomes; mutations that affect these hairs [e.g., *multiple wing hairs* (**Fig. D.12**)] have been used extensively as markers. Another marker that *Drosophila* researchers have recently employed in many experiments is **green fluorescent protein** (**GFP**) from jellyfish (see Fig. 20.13 of the main textbook and the photo on page 75 of this chapter). A GFP gene driven by a promoter expressed in all cells (the ubiquitin promoter) has been introduced into flies by P-element transformation. In heterozygotes for this element, mitotic exchange will produce two daughter cells, one lacking the element and one containing two copies of (i.e., homozygous for) the element. Clones developing from the former type of cell will lack GFP fluorescence, whereas clones derived from GFP homozygous cells will be recognizable by their brighter fluorescence than the surrounding heterozygous cells.

We discuss here three of the many important applications of genetic mosaics in the analysis of *Drosophila* development.

Testing the Developmental Potential of Cells

Clonal analysis has been used widely to determine the range of structures to which a cell and its mitotic descendants are able to contribute. The logic is as follows: Once the potential of a cell is restricted to a specific tissue or region, then

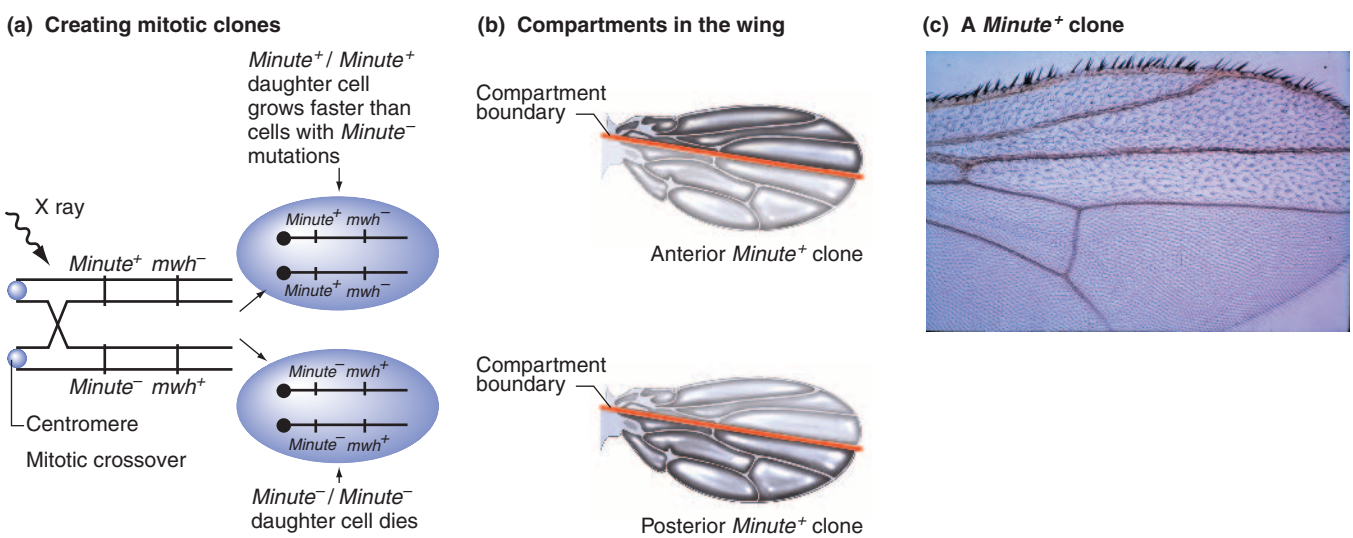


Figure D.12 **Mosaic analysis demonstrates the existence of developmental compartments.** **(a)** X-ray-induced mitotic crossovers in cells with the genotype at the *left* can generate daughter cells homozygous for $Minute^+$ and mwh^-. These $Minute^+$ cells and their mitotic descendants will outgrow all other cells, and they are marked by *multiple wing hairs*$^-$ (mwh^-) which causes each wing cell to produce several tiny hairs instead of the normal one. The homozygous $Minute^-$ daughter cells die. **(b)** Large $Minute^+$ clones (*dark gray*) can fill the anterior or the posterior half of the wing, but they will not cross the compartment boundary. **(c)** Photograph of an anterior clone; the wing cells in the clone are large (because they are $Minute^+$) and have several small hairs (as they are homozygous for mwh^-).

all of the mitotic descendants of that cell will contribute only that one tissue or region. In contrast, if a cell has not undergone such a restriction, then its mitotic descendants will have the ability to contribute to two or more different tissues or regions of the body.

To test the developmental potential of cells, geneticists have devised a method to give clones a growth advantage over neighbouring cells by employing dominant mutations known as *Minutes.* There are about 50 *Minute* genes scattered throughout the genome; those that have been studied in detail turn out to encode ribosomal proteins. Heterozygotes for mutations in these genes show developmental delay and reduced bristle size, presumably because the *Minute* mutations cause a reduced rate of protein synthesis. Investigators use mitotic recombination to generate $Minute^+$ / $Minute^+$ cells in a $Minute^+$ / $Minute^-$ background (**Fig. D.12**). These wild-type (+ / +) cells have a strong growth advantage over their neighbors and produce very large clones. Remarkably, these large $Minute^+$ clones can fill either the anterior half of the wing or the posterior half, but they never cross a boundary line separating the two. These two parts of the wing thus constitute two distinct regions known as **lineage compartments.** This experiment indicates that the $Minute^+$/ $Minute^+$ cells formed at the time of mitotic recombination already have a restricted developmental potential: Their descendants can populate either the posterior wing compartment or the anterior wing compartment, but not both. By inducing mitotic recombination later in embryogenesis, researchers observed that the developmental potential of clones of cells becomes increasingly restricted to smaller and smaller subcompartments (for example, the anterior ventral or anterior dorsal regions of the wing) as development proceeds.

Using Mitotic Recombination to Determine When and Where a Gene Must Be Expressed to Generate a Normal Phenotype

Ommatidia are the facets of the fly's compound eye (**Fig. D.13a**). *Drosophila* geneticists have used mosaic analysis to study the normal development of the ommatidia. During the development of the eye, undifferentiated cells are recruited in a series of steps to become the eight light-sensing cells or photoreceptors (designated R1–R8) of each eye facet. The last photoreceptor cell to be recruited in each group is R7 (**Fig. D.13b**). To understand how R7 is specified, geneticists have isolated mutations that cause ommatidia to develop without an R7 cell. One gene defined by such mutations is *sevenless* (*sev*); another is *bride of sevenless* (*boss*). In what cells must these two genes be expressed if R7 is to become part of the eye facet?

To answer this question for the *sevenless* gene, biologists generated mosaics with marked cells that were mutant for both *white* and *sevenless* in an eye where other cells carried the dominant wild-type alleles for the two genes. Some facets contained a mix of red and white photoreceptor cells (Fig. D.13c). But in every facet with an R7 cell, the R7 cell was red (that is, it had both the w^+ and sev^+ genes), even if all the other photoreceptor cells in the same facet were homozygous for w^- and sev^-. There were no facets in which the R7 cell was white. These results suggest that the requirement for the sev^+ gene product is **cell autonomous:** To develop correctly, a presumptive R7 cell must make the Sev^+ protein, and it does not matter whether the protein is made in any other photoreceptor cell in the same facet (**Fig. D.13c**). The Sev^+ protein therefore affects only the cell in which it is made.

By contrast, when researchers performed the same type of analysis for *boss,* the results were very different. An R7 cell develops properly only if an R8 cell makes the $Boss^+$ protein; this is true even when the R7 cell itself is $boss^-$. If the R8 cell is $boss^-$, however, the facet does not have an R7 cell. These results show that the $Boss^+$ protein is not cell autonomous. Instead, it is produced in one cell (R8), but affects the differentiation of a different cell (R7).

If a gene is cell autonomous, only cells with a mutant genotype will express the mutant phenotype, which indicates that the gene product acts within the cell in which it is made. With a *nonautonomous gene,* the phenotype of a cell does not depend strictly on its own genotype; rather, its phenotype depends on the genotype of neighboring cells. Nonautonomy indicates that the gene product involved can influence cells other than those in which it is made (for example, by diffusing to other cells). The cloning of the *sev* and *boss* genes made it possible to understand in molecular terms the different behaviors of the two genes in the mosaic analysis (**Fig. D.13d**). The *sev* gene encodes a receptor protein that is embedded in the membrane of the developing R7 cell, while the *boss* gene encodes a protein located on the surface of the R8 cell. The binding of the $Boss^+$ protein on R8 (the ligand) to the Sev^+ receptor on the presumptive R7 cell initiates a cascade of gene activity (a signal transduction pathway; see Chapters 18 and 19 of the main textbook) in the potential R7 cell that triggers its differentiation into a mature R7 cell.

Using Mosaic Analysis to Determine Whether Genes Function at Multiple Times in Development

Suppose you are studying a gene you know contributes to early embryogenesis; animals homozygous for mutations in this gene die before they hatch into larvae. You want to know whether or not this gene plays a role in the production of structures that appear much later in development, for example, the adult eyes. You cannot answer this question by simply looking at the homozygous mutant, because it dies before the eyes form. The use of mosaics circumvents this problem. Mitotic recombination is used to produce clones of cells homozygous for the mutation in adults that are otherwise heterozygous for the mutation and hence, viable. Assuming the gene studied is autonomous in mosaics, the role of the

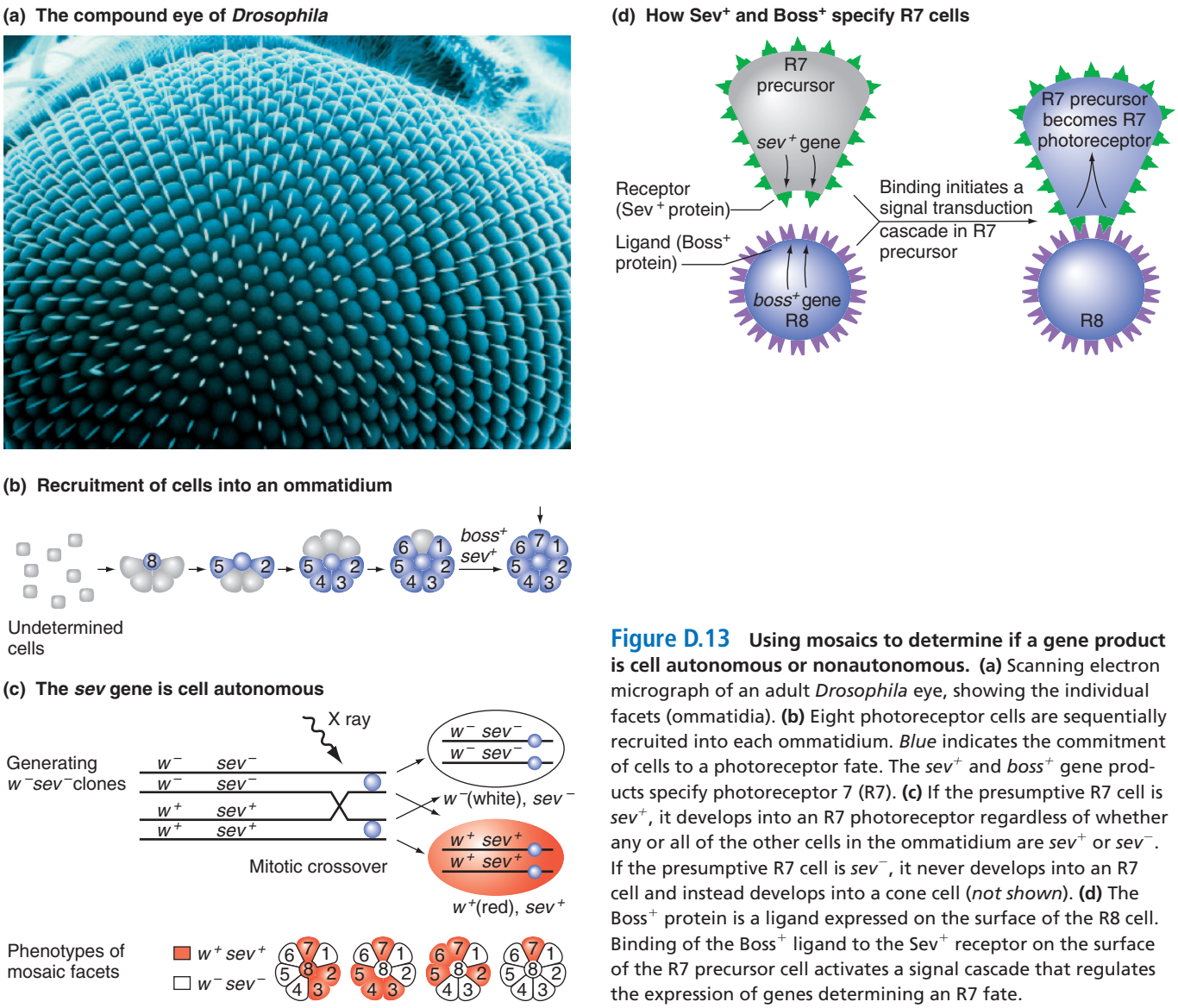


Figure D.13 **Using mosaics to determine if a gene product is cell autonomous or nonautonomous.** **(a)** Scanning electron micrograph of an adult *Drosophila* eye, showing the individual facets (ommatidia). **(b)** Eight photoreceptor cells are sequentially recruited into each ommatidium. *Blue* indicates the commitment of cells to a photoreceptor fate. The sev^+ and $boss^+$ gene products specify photoreceptor 7 (R7). **(c)** If the presumptive R7 cell is sev^+, it develops into an R7 photoreceptor regardless of whether any or all of the other cells in the ommatidium are sev^+ or sev^-. If the presumptive R7 cell is sev^-, it never develops into an R7 cell and instead develops into a cone cell (*not shown*). **(d)** The $Boss^+$ protein is a ligand expressed on the surface of the R8 cell. Binding of the $Boss^+$ ligand to the Sev^+ receptor on the surface of the R7 precursor cell activates a signal cascade that regulates the expression of genes determining an R7 fate.

gene in eye development can then be inferred from the phenotype of the homozygous mutant clones within the eyes.

A similar issue is raised by the fact that in *Drosophila,* many genes are expressed in the female germ line and their products are deposited into the developing egg. For most such genes, the maternally supplied products are sufficient to sustain development into the larval stages even if the zygote itself completely lacks the gene. To determine the effect on development of a complete lack of a gene product, it is thus usually necessary to generate homozygous mutant zygotes from homozygous mutant mothers. However, this strategy cannot be employed for lethal mutations, because homozygous mutant mothers cannot be obtained.

To circumvent this problem, geneticists generate homozygous mutant clones in the germ cells of heterozygous mothers. Since germ cells give rise to the eggs, such clones often produce eggs completely lacking the gene product of interest. As **Fig. D.14** shows, geneticists use a dominant mutation called ovo^D to ensure that the only eggs produced by heterozygous mothers are made from homozygous mutant germ-line cells. Researchers can increase the frequency of mutant clones by using the FLP/*FRT* recombination system depicted previously in Figure D.11. For many genes, homozygous mutant zygotes developing from eggs made by homozygous mutant germ-line cells show a much more severe phenotype than homozygous mutant zygotes produced by mothers having heterozygous germ lines. This result indicates the importance of the maternally supplied protein to early development.

Ectopic Expression

Although the effects of eliminating function (as in a null mutant) are of paramount importance in assessing the function of a gene, it is often helpful to determine the effects of

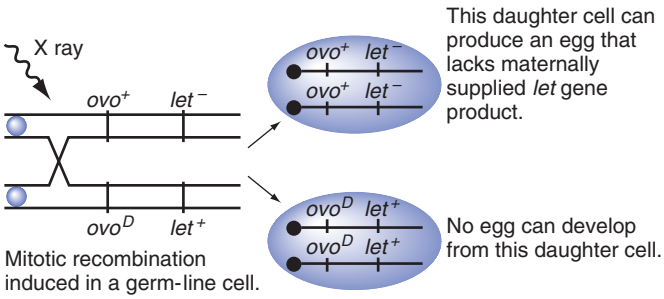


Figure D.14 Using mosaics made by mitotic recombination to determine the importance of a maternally supplied gene product for embryonic development. A female of the genotype on the *left* cannot produce any eggs because of the dominant ovo^D mutation, unless mitotic recombination creates an ovo^+/ovo^+ germ-line cell. Such a cell will also be homozygous for the lethal mutation (let^-) in question. If maternally supplied let^+ gene product is needed for early embryogenesis, fertilized eggs cannot develop into adults.

ectopic expression; that is, expression at an abnormal time or outside the cells or tissue where the gene is normally expressed. One of the most remarkable examples of the use of technology to achieve ectopic expression was described in Chapter 20 of the main textbook. As explained in that discussion, when investigators placed a *Drosophila* cDNA sequence encoding the Eyeless (Ey) protein under the control of a heat-shock promoter and introduced this artificial gene into the genome by P-element-mediated transformation, they could turn on expression of the Ey protein in all cells by subjecting the flies to elevated temperatures. Figure 20.10 in the main textbook illustrates the dramatic results: the development of ectopic eye tissue virtually anywhere on the fly's body—such as on the antennae, legs, or wings! The experiment showed that the Ey protein is a master regulator able to activate a program of eye development. This conclusion fits with the fact that the Ey protein appears to be a transcription factor that can turn on a number of target genes.

Although placing a gene under heat-shock control was the first method used to generate ectopic expression, a much improved technique emerged in the 1990s. In the newer method, investigators replace the *E. coli lacZ* reporter normally used in enhancer trap vectors (see Fig. D.10) with a yeast gene that encodes the GAL4 transcriptional activator protein (**Fig. D.15**). Although GAL4 has no effect on its own in *Drosophila,* it can function in flies to activate transcription from genes that carry a yeast DNA sequence called the *GAL4 Upstream Activator Sequence* (UAS_G). To drive ectopic expression of a gene under study, researchers first clone the gene's cDNA downstream of a UAS_G and a promoter, and they introduce this synthetic gene into the germ line by P-element-mediated transformation. UAS_G-cDNA transgenes are not transcribed in the absence of the GAL4 protein and so have no effect on the flies that carry them. However, when *UAS*-cDNA flies are crossed to a *GAL4* enhancer trap line, the cDNA is expressed in the F_1 progeny, specifically in those cells making GAL4 protein. The pattern of *GAL4* expression is determined by enhancers ("drivers") near the site where the *GAL4*-containing P element inserted into the genome. Because a large number of *GAL4* enhancer trap lines expressed in many different patterns are now available, the *GAL4* method for ectopic expression is a highly versatile tool for studies of *Drosophila* development.

The *Drosophila* Genome Project

Drosophila, like other model organisms, is the subject of a very active genome project. The sequencing of the euchromatic portion of the genome is now complete.

What the *Drosophila* Genome Sequence Has Taught Us

About 13,600 known or predicted genes are present. This gene number compares with 6000–6500 in yeast, 18,425 in the nematode, and 25,000–30,000 in humans. The average gene density in *Drosophila* is about one gene for every 9 kb. Surprisingly, the *Drosophila* gene set appears more similar

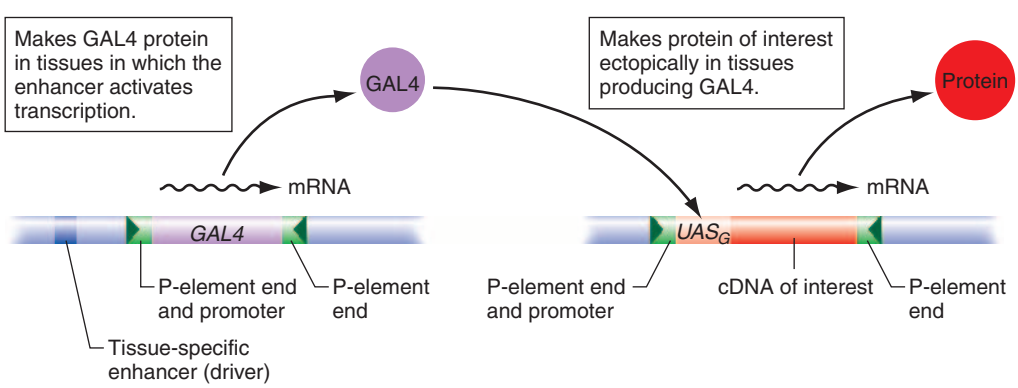


Figure D.15 Driving ectopic expression in particular cell types. This technique requires flies containing two constructs. The first is an enhancer trap that expresses GAL4 protein in tissues dictated by nearby enhancers (drivers). The second construct contains a cDNA of interest under control of UAS_G. The protein of interest will be made ectopically, specifically in those tissues producing GAL4.

to that of mammals than to that of nematodes: About half of fly proteins show homology with mammalian proteins, whereas only about a third of fly proteins appear to be homologous to nematode proteins. A key advantage of *Drosophila* for genetic experiments is that most fly genes are present in just one copy in the genome. It is thus more likely that a mutation in *Drosophila* will cause an aberrant phenotype than a mutation in mammals, whose genomes often contain several related genes that have overlapping or redundant functions.

In a recent compilation, 61% of human disease genes were found to have apparent homologs in *Drosophila.* This conservation suggests that flies have a bright future as a model for understanding many human genetic diseases. One of the first attempts to utilize *Drosophila* in this way was in the study of Huntington disease, a late-onset neurodegenerative disease caused by the expansion of CAG trinucleotide repeat numbers within the *huntingtin* gene (see the Genetics and Society box in Chapter 7 of the main textbook). These repeat arrays encode polyglutamine, which causes the neuronal cell death responsible for disease symptoms. A *huntingtin* gene containing an expanded repeat array was introduced into *Drosophila* by P-element mediated transformation. When this transgene was expressed in neural cells using the GAL4 method, it caused neural degeneration very similar to that seen in Huntington patients. To identify other genes that might play a role in the pathogenesis of Huntington, investigators have used this *Drosophila* strain to screen for new mutations in other genes that suppress or enhance the neural degeneration phenotype. The candidate genes identified in these screens are now being tested to determine whether they also play important roles in humans. In addition, the *Drosophila huntingtin* strain is being used to screen new drugs for their ability to suppress the neural degeneration phenotype. Currently, many other human disease syndromes, even including such complex disorders such as drug and alcohol addiction, are being studied in *Drosophila* by the same basic methods.

The genome sequence has provided other major surprises. One is that about 30% of the genes in *Drosophila* appear unrelated to genes in other organisms or to other genes within *Drosophila.* The functions of most of these fast-evolving "orphan genes" remain to be determined. Yet another surprise is the total gene number. Earlier genetic studies had indicated that roughly 4000 genes are essential for viability in *Drosophila.* Thus only about a third of the genes in the fly genome are essential, a far smaller proportion than had previously been thought. This number is particularly surprising when you consider that most of the genes are present only once in the genome. If so many genes are present in only a single copy and yet are not essential for viability, what are they doing?

The availability of the *Drosophila* genome sequence is having a major impact on the practice of *Drosophila* genetics. First, investigators have located many polymorphisms in the genome that have enormously expedited the mapping of point mutations to the affected gene. Second, using DNA microarrays, fly geneticists can examine global changes in gene transcription caused by mutations in key regulatory genes or by environmental changes, as discussed in Chapter 12 of the main textbook. This approach promises to revolutionize our understanding of gene regulatory networks. Finally, investigators can "mine" the *Drosophila* genome sequence using the methods of bioinformatics. For the first time, they can identify all the members of gene families and study their individual and overlapping functions. In a different type of data mining, researchers are identifying important *cis*-regulatory regions by searching the genome sequence for small regions near the 5′ ends of genes that contain clustered binding sites for known transcription factors.

Recently, the genome sequences for ten additional *Drosophila* species have been reported. These genome sequences have been a tremendous help in the search for regulatory sequence modules, as these modules typically lie within regions of high sequence conservation.

A New Direction for the *Drosophila* Genome Project: Obtaining Mutations in all Fly Genes

The goals of the *Drosophila* genome project go well beyond genome sequencing and seek to exploit the many genetic advantages of the fly to obtain functional as well as structural information about genes. One approach has been a large-scale effort to isolate transposon insertion alleles of as many genes as possible. Although the P element has been the primary transposon used in these studies, other transposons called *piggyBac, Hermes,* and *Minos,* are also being used for insertional mutagenesis. Studies of the phenotypes associated with these mutations have provided important information about the biological functions of the disrupted genes. Many of the transposons employed in the genome project are enhancer trap elements that enable rapid determination of the expression patterns of the disrupted genes (see Fig. D.10). So far, investigators have isolated transposon insertions for about 65% of the currently annotated genes in *Drosophila.*

For many years the most significant drawback to *Drosophila* as an experimental system was the lack of any method for generating targeted gene knockouts; that is, techniques to block the function of particular genes of interest. In the past few years, two such methods have been developed. In the first, knockouts are produced by homologous recombination. The procedure, which is quite complex, is as follows.

Researchers begin by engineering a clone containing the gene of interest in several ways (**Fig D.16a**). (1) The cloned gene is disrupted by insertion of a marker gene (usually w^+) within it. (2) The disrupted gene is then flanked on both sides by an 18 bp sequence that is cut by a site-specific endonuclease from yeast called I-SceI. (3) The fragment is then further modified by flanking the I-SceI

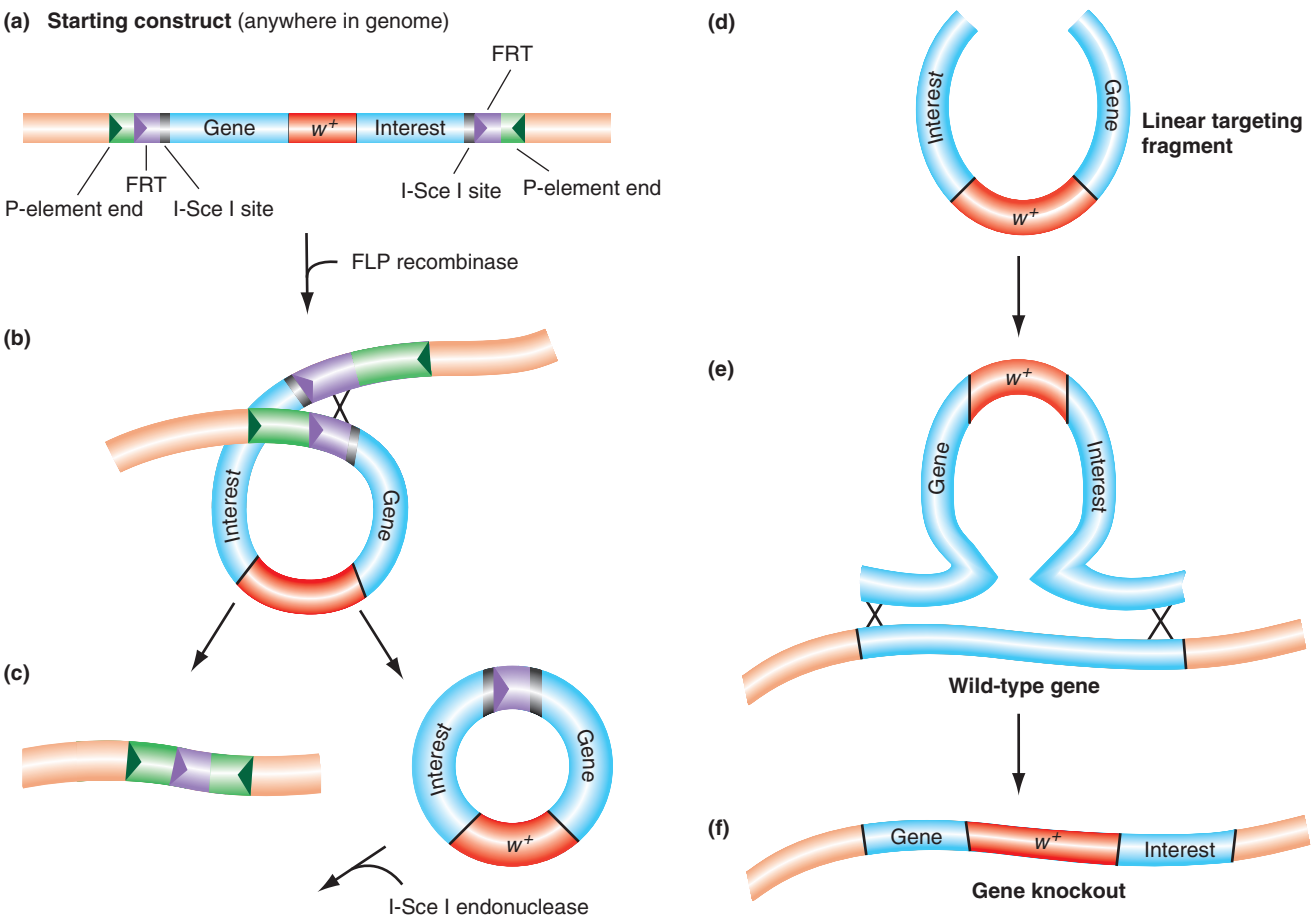


Figure D.16 **Targeted gene knockouts by homologous recombination.** Heat shock induces synthesis of FLP recombinase and I-SceI endonuclease from appropriate transgenes (review Fig. D.11). FLP excises an engineered gene of interest as a circle; I-SceI cuts the circle at the I-sites to produce a recombinogenic linear DNA fragment. Recombination replaces the endogenous wild-type gene with the defective engineered gene. This figure shows only one of several recombination events that could lead to a gene knockout.

sites with *FRT* sites (which are recognized by the FLP recombinase of yeast as described previously). A construct containing this triply modified gene is then inserted at random into the *Drosophila* genome by P-element-mediated transformation.

The next parts of the procedure are designed to generate a linear fragment containing the altered gene within the germ line of flies containing the construct (**Fig. D.16b–d**). The rationale is that linear fragments of DNA are known to stimulate recombination. Linearization requires that flies carrying the insert are crossed to a strain carrying two P-element transgenes: one with the coding sequence for FLP recombinase, the other with the coding sequence for I-SceI. Both of these P-element transgenes are under the control of heat-shock promoters. When progeny carrying all three transgenes are heat shocked, the FLP recombinase that is produced catalyzes recombination between the FRT sites flanking the introduced gene fragment. As a result, this fragment excises from the chromosome as a circle. I-SceI (whose synthesis is also stimulated by heat shock) then opens this circle by cutting at the introduced cut sites. The linearized piece ultimately recombines with the wild-type endogenous gene from which the fragment is derived, causing a targeted knockout (**Fig. D.16e–f**).

The gene targeting method just described is quite laborious, so it is fortunate that a more convenient and flexible way to knock out gene function has been developed that depends upon double-stranded RNA interference (RNAi). As described in Chapter 20, dsRNAs (double-stranded RNAs) can reduce gene expression in diverse organisms by targeting homologous mRNAs for degradation. Injection of dsRNAs into embryos effectively disrupts homologous gene expression in the embryo, but injection of dsRNAs have at best only modest effects in larvae or adults. To knock out gene expression in later developmental stages, investigators construct transgenes that encode inverted repeat or "snap-back" RNAs (**Fig. D.17**). When these transgenes are transcribed in the animal, the snap-back RNAs form double-stranded regions that can almost completely suppress expression of the endogenous gene.

The development of gene targeting techniques for *Drosophila* is very exciting, as these methods should

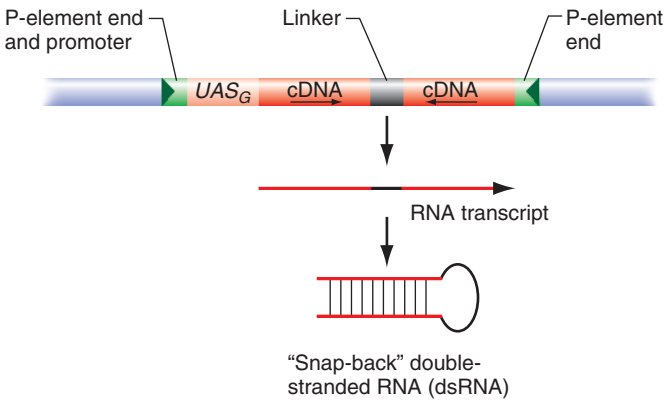


Figure D.17 RNA interference (RNAi) in *Drosophila*. Investigators clone a cDNA of interest in forward and reverse orientations, separated by a 10–100 bp linker of random DNA. When these sequences are transcribed, the resulting RNA can "snap back" by complementary base pairing to form double-stranded RNA (dsRNA) that can inhibit gene expression. Here, transcription is controlled by UAS_G so that dsRNA is made only in GAL4-producing tissues (see Fig. D.15).

facilitate the systematic inactivation of all *Drosophila* genes. By combining these technologies with those for driving ectopic expression, we should have in the not-too-distant future a nearly complete picture of basic *Drosophila* gene functions. Even so, there will still be much to do, since experience to date suggests that many genes will have no obvious effect when inactivated or ectopically expressed.

D.4 The Genetic Analysis of Body Plan Development in *Drosophila:* A Comprehensive Example

The immensely powerful combination of genetic and molecular techniques available in *Drosophila* has revolutionized our understanding of development. And because most of the developmental regulators identified in *Drosophila* are highly conserved in other animals, the *Drosophila* model is having a profound influence on other fields, including medicine. In fact many developmentally important genes in humans were first identified as homologs of genes found in *Drosophila.* A vivid illustration of the impact of *Drosophila* is the work on genes controlling the fly's body plan.

Contributions from many laboratories have revealed how the body of *Drosophila* becomes specialized along both its anterior-posterior (AP) and dorsal-ventral (DV) axes. Together, these studies constitute one of the most successful genetic dissections ever performed. Here we consider only development of the AP body axis, that is, how the body differentiates along the line running from the head to the tail of the animal.

Very early in development, the action of a large group of genes known as the **segmentation genes** subdivides the body into an array of essentially identical body segments. Expression of a second set of genes called **homeotic genes** then assigns a unique identity to each body segment. In studies of the genes that specify and regulate development of the AP body axis, researchers have tried to answer two basic questions. First, how does the developing animal establish the proper number of body segments? And second, how does each body segment "know" what kinds of structures it should form and what role it should play in the animal's biology?

Early Development of the Basic Body Plan

To understand how the segmentation and homeotic genes function, it is helpful to consider some of the basic events that take place in the first few hours of *Drosophila* development (**Fig. D.18a** and **b**). The egg is fertilized in the uterus as it is being laid, and the meiotic divisions of the oocyte nucleus, which had previously arrested in the metaphase of meiosis I, resume at this time. After fusion of the haploid male and female pronuclei, the diploid zygotic nucleus of the embryo undergoes 13 rounds of nuclear division at an extraordinarily rapid rate, with the average time of mitotic cycles 2 through 9 being only 8.5 minutes. Nuclear division in early *Drosophila* embryos, unlike most mitoses, is not accompanied by cell division, so that the early embryo becomes a multinucleate syncytium. During the first eight division cycles, the multiple nuclei are centrally located in the egg; during the ninth division, most of the nuclei migrate out to the cortex—just under the surface of the embryo—to produce the **syncytial blastoderm.** During the tenth division, nuclei at the posterior pole of the egg are enclosed in membranes that invaginate from the egg cell membrane to form the first embryonic cells; these "pole cells" are the primordial germ cells. At the end of the thirteenth division cycle, about 6000 nuclei are present at the egg cortex.

During the interphase of the fourteenth cycle, membranes in the egg's cortex grow inward between these nuclei, creating an epithelial layer called the **cellular blastoderm** that is one cell deep (**Figs. D.18** and **D.19a**). The embryo completes formation of the cellular blastoderm about 3 hours after fertilization. At the cellular blastoderm stage, no regional differences in cell shape or size are apparent (with the exception of the pole cells at the posterior end). Experiments in which blastoderm cells have been transplanted from one location to another, however, show that despite this morphological uniformity, the segmental identity of the cells has already been determined. Consistent with this finding, molecular studies reveal that most segmentation and homeotic genes function during or even before the cellular blastoderm stage.

Immediately after cellularization, **gastrulation** and establishment of the embryonic germ layers begin. The

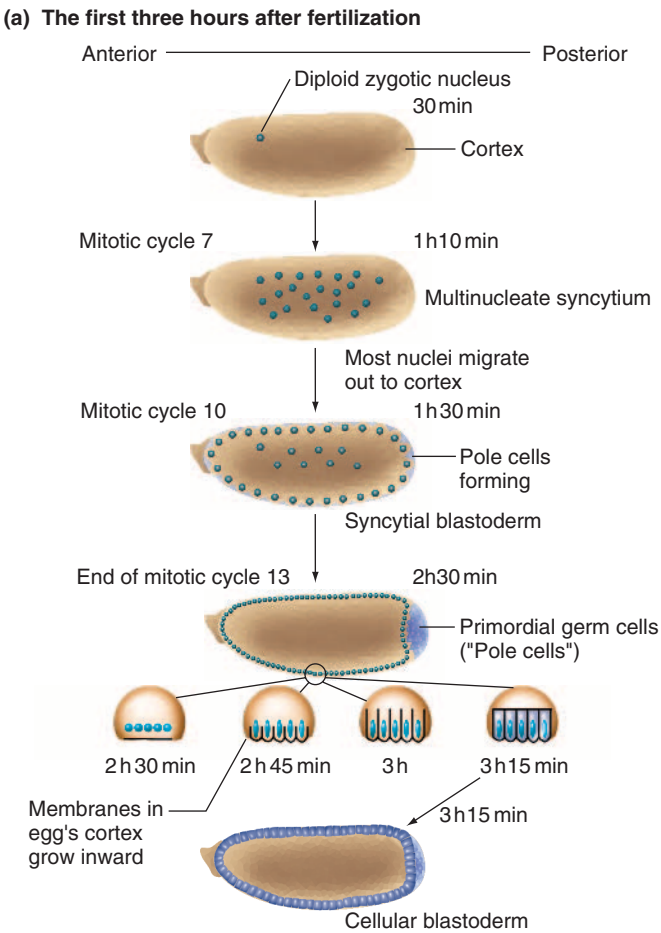


Figure D.18 **Early *Drosophila* development: From fertilization to formation of the cellular blastoderm.** **(a)** The original zygotic nucleus undergoes 13 very rapid mitotic divisions in a single syncytium. A few nuclei at the posterior end of the embryo become the germ-line *pole cells.* At the *syncytial blastoderm* stage, the egg surface is covered by a monolayer of nuclei (except for the pole cells at the posterior end). At the end of the thirteenth division cycle, cell membranes enclose the nuclei at the cortex into separate cells to produce a *cellular blastoderm.* **(b)** Photomicrographs of early embryonic stages stained with a fluorescent dye for DNA. (For another view of the syncytial blastoderm, see Fig. 4.10 of main textbook.)

mesoderm forms by invagination of a band of midventral cells that extends most of the length of the embryo. This infolding (the ventral furrow; **Fig. D.19b**) produces an internal tube whose cells soon divide and migrate to produce a mesodermal layer. The *endoderm* forms by distinct invaginations anterior and posterior to the ventral furrow. The cells of these invaginations migrate over the yolk to produce the gut. Finally, the nervous system arises from neuroblasts that segregate from bilateral zones of the ventral *ectoderm.* The first visible signs of segmentation are periodic bulges in the mesoderm, which appear about 40 minutes after gastrulation begins. Within a few hours of gastrulation, the embryo is divided into clear-cut body segments that will become the posterior three head segments, three thoracic segments, and eight major abdominal segments of the larva that hatches from the eggshell (**Fig. D.19c**). Even though the animal eventually undergoes metamorphosis to become an adult fly, the same basic body plan is conserved in the adult stage (**Fig. D.19d**).

The genes responsible for the formation of segments fall into four classes and function in a regulatory hierarchy that progressively subdivides the embryo into successively smaller units. In the order of their expression, these classes are (1) maternal genes, (2) gap genes, (3) pair-rule genes, and (4) segment-polarity genes.

Specification of Segment Number: Maternal Genes Interact to Produce Morphogen Gradients

There is very little transcription of genes in the embryonic nuclei between fertilization and the end of the 13 rapid syncytial divisions that immediately follow fertilization. Because of this near (but by no means total) lack of transcription, developmental biologists suspected that formation of the basic body plan initially requires **maternally supplied components** deposited by the mother into the egg during oogenesis. How could they identify the genes encoding these maternally supplied components? Christiane Nüsslein-Volhard and Eric Wieschaus realized that the embryonic phenotype determined by such genes does not depend on the embryo's own genotype; rather it is determined by the genotype of the mother. By asking whether mothers homozygous for candidate mutations would produce defective embryos, they devised genetic screens to identify recessive mutations in maternal genes that influence embryonic development; these recessive mutations are often called **maternal-effect mutations.**

To carry out their screens, Nüsslein-Volhard and Wieschaus established individual balanced stocks for thousands of mutagen-treated chromosomes, and then they examined the phenotypes of embryos obtained from homozygous mutant mothers. They focused their attention on stocks in which homozygous mutant females were sterile, because they anticipated that the absence of maternally supplied

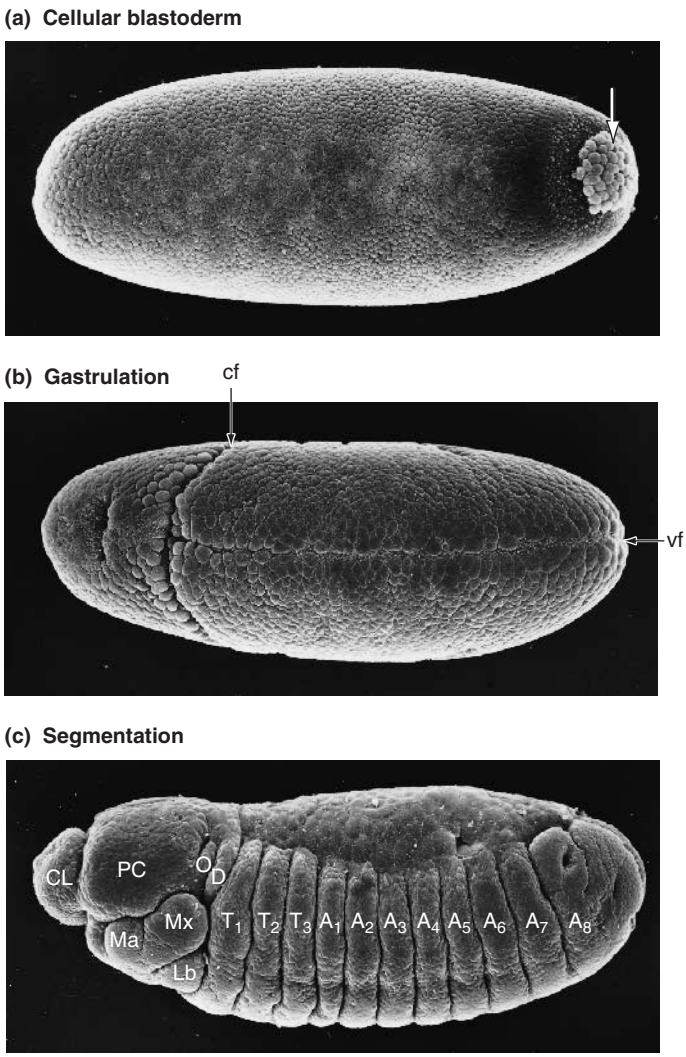


Figure D.19 ***Drosophila*** **development after formation of the cellular blastoderm.** **(a)** Scanning electron micrograph of a cellular blastoderm; individual cells are visible at the periphery of the embryo, and the pole cells at the posterior end can be distinguished (*arrow*). **(b)** A ventral view of some of the furrows that form during gastrulation, roughly 4 hours after fertilization: vf, ventral furrow; cf, cephalic furrow. **(c)** By 10 hours after fertilization, it is clear that the embryo is subdivided into segments. Ma, Mx, Lb are the three posterior head segments (mandibular, maxillary, and labial, respectively). CL, PC, O, and D refer to nonsegmented regions of the head. The three thoracic segments (labeled T_1, T_2, and T_3) are the prothorax, the mesothorax, and the metathorax, respectively, while the abdominal segments are labeled A_1–A_8. **(d)** The identity of embryonic segments (*left*) is preserved through the larval stages and is also retained through metamorphosis into the adult (*right*). For simplicity, head segments are not distinguished here.

components needed for the earliest stages of development would result in embryos so defective that they could never grow into adults. Through these large-scale screens, Nüsslein-Volhard and Wieschaus identified a large number of maternal genes that are required for the normal patterning of the body. For this and other work, they shared the Nobel Prize for physiology or medicine with Edward B. Lewis—whose contributions we describe later. We focus here on two groups of the genes they found. One group is required for normal patterning of the embryo's anterior; the other group is required mainly for normal posterior patterning. The genes in these two groups are the first genes activated in the process that determines segment number.

The finding that separate groups of maternal genes control anterior and posterior patterning in the embryo is consistent with the conclusions of classical embryological experiments. Studies in which polar cytoplasm from the embryo's ends was transplanted, or in which preblastoderm embryos were separated into two halves by constriction of the embryo with a fine thread, suggested that the insect body axis is patterned during cleavage by the interaction of two signaling centers located at the anterior and posterior poles of the egg. In a specific model, Klaus Sander proposed that each pole of the egg produces a different substance, and that these substances form opposing gradients by diffusion. He suggested that the concentrations of these substances then determine the types of structures produced at each position along the body axis. Molecular characterization of the maternal genes of the anterior and posterior groups indicates that the Sander model for body axis patterning is correct. Substances that define different cell fates in a concentration-dependent manner are known as **morphogens.**

In *Drosophila,* the *bicoid* (*bcd*) Gene Encodes the Anterior Morphogen

Embryos from mothers homozygous for null alleles of *bcd* lack all head and thoracic structures. The protein product of *bcd* is a DNA-binding transcription factor whose transcript is localized near the anterior pole of the egg cytoplasm (**Fig. D.20a**). Translation of the *bcd* transcripts takes place after fertilization. The newly made Bcd protein diffuses

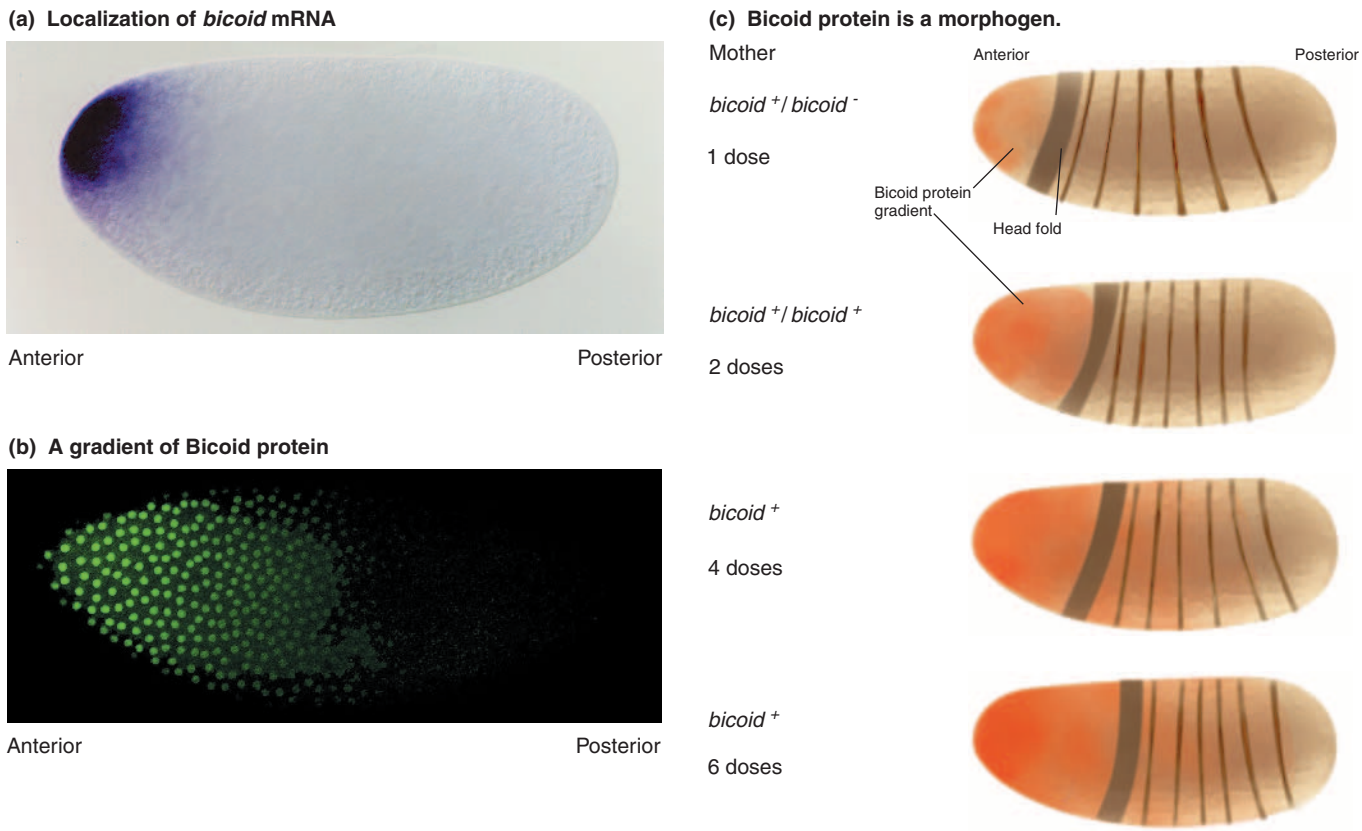


Figure D.20 **The Bicoid protein is the anterior morphogen.** **(a)** The *bicoid* (*bcd*) mRNA (visualized by *in situ* hybridization in *purple*) concentrates at the anterior tip of the embryo. **(b)** The Bicoid (Bcd) protein (seen by *green* antibody staining) is distributed in a gradient: high at the anterior end and trailing off toward the posterior end. It can be seen that the Bcd protein (a transcription factor) accumulates in the nuclei of this syncytial blastoderm embryo. **(c)** The greater the maternal dosage of bcd^+, the higher the concentration of Bcd protein in the embryo, and the more of the embryo that is devoted to anterior structures. Head structures will develop anterior of the head fold invagination; thoracic and abdominal structures posterior to it.

from its source at the pole to produce a high-to-low, anterior-to-posterior concentration gradient that extends over the anterior two-thirds of the embryo by the ninth division cycle (**Fig. D.20b**). This gradient determines most aspects of head and thorax development.

One of the first lines of evidence that the Bcd protein functions as a morphogen came from experiments in which the maternal dosage of the *bcd* gene varied (Fig. D.20c). Mothers that carried only one dose of the *bcd* gene, instead of the normal diploid dose, incorporated about half the normal amount of *bcd* RNA into their eggs. As a result, translation yielded less Bcd protein, and the Bcd gradient was shallower and shifted to the anterior. In these Bcd-deficient embryos, the thoracic segments developed from more anterior regions than normal, and less of the body was devoted to the head. The opposite effect occurred in mothers carrying extra doses of the *bcd* gene. These and other observations suggested that the level of Bcd protein is a key to the determination of head and thoracic fates in the embryo. Three other genes work with *bcd* in the anterior group of maternal genes; the function of the protein products of these three genes is to localize *bcd* transcripts to the egg's anterior pole.

The Bcd protein itself works in two ways: as a transcription factor that helps control the transcription of genes farther down the regulatory pathway, and as a translational repressor. The target of its repressor activity is the transcript of the *caudal* (*cad*) gene, which also encodes a DNA-binding transcription factor. The *cad* transcripts are uniformly distributed in the egg before fertilization, but because of translational repression by the Bcd protein, translation of these transcripts produces a gradient of Cad protein that is complementary to the Bcd gradient. That is, there is a high concentration of Cad protein at the posterior end of the embryo and lower concentrations toward the anterior (**Fig. D.21**). The Cad protein plays an important role in activating genes expressed later in the segmentation pathway to generate posterior structures.

The *nanos* (*nos*) Gene Encodes the Primary Posterior Morphogen

The *nos* RNA is localized to the posterior egg cytoplasm by proteins encoded by other posterior group maternal genes. Like *bcd* RNAs, *nos* transcripts are translated during the

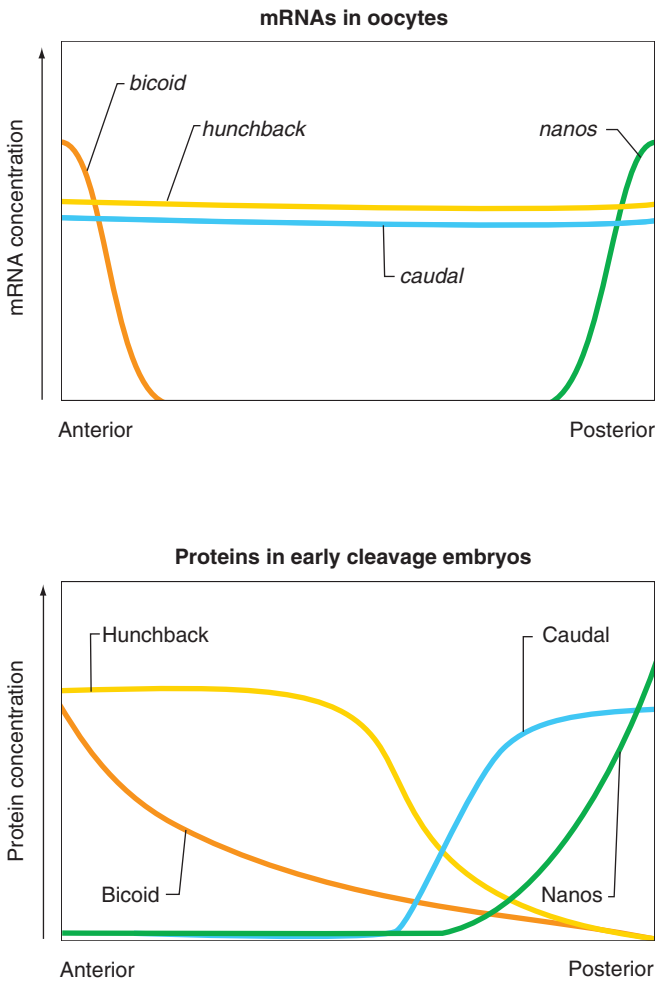


Figure D.21 **Distribution of the mRNA and protein products of maternal-effect genes within the early embryo.** *Top:* In the oocyte prior to fertilization, *bicoid* (*bcd*) mRNA is concentrated near the anterior tip and *nanos* (*nos*) mRNA at the posterior tip, while maternally supplied *hunchback* (*hb*) and *caudal* (*cad*) mRNAs are uniformly distributed. *Bottom:* In early cleavage stage embryos, the Bicoid (Bcd) and Hunchback (Hb) proteins are found in concentration gradients high at the anterior and lower toward the posterior (A to P), while the Nanos (Nos) and Caudal (Cad) proteins are distributed in opposite P-to-A gradients.

cleavage stages. After translation, diffusion produces a posterior-to-anterior Nos protein concentration gradient. The Nos protein, unlike the Bcd protein, is not a transcription factor; rather, the Nos protein functions only as a translational repressor. Its major target is the maternally supplied transcript of the *hunchback* (*hb*) gene, which is deposited in the egg during oogenesis and is uniformly distributed before fertilization. For development to occur properly, the Hb protein (which is another transcription factor) must be present in a gradient with high concentrations at the embryo's anterior and low concentrations at the posterior. The Nos protein, which represses the translation of *hb* maternal mRNA and is present in a posterior-to-anterior concentration gradient, helps construct the anterior-to-posterior Hb gradient by lowering the concentration of the Hb protein toward the embryo's posterior pole (**Fig. D.21**). The embryo also has a second mechanism for establishing the Hb protein gradient that functions somewhat later: It only transcribes the *hb* gene from zygotic nuclei in the anterior region (see following).

Figure D.21 summarizes how the coordinated activity of the maternal genes establishes polarities in the embryo that eventually result in proper embryonic segmentation. In addition to distributing the Bcd and Cad transcription factors in gradients within the embryo, the maternal genes ensure that the maternally supplied *hb* RNA is not translated in the posterior part of the embryo.

Specification of Segment Number Through the Activation of Zygotic Genes in Successively More Sharply Defined Regions

The maternally determined Bcd, Hb, and Cad protein gradients control the spatial expression of zygotic segmentation genes. Unlike the products of maternal genes, whose mRNAs are placed in the egg during oogenesis, the products of zygotic genes are transcribed and translated from DNA in the nuclei of embryonic cells descended from the original zygotic nucleus. The expression of zygotic segmentation genes begins in the syncytial blastoderm stage, a few division cycles before cellularization (roughly cycle 10).

Most of the zygotic segmentation genes were identified in a second mutant screen also carried out in the late 1970s by Christiane Nüsslein-Volhard and Eric Wieschaus. In this screen, the two *Drosophila* geneticists placed individual ethyl methane sulfonate (EMS)-mutagenized chromosomes into balanced stocks and then examined homozygous mutant embryos from these stocks for defects in the segmentation pattern of the embryo. Such homozygous mutant embryos were so aberrant that they were unable to grow into adults; thus the mutations causing these defects would be classified as recessive lethals. After screening several thousand such stocks for each of the *Drosophila* chromosomes, they identified three classes of zygotic segmentation genes: gap genes (9 different genes); pair-rule genes (8 genes); and segment-polarity genes (about 17 genes). These three classes of zygotic genes fit into a hierarchy of gene expression.

Gap Genes Are Expressed in Broad Zones Along the Anterior-Posterior Axis

The gap genes are the first zygotic segmentation genes to be transcribed. Embryos homozygous for mutations in the gap genes show a gap in the segmentation pattern caused by an absence of particular segments that correspond to the position at which each gene is transcribed (**Fig. D.22**). How

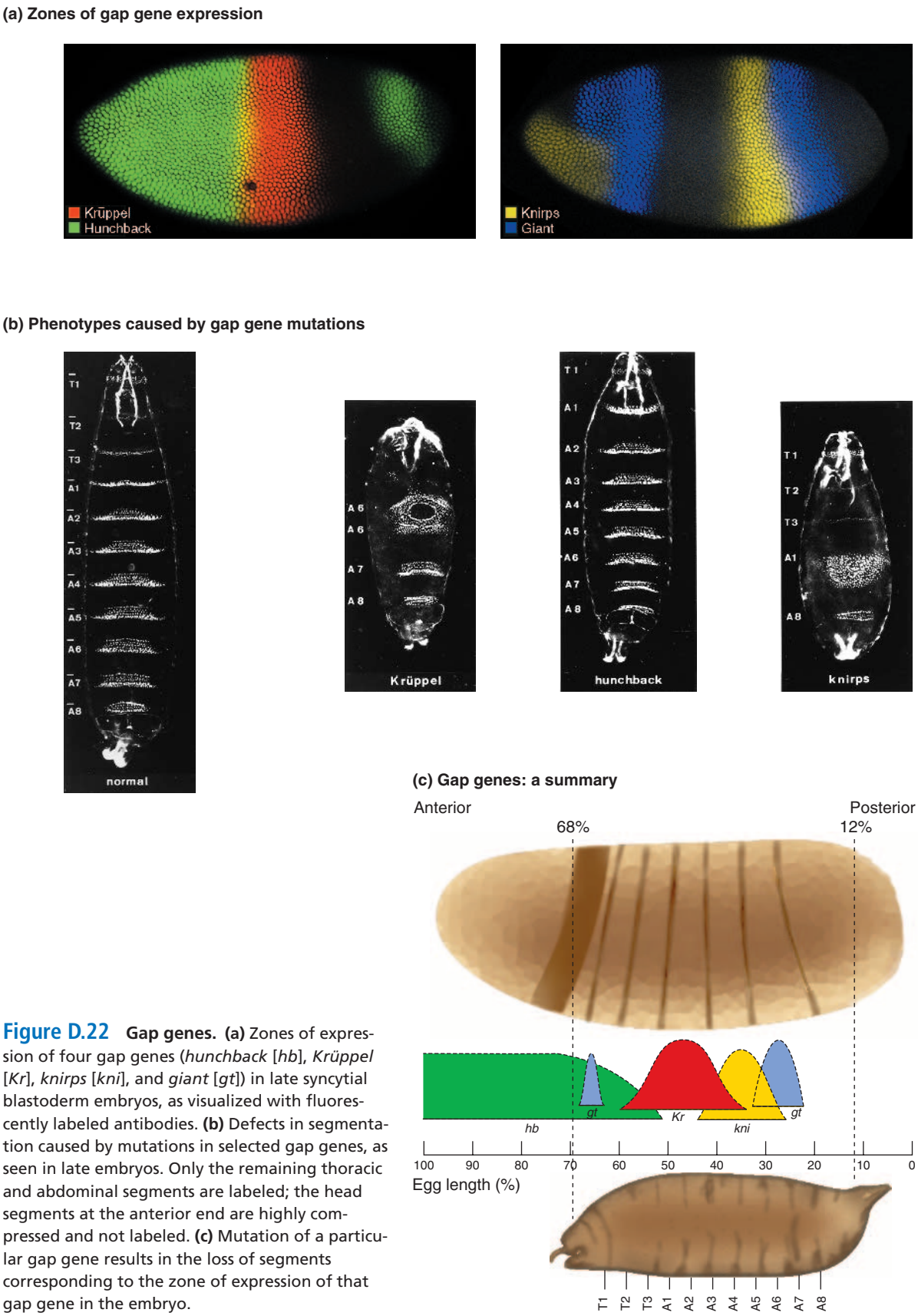


Figure D.22 **Gap genes.** **(a)** Zones of expression of four gap genes (*hunchback* [*hb*], *Krüppel* [*Kr*], *knirps* [*kni*], and *giant* [*gt*]) in late syncytial blastoderm embryos, as visualized with fluorescently labeled antibodies. **(b)** Defects in segmentation caused by mutations in selected gap genes, as seen in late embryos. Only the remaining thoracic and abdominal segments are labeled; the head segments at the anterior end are highly compressed and not labeled. **(c)** Mutation of a particular gap gene results in the loss of segments corresponding to the zone of expression of that gap gene in the embryo.

do the maternal transcription factor gradients ensure that the various gap genes are expressed in their broad zones at the proper position in the embryo? Part of the answer is that the binding sites in the promoter regions of the gap genes have different affinities for the maternal transcription factors. For example, some gap genes are activated by the Bcd protein (the anterior morphogen). Gap genes such as *hb* with low-affinity Bcd protein-binding sites will be activated only in the most anterior regions, where the concentration of Bcd is at its highest; by contrast, genes with high-affinity sites will be activated farther toward the posterior pole.

Another part of the answer is that the gap genes themselves encode transcription factors that can influence the expression of other gap genes. The *Krüppel* (*Kr*) gap gene, for example, appears to be turned off by high amounts of Hb protein at the anterior end of its band of expression; activated within its expression band by Bcd protein in conjunction with lower levels of Hb protein; and turned off at the posterior end of its expression zone by the products of the *knirps* (*kni*) gap gene (**Fig. D.22c**). (Note that the *hb* gene is usually classified as a gap gene, despite the maternal supply of some *hb* RNA, because the protein translated from the transcripts of zygotic nuclei actually plays the more important role.)

Pair-Rule Genes Subdivide the Body into Units That Are Two Segments in Length

After the gap genes have divided the body axis into rough, generalized regions, activation of the pair-rule genes generates more sharply defined sections. These genes encode transcription factors that are expressed in seven stripes in preblastoderm and blastoderm embryos (**Fig. D.23a**). The stripes have a two-segment periodicity; that is, there is one stripe for every two segments. Mutations in pair-rule genes cause the deletion of similar pattern elements from every alternate segment. For example, larvae mutant for *fushi tarazu* ("segment deficient" in Japanese) lack parts of abdominal segments A1, A3, A5, and A7 (see Fig. 20.2 of the main textbook). Mutations in *even-skipped* cause loss of even-numbered abdominal segments.

There are two classes of pair-rule genes: primary and secondary. The striped expression pattern of the three primary pair-rule genes depends on the transcription factors encoded by the maternal genes and the zygotic gap genes. Specific elements within the upstream regulatory region of each pair-rule gene drive the expression of that pair-rule gene within a particular stripe. For example, as Fig. D.23b and c shows, the DNA region responsible for driving the expression of *even-skipped* (*eve*) in the second stripe contains multiple binding sites for the Bcd protein and the proteins encoded by the gap genes *Krüppel, giant* (*gt*), and *hb.* The transcription of *eve* in this stripe of the embryo is activated by Bcd and Hb, while it is repressed by Gt and Kr; only in the stripe 2 region are Gt and Kr levels low enough and Bcd and Hb levels high enough to allow activation of

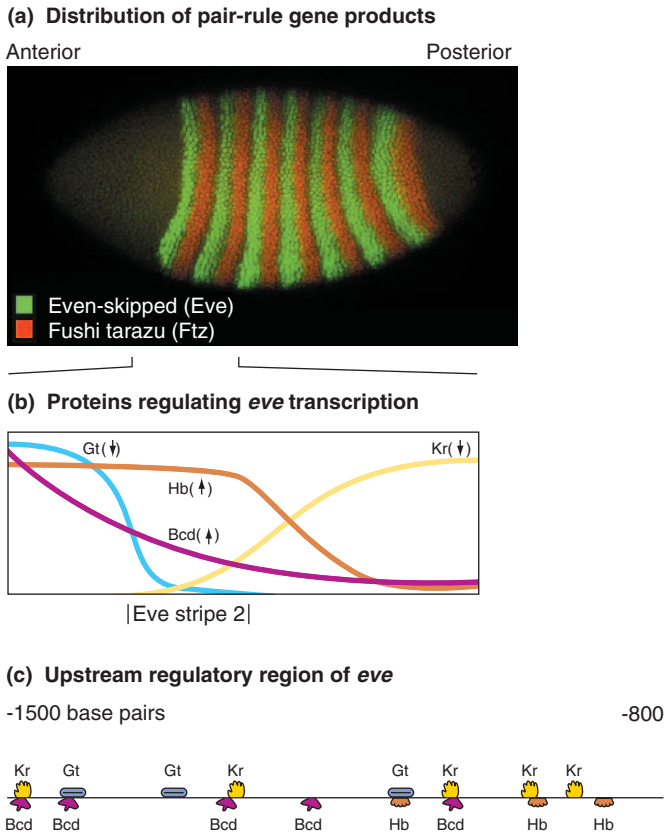


Figure D.23 **Pair-rule genes. (a)** Zones of expression of the proteins encoded by the pair-rule genes *fushi tarazu* (*ftz*) and *even-skipped* (*eve*) at the beginning of the cellular blastoderm stage. Each gene is expressed in seven stripes. Eve stripe 2 is the second green stripe from the *left.* **(b)** The formation of Eve stripe 2 requires activation of *eve* transcription by the Bcd and Hb proteins and repression at its left and right ends by Gt and Kr proteins respectively. **(c)** The 700 bp-long upstream regulatory region of the *eve* gene that directs the Eve second stripe contains multiple binding sites for the four proteins shown in part (b).

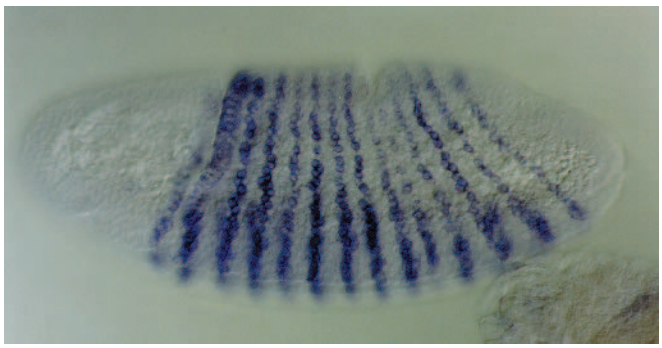
the element driving *eve* expression. In contrast with the primary pair-rule genes, the five pair-rule genes of the secondary class are controlled largely by interactions with transcription factors encoded by other pair-rule genes.

Segment Polarity Genes Occupy the Lowest Level of the Segmentation Hierarchy

Many segment polarity genes are expressed in stripes that are repeated with a single segment periodicity, that is, there is one stripe per segment (**Fig. D.24a**). Mutations in segment polarity genes cause deletion of part of each segment, often accompanied by mirror-image duplication of the remaining parts. The segment polarity genes thus function to determine patterns that are repeated within each segment.

The regulatory system that directs the expression of segment polarity genes in a single stripe per segment is quite complex. In general, the transcription factors encoded

(a) Distribution of engrailed RNA



(b) Segment polarity genes establish compartment borders.

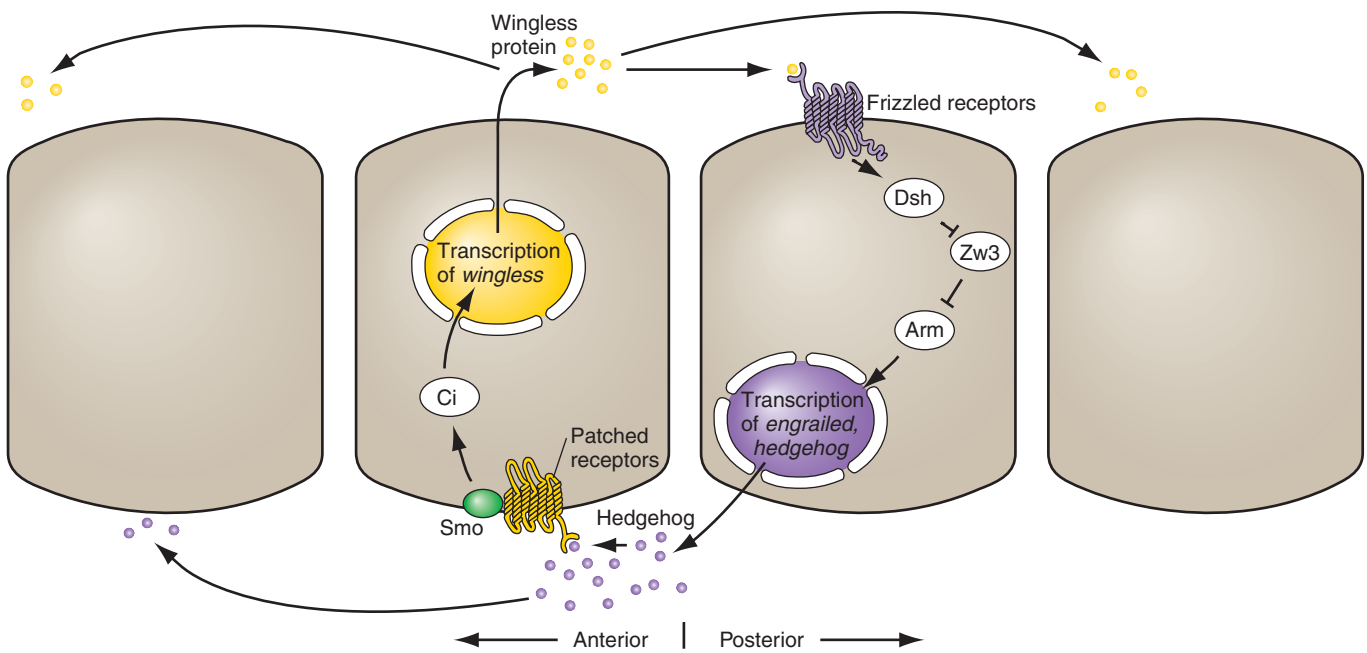


Figure D.24 **Segment polarity genes.** **(a)** Wild-type embryos express the segment polarity gene *engrailed* in 14 stripes. **(b)** The border between a segment's posterior and anterior compartments is governed by the *engrailed* (*en*), *wingless* (*wg*), and *hedgehog* (*hh*) segment polarity genes. Regulation by pair-rule genes ensures that cells in posterior compartments express *en.* The En protein activates the transcription of the *hh* gene, which encodes a secreted protein ligand. Binding of this Hh protein to the Patched receptor in the adjacent anterior cell initiates a signal transduction pathway (through the Smo and Ci proteins) leading to the transcription of the *wg* gene. Wg is also a secreted protein that binds to a different receptor in the posterior cell, which is encoded by *frizzled.* Binding of the Wg protein to this receptor initiates a different signal transduction pathway (including the Dsh, Zw3, and Arm proteins) that stimulates the transcription of *en* and of *hh.* The result is a reciprocal loop stabilizing the alternate fates of adjacent cells at the border.

by pair-rule genes initiate the pattern by directly regulating certain segment polarity genes. Interactions between various cell polarity genes then maintain this periodicity later in development. Significantly, activation of segment polarity genes occurs after cellularization of the embryo is complete. Hence, the diffusion of transcription factors within the syncytium, of central importance in segmentation up to this point, ceases to play a role. Instead, intrasegmental patterning is determined mostly by the diffusion of secreted proteins between cells.

Two of the segment polarity genes, *hedgehog* (*hh*) and *wingless* (*wg*), encode secreted proteins. These proteins, together with the transcription factor encoded by the *engrailed* (*en*) segment polarity gene, are responsible for many aspects of segmental patterning (**Fig. D.24b**). A key component of this control is that a one-cell-wide stripe of cells secreting the Wg protein is adjacent to a stripe of cells expressing the En protein and secreting the Hh protein. The interface of these two types of cells is a self-reinforcing, reciprocal loop. The Wg protein secreted by the more anterior of the two adjacent stripes of cells is required for the continued expression of *hh* and *en* in the adjacent posterior stripe. The Hh protein secreted by the more posterior stripe of cells maintains expression of *wg* in the anterior stripe.

Gradients of Wg and Hh proteins made from these adjacent stripes of cells control many aspects of patterning in the remainder of the segment. This is because the products of both *wg* and *hh* appear to function as morphogens; that is, responding cells appear to adopt different fates depending on the concentration of Wg or Hh protein to which they are exposed.

Other segment polarity genes encode proteins involved in signal transduction pathways initiated by the binding of Wg and Hh proteins to receptors on cell surfaces. Recall from Chapters 19 and 20 of the main textbook that signal transduction pathways enable a signal received from a receptor on the cell's surface to be converted through a series of intermediate steps to a final intracellular regulatory response, usually the activation or repression of particular target genes. The signal transduction pathways initiated by the Wg and Hh proteins determine the ability of cells in portions of each segment to differentiate into the particular cell types characteristic of that segment portion.

Homologs of the segment polarity genes are key players in many important patterning events in vertebrates. For example, the chicken *sonic hedgehog* gene (related to the fly *hh*) is critical for the initiation of the left-right asymmetry in the early chicken embryo as well as for the processes that determine the number and polarity of digits produced by the limb buds. The mammalian homolog of this gene has the same conserved functions.

Summary of Segment Number Specification

For each class of segmentation gene, the pattern of expression of its members is controlled either by classes of genes higher in the hierarchy or by members of the same class, never by genes of a lower class (**Fig. D.25**). In this regulatory cascade, the maternal genes control the gap and pair-rule genes, the gap genes control themselves and the pair-rule genes, and the pair-rule genes control themselves and the segment polarity genes.

The expression of genes in successively lower parts of the hierarchy is restricted to more sharply defined embryonic regions. During the early embryonic cycles of nuclear divisions, the spatial restriction of gene expression is a response to maternally established transcription factor gradients. In the syncytial blastoderm embryo, refinement of the spatial restriction of gene expression depends on the hierarchical action of the transcription factors encoded by the zygotic genes. Once cellularization is complete, intercellular communication mediated by secreted proteins encoded

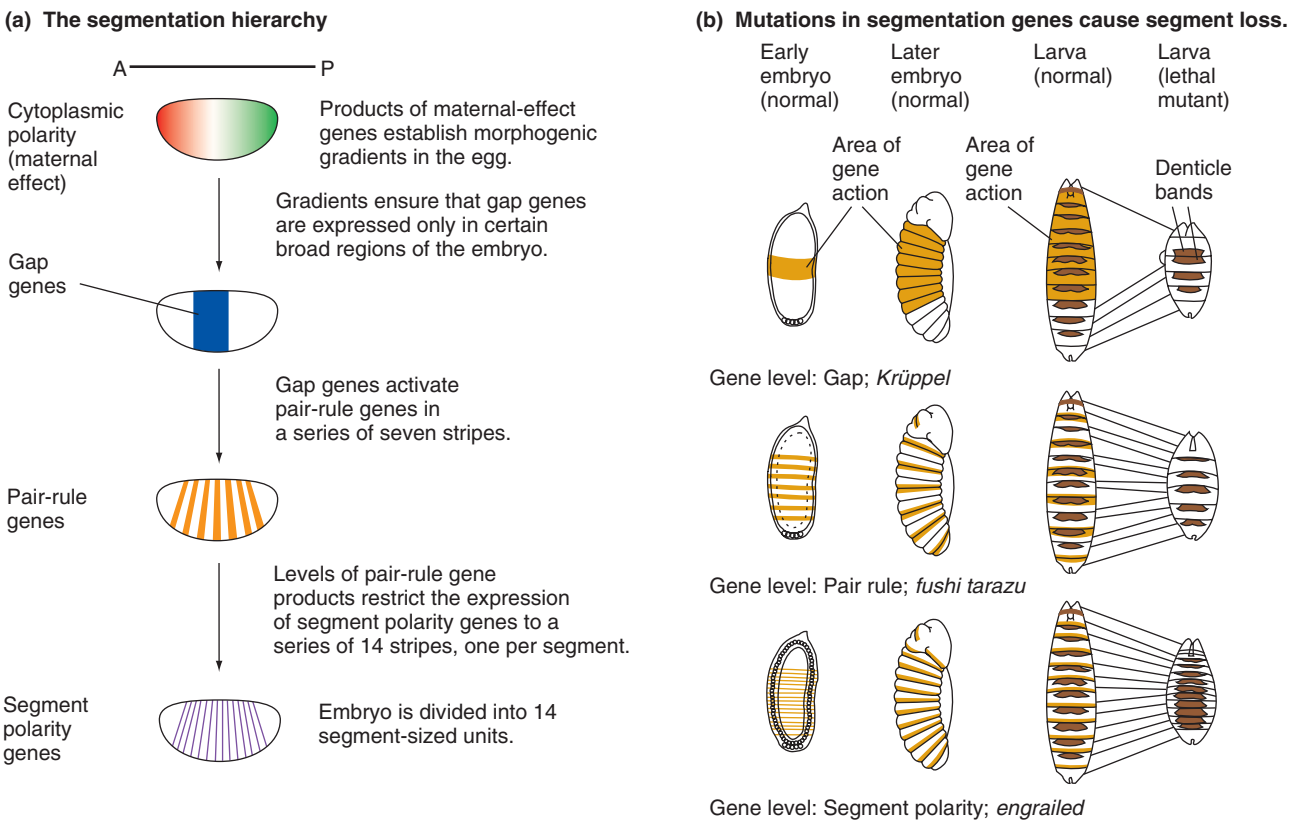


Figure D.25 **A summary of the genetic hierarchy leading to segmentation in *Drosophila*.** **(a)** Genes in successively lower parts of the hierarchy are expressed in narrower bands within the embryos. **(b)** Mutations in segmentation genes cause the loss of segments that correspond to regions where the gene is expressed (*shown in yellow*). The denticle bands (*dark brown*) are features that help researchers identify the segments remaining in the mutant animals.

by certain segment polarity genes becomes the major mode of pattern formation. Although the cellular blastoderm, when viewed from the outside, looks like a uniform layer of cells (as seen in Fig. D.19a), the coordinated action of the segmentation genes has already divided the embryo into segment primordia. A few hours after gastrulation, these primordia become distinguishable morphologically as clear-cut segments (as in Fig. D.19c).

Each Segment Establishes Its Own Unique Identity Through the Activation of Homeotic Genes

After the segmentation genes have subdivided the body into a precise number of segments, the homeotic genes help assign a unique identity to each segment. They do this by functioning as master regulators that control the transcription of batteries of genes responsible for the development of segment-specific structures. The homeotic genes themselves are regulated by the gap, pair-rule, and segment polarity genes so that at the cellular blastoderm stage, or shortly thereafter, each homeotic gene becomes expressed within a specific subset of body segments. Most homeotic genes then remain active through the rest of development, functioning continuously to direct proper segmental specialization.

Mutations in homeotic genes, referred to as **homeotic mutations,** cause particular segments, or parts of them, to develop as if they were located elsewhere in the body. Because some of the mutant homeotic phenotypes are quite spectacular, researchers noticed them very early in *Drosophila* research. In 1915, for example, Calvin Bridges found a mutant he called *bithorax* (*bx*). In homozygotes for this mutation, the anterior portion of the third thoracic segment (T3) develops like the anterior second thoracic segment (T2); in other words, this mutation transforms part of T3 into the corresponding part of T2, as illustrated in **Fig. D.26a**. This mutant phenotype is very dramatic, as T3 normally produces only small club-shaped balancer organs called *halteres,* whereas T2 produces the wings. Another homeotic mutation is *postbithorax* (*pbx*), which affects only posterior T3, causing its transformation into posterior T2. (Note that in this context, *Drosophila* geneticists use the term "transformation" to mean a change of body form.) In the *bx pbx* double mutant, all of T3 develops as T2 to produce the now famous four-winged fly (**Fig. D.26b**).

In the last half of the twentieth century, researchers isolated many other homeotic mutations, most of which map within one of two gene clusters. Mutations affecting segments in the abdomen and posterior thorax lie within a cluster known as the **bithorax complex (BX-C);** mutations affecting segments in the head and anterior thorax lie within the **Antennapedia complex (ANT-C)** (**Fig. D.27**).

The Bithorax Complex

Edward B. Lewis shared the 1995 Nobel Prize for physiology or medicine with Christiane Nüsslein-Volhard and

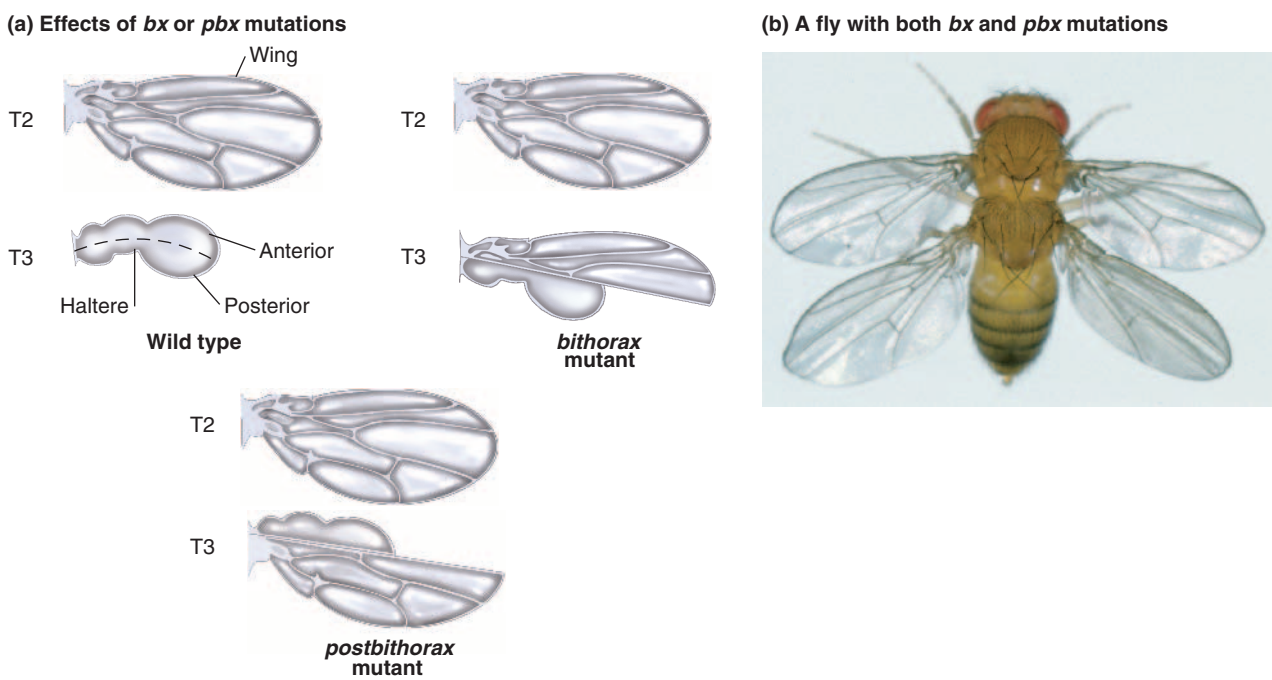


Figure D.26 **Homeotic transformations. (a)** In animals homozygous for the mutation *bithorax* (*bx*), the anterior compartment of T3 (the third thoracic segment that makes the haltere) is transformed into the anterior compartment of T2 (the second thoracic segment that makes the wing). The mutation *postbithorax* (*pbx*) transforms the posterior compartment of T3 into the posterior compartment of T2. **(b)** In a *bx pbx* double mutant, T3 is changed entirely into T2. The result is a four-winged fly.

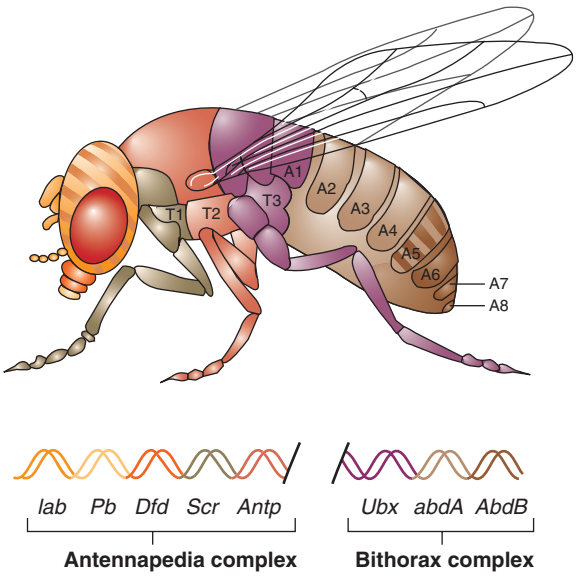


Figure D.27 **Homeotic selector genes.** Two clusters of genes on chromosome 3—the Antennapedia complex and the bithorax complex—are responsible for determining most aspects of segment identity. Interestingly, the order of genes in these complexes is the same as the order of the segments each gene controls.

Eric Wieschaus for his extensive genetic studies of the BX-C. In his work, Lewis isolated BX-C mutations that, like *bx* and *pbx,* affected the posterior thorax; he also found novel BX-C mutations that caused anteriorly directed transformations of each of the eight abdominal segments. Lewis named mutations affecting abdominal segments *infra-abdominal* (*iab*) mutations, and he numbered these according to the primary segment they affect. Thus, *iab-2* mutations cause transformations of A2 toward A1, *iab-3* mutations cause transformations of A3 toward A2, and so forth.

Researchers initiated molecular studies of the BX-C in the early 1980s, and in 15 years they not only extensively characterized all of the genes and mutations in the BX-C at the molecular level, but they also completed the sequencing of the entire 315 kb region. **Figure D.28** summarizes the structure of the complex. A remarkable feature of the BX-C is that mutations map in the same order on the chromosome as the anterior-posterior order of the segments each mutation affects. Thus, *bx* mutations, which affect anterior T3, lie near the left end of the complex, while *pbx* mutations, which affect posterior T3, lie immediately to their right. In turn, *iab-2,* which affects A2, is to the right of *pbx,* and is followed by *iab-3,* which affects A3.

Because the *bx, pbx,* and *iab* elements are independently mutable, Lewis thought that each was a separate gene. However, the molecular characterization of the region revealed that the BX-C actually contains only three protein-coding genes: *Ultrabithorax* (*Ubx*), which controls the identity of T3 and A1; *abdominal-A* (*abd-A*), which controls the identities of A2–A4; and *Abdominal-B* (*Abd-B*), which controls the identities of A5–A8 (Fig. D.28). The expression patterns of these genes are consistent with their roles. *Ubx* is expressed in segments T3–A8 (but most strongly in A1); *abd-A* is expressed in A2–A8 (most strongly in A2–A4); and *Abd-B* in A5–A8. The *bx, pbx,* and *iab* mutations studied by Lewis affect large *cis*-regulatory regions that control the intricate spatial and temporal expression of these genes within specific segments. The average size of each of these regulatory regions is about 15 kb. Researchers have not yet completely worked out the mechanisms by which single BX-C genes control multiple segment identities. However, it is likely that segment identity is determined primarily by the level or timing of BX-C gene expression.

The Antennapedia Complex and the Homeobox

Genetic studies in the early 1980s performed by Thomas Kaufman and colleagues showed that a second homeotic gene cluster, the Antennapedia complex (ANT-C), specifies the identities of segments in the head and anterior thorax of *Drosophila.* The five homeotic genes of the ANT-C are *labial* (*lab*), which is expressed in the intercalary region; *proboscipedia* (*Pb*), expressed in the maxillary and labial segments; *Deformed* (*Dfd*), expressed in the mandibular and maxillary segments; *Sex combs reduced* (*Scr*), expressed in the labial and T1 segments; and *Antennapedia* (*Antp*), expressed mainly in T2, although it is also active at lower levels in all three thoracic and most abdominal segments. (Figure D.19c shows these head and thoracic segments, while Figure D.27 illustrates the order of the homeotic genes in the ANT-C.) As with the BX-C, the order of genes in the ANT-C is the same (with the exception of *pb*) as the order of segments each controls. In addition to homeotic genes, the ANT-C contains several other important genes, including *zerknullt* (*zen*), which specifies dorsal embryonic structures; *fushi tarazu* (*ftz*), a segmentation gene of the pair-rule class; and *bicoid* (*bcd*), which, as we have seen, encodes the maternally supplied anterior determinant.

In the course of characterizing the ANT-C at the molecular level, researchers detected a region of homology between *Antp* and *ftz.* They subsequently found closely related sequences in all of the homeotic genes of the ANT-C and the BX-C, as well as in *zen* and *bcd* and many genes important for development located outside the homeotic gene complexes. In each gene, the region of homology, called the **homeobox,** is 180 bp in length and is located within the protein-coding sequence of the gene. The 60 amino acids encoded by the homeobox constitute a DNA-binding domain called the **homeodomain** that is structurally related to the helix-turn-helix motif of many bacterial regulatory proteins (review Fig. 19.16 of the main textbook). We now know that the homeotic proteins of the ANT-C and BX-C are transcription factors in which the

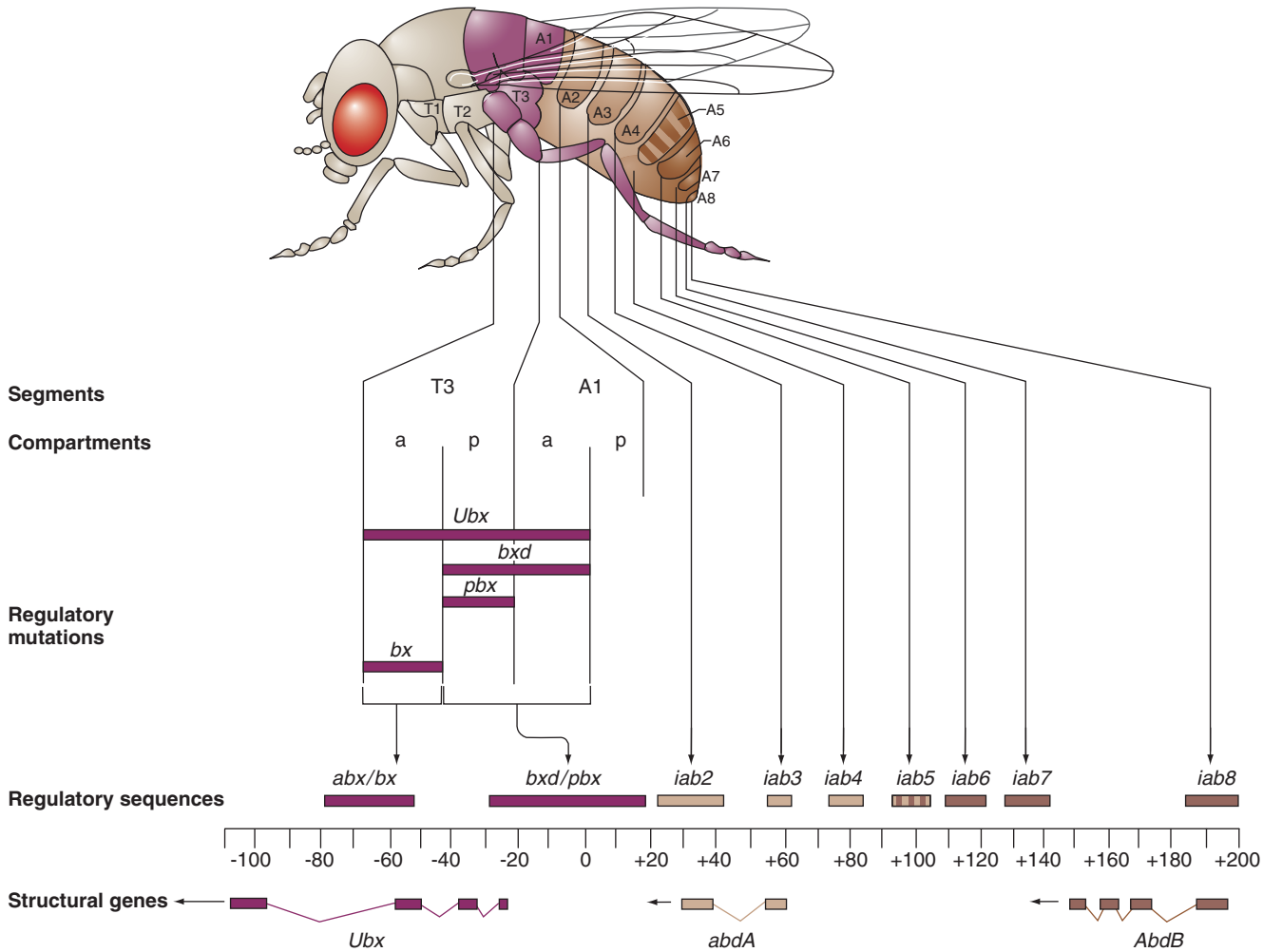


Figure D.28 **A close-up view of the 300 kb of the bithorax complex.** There are only three homeotic genes in this complex: *Ubx, abd-A,* and *Abd-B.* Many homeotic mutations such as *bx* and *pbx* affect regulatory regions that influence the transcription of one of these three genes in particular segments. For example, *bx* mutations (*at the left end of the complex*) prevent the transcription of *Ubx* in the anterior compartment of the third thoracic segment, while *iab-8* mutations (*far right*) affect the transcription of *Abd-B* in segment A8. Note that the order of these regulatory regions corresponds to the anterior-to-posterior order of segments in the animal.

homeobox-encoded homeodomain is responsible for the sequence-specific binding of the proteins to their target genes. Surprisingly, however, DNA-binding studies have shown that most of these homeodomains have very similar binding specificities. The homeodomains of the Antp and Ubx proteins, for example, bind essentially the same DNA sequences. Since different homeotic proteins are thought to regulate specific target genes, this lack of DNA-binding specificity seems paradoxical. Much current research is directed toward understanding how the homeotic proteins target specific genes that dictate different segment identities.

Connections

The discovery of the homeobox was one of the most important advances in developmental biology in the last two decades of the twentieth century. Some scientists compare the homeobox to the Rosetta stone (which led to the deciphering of Egyptian hieroglyphics) because it has enabled the isolation by homology of many other developmentally important genes in both *Drosophila* and other organisms. In the late 1980s and early 1990s, for example, the biological community was astonished to learn that the mouse contains clustered homeobox genes with clear homology to the ANT-C and BX-C genes of *Drosophila.* In the mouse, the homeobox genes are called *Hox* genes. The finding of

Hox genes in mice is yet another demonstration of the evolutionary conservation of gene function. Reference E presents a detailed discussion of homeobox genes in the mouse.

Some geneticists have voiced the opinion that the future of *Drosophila* research is limited because the four- to eightfold higher gene number in mice, humans, and other mammals means that mammals have numerous genes with no homologs in *Drosophila.* However, more recent studies indicate that this difference in gene number results mainly from the fact that mice and humans often have several closely related genes that all show homology to a single gene in flies. Although many of the multiple homologs in mice have undergone some functional divergence, in most cases, there is probably strong redundancy of function. This is certainly true of the mouse *Hox* genes. Since redundant genes are very difficult to detect by mutation, it is likely that *Drosophila*'s low gene number (and hence low level of genetic redundancy) will prove to be a major experimental advantage, rather than a disadvantage. The approach of identifying genes by mutation in *Drosophila,* followed by isolation and study of homologs in mice and humans, is likely to remain a central research paradigm. Thus, the evolutionary conservation of gene function combined with the many experimental advantages of the fruit fly ensure that *Drosophila* will remain at the forefront of biological research for the foreseeable future.

Essential Concepts

1. Several characteristics of *Drosophila* are useful for genetic analysis.
   a. The life cycle is relatively short, and each fly can produce thousands of progeny.
   b. A moderate amount of crossing-over occurs in females and no crossing-over occurs in males.
   c. *Balancer chromosomes* help preserve linkage.
   d. *P-element transposons* are tools of molecular manipulation useful for transformation, gene tagging, and enhancer trapping.
   e. *Genetic mosaics* resulting from mitotic recombination help track the roles of individual genes in the development of specific body structures.
   f. *Ectopic expression* can help pin down the function of a gene by making it possible to examine the effects of overexpressing the gene product or expressing the gene product in the wrong tissues.
2. The *Drosophila* Genome Project has already sequenced the euchromatic portion of the genome. Analysis of this sequence suggests that there are approximately 13,600 genes in the fly genome. A major remaining goal of the *Drosophila* Genome Project is to obtain information about the function of each of these genes. Large collections of transposon-induced mutations, as well as new techniques for targeted gene knockouts by homologous recombination and RNA intereference, should make the inactivation of each of these 13,600 genes possible in the near future.
3. The development of the anterior-posterior (AP) axis in the fruit fly depends on the coordinated action of the *segmentation genes,* which divide the body into three head, three thoracic, and eight abdominal segments; and the *homeotic genes,* which assign a unique identity to each segment.
4. The four classes of segmentation genes are expressed in the order maternal, gap, pair-rule, and segment polarity. The *maternal genes* produce gradients of *morphogens.* The *gap genes,* the first zygotic segmentation genes, are expressed in broad zones along the anterior-posterior axis. The *pair-rule genes* subdivide those broad areas into units that span two of the ultimate body segments. The *segment polarity genes* subdivide the two-segment units into individual segments. In the hierarchy of segmentation gene expression, each gene class is controlled by classes of genes higher in the hierarchy or by members of the same class.
5. The homeotic genes are master regulators of other genes that control the development of segment-specific structures. The *homeobox* in each homeotic gene encodes a *homeodomain* that allows the gene products to bind to specific target genes and control their expression. The homeotic genes are, in turn, regulated by the earlier acting gap, pair-rule, and segment polarity genes. Homeotic mutations cause particular segments, or parts of them, to develop as if they were located elsewhere in the body. Most homeotic mutations map to the bithorax and the Antennapedia complexes.
6. The finding of homeobox genes in other organisms such as the mouse has made it possible to identify developmentally important genes shared by all animals, demonstrating the evolutionary conservation of gene function.

Solved Problems

I. *Drosophila* geneticists try to minimize the work needed to maintain the many lines of flies they investigate. Ideally, they establish genetic stocks in which the flies in a culture bottle mate with each other so they produce the same kinds of progeny generation after generation. Because females homozygous for a maternal-effect lethal (*mel*) mutation are sterile (their offspring die during embryogenesis), such mutations can be propagated only when females are heterozygous. A particular *mel* mutation on the second chromosome (an autosome) is kept as a heterozygote with a balancer chromosome carrying the *Cy* mutation, which is a recessive lethal but also produces a Curly wing phenotype when heterozygous.

a. Suppose you start with equal numbers of male and female flies, both of genotype *mel/Cy.* What kinds of progeny would be produced in the next generation, and in what proportion? Would the culture bottles containing these progeny constitute a genetic stock?

b. What kind of flies would you select from the culture bottle to study the phenotypic effect of the maternal-effect mutation in detail?

Answer

This problem involves an understanding of balancer chromosomes and maternal-effect lethal mutations.

a. First determine all the ways gametes from these flies could combine and what the phenotype (including sex) would be. From a mating of *mel/Cy* females with *mel/Cy* males, the following kinds of zygotes would be formed: 1/8 *mel/mel* females (which survive but are sterile), 1/4 *mel/Cy* females (which survive and are fertile), 1/8 *Cy/Cy* females (which die before emerging as adults), 1/8 *mel/mel* males (which survive and are fertile because the mutation affects female but not male fertility), 1/4 *mel /Cy* males (which survive and are fertile), and 1/8 *Cy/Cy* males (which die before emerging as adults).

Of the adults emerging from these zygotes, you would find: 1/6 mel/mel *females (sterile with wild-type wings), 1/3* mel/Cy *females (fertile with Curly wings), 1/6* mel/mel *males (fertile with wild-type wings), 1/3* mel/Cy *females (fertile with Curly wings).* If all the flies in this culture bottle were allowed to mate at random, the same types of adult flies would emerge in the subsequent generation (you should work this out for yourself, remembering that only the heterozygous females are fertile). As a result, *this bottle constitutes a genetic stock: The same flies will be present in every generation.*

b. *You want* mel/mel *females, which can be recognized as the non-Curly (wild-type) females.* These females could be mated to any type of fertile male and the developing progeny examined to study the maternal effect lethal phenotype.

II. A particular gene in *Drosophila* called *no eclosion* (*nec*) encodes an enzyme needed for the fly to escape from the pupal case. The wild-type allele (nec^+) is dominant to a null mutant allele (nec^-). Animals homozygous for nec^- die because they cannot get out of the pupal case. The nec^- mutation (on an otherwise normal chromosome) is maintained in a stock over the autosomal balancer chromosome TM6C, which carries the nec^+ allele and two dominant markers: *Stubble,* which affects bristle shape in adults, and *Tubby,* which causes larvae and pupae to be short and squat. *Stubble* also behaves as a recessive lethal mutation, causing death during embryonic development (that is, before the larva hatches). Several deletions (A–D) in the vicinity of the *nec* gene were generated by X rays. Each of these deletions is maintained in a stock over the same TM6C balancer chromosome. Crosses are made between the nec^-/TM6C stock and each of the deletion/TM6C stocks. Crosses with deletions A and D yield only Stubble adult flies, while crosses with deletions B and C yield both wild-type (non-Stubble) as well as Stubble flies.

a. Which deletions remove the nec^+ gene?

b. If 1000 zygotes were generated from the crosses with deletions A and D that yielded only Stubble flies, how many of these zygotes should eventually produce adult flies?

c. Of the 1000 zygotes produced in part (b), how many would be expected to develop into animals that die in the pupal case? Would these pupal cases be short and squat (Tubby) or long and thin (wild type)?

d. From the preceding crosses with deletions B and C that yielded both wild-type (non-Stubble) as well as Stubble flies, what percentage of the adult flies that emerge should be non-Stubble?

e. *Stubble* and *Tubby* are located 10 m.u. apart. In a cross between nec^+/TM6C females and males true breeding wild type for all genes under consideration, how many of the 1000 adult progeny should have emerged from long, thin pupal cases and had Stubble bristles?

Answer

This problem involves an understanding of balancer chromosomes, the *Drosophila* life cycle, and complementation analysis.

a. In the cross between *nec*$^-$/TM6C and the deletion/TM6C strains, the following zygotes would be produced in equal proportions, each 25% of the total. For simplicity, these are labeled as classes 1–4.

Class 1	Class 2	Class 3	Class 4
nec$^-$ / deletion	*nec*$^-$ / TM6C	deletion / TM6C	TM6C / TM6C

Regardless of the relationship between *nec*$^+$ and a particular deletion, classes 2 and 3 should always survive, with animals developing as short, squat larvae and pupae (Tubby) that emerge as Stubble adults. In addition, class 4 should always die during embryonic development. Class 1 will form wild-type larvae, pupae, and adults if *nec*$^-$ and the deletion in question complement each other. If *nec*$^-$ and the deletion in question fail to complement each other, class 1 will not emerge from the pupal case; the pupal case itself will be long and thin.

Deletions A and D remove the nec *gene.* Because there are no wild-type adults in the matings with deletions A and D, we know that class 1 dies (these animals have no functional *nec*$^+$ product). No complementation occurred, so the deletions include the *nec* gene. With deletions B and C, there was complementation (class 1 survived) so the *nec* gene is not within the deleted areas.

b. We are looking here at crosses in which class 1 dies. In these crosses, zygotes of classes 1 and 4 would not develop into adults, but those of classes 2 and 3 would. Thus, *500 zygotes (1/2 of the total) would produce adult flies.*

c. Looking at the same kinds of crosses as in part (b), only class 1 would die in the pupal case. Thus, *1/4 of the zygotes (250 animals) would die as pupae.* These animals do not have the balancer chromosome, so *their pupal cases would be long and thin.*

d. Here, we are looking at crosses in which class 1 survives and produces wild-type animals. In these crosses, classes 1, 2, and 3 would produce adults. Of these, only class 1 should be wild type (non-Stubble), while classes 2 and 3 would have the *Stubble*-bearing TM6C balancer. Thus, of the emerged adults, *1/3 should be non-Stubble.*

e. *None.* Because of crossover suppression between the balancer chromosome and its homolog, viable recombinants that separate the *Tubby* and *Stubble* dominant markers on the balancer chromosome cannot be obtained.

Problems

D-1 For each of the terms in the left column, choose the best matching phrase in the right column.

a. morphogens	1. contains multiple overlapping inversions
b. chromocenter	2. reorganization of the body plan
c. balancer chromosome	3. making gene products in tissues where they are not normally made
d. enhancer trapping	4. define cell fate in a concentration-dependent manner
e. ectopic expression	5. heterochromatin in polytene chromosomes
f. metamorphosis	6. locating regulatory regions based on patterns of expression of a reporter gene

D-2 A balancer chromosome that is often used in *Drosophila* genetics is called TM3. This balancer chromosome is a variant of the third chromosome (an autosome), containing multiple inversions and the dominant mutation *Stubble,* which also has a recessive lethal phenotype. Ten different third chromosomes carrying various recessive lethal point mutations were isolated and kept in stock as TM3 heterozygotes (that is, lethal mutation/TM3). These stocks were then crossed to each other, and the number of wild-type progeny per 150 total adult progeny emerging from each of these crosses is recorded here.

Mutation	1	2	3	4	5	6	7	8	9	10
1	0									
2	47	0								
3	51	49	0							
4	45	54	50	0						
5	53	44	51	48	0					
6	62	47	39	59	39	0				
7	55	50	43	0	48	52	0			
8	53	44	42	0	60	54	0	0		
9	56	42	38	57	0	50	51	48	0	
10	0	58	59	38	46	49	51	49	55	0

a. Does this experiment measure complementation or recombination between recessive lethal mutations?

b. In the cross between mutant 1 and mutant 2, there are 103 adult progeny that are not wild type. What is their phenotype?

c. In the cross between mutant 1 and mutant 2, there are an additional 50 zygotes that never became adults (and therefore were not counted). What is the genotype of these zygotes?

d. How many genes are represented in this collection of 10 mutations? Which mutations are in the same genes?

e. Five different deletions in the third chromosome were isolated and balanced by TM3. These deletion

stocks were crossed to the TM3-balanced point mutation stocks described previously. The emergence of wild-type progeny is indicated in the following chart as +; the absence of wild-type progeny is indicated by −. Does this second experiment measure complementation or recombination between the recessive lethal mutations and the deletions?

	Point mutation number									
Deletion	1	2	3	4	5	6	7	8	9	10
A	+	−	+	−	+	−	−	−	+	+
B	−	−	−	−	+	−	−	−	+	−
C	−	+	−	+	−	+	+	+	−	−
D	+	+	−	−	+	−	−	−	+	+
E	−	+	−	−	+	+	−	−	+	−

f. Using this new information, show the best possible map for the 10 point mutations and 5 deletions in the region.

D-3 For a particular P-element mutagenesis experiment, you want to obtain new insertions of a P element marked with the w^+ gene into any of the autosomes. Starting with a stock of flies containing a w^+ marked P element on the X chromosome, design an experiment that would enable you to obtain flies containing P-element insertions in autosomes.

D-4 Which of the following animals would show mutant phenotypes in somatic cells as a result of a *Drosophila* cross involving hybrid dysgenesis?

a. progeny of a P male × M female cross
b. progeny of a M male × P female cross
c. progeny of the progeny of a P male × M female cross
d. progeny of the progeny of a M male × P female cross
e. progeny of the progeny of a P male × P female cross

D-5 Which of the following is not a property of the *hunchback* gene in *Drosophila?*

a. The *hunchback* mRNA is uniformly distributed in the egg by the mother.
b. Transcription of *hunchback* is enhanced by Bicoid (the anterior morphogen).
c. Translation of the *hunchback* mRNA is inhibited by Nanos (the posterior morphogen).
d. The Hunchback protein eventually is distributed in a gradient (anterior high; posterior low).
e. Hunchback protein directs the distribution of *bicoid* mRNA.

D-6 The *hunchback* gene contains a promoter region, the structural region (the amino-acid coding sequence), and a 3′ untranslated region (DNA that will be transcribed into sequences appearing at the 3′ end of the mRNA that are not translated into amino acids).

a. What important sequences required to control *hunchback* gene expression are found in the promoter region of *hunchback?*
b. There is another important kind of sequence that turns out to be located in the part of the gene transcribed as the 3′ UTR (untranslated region) of the *hunchback* mRNA. What might this sequence do?

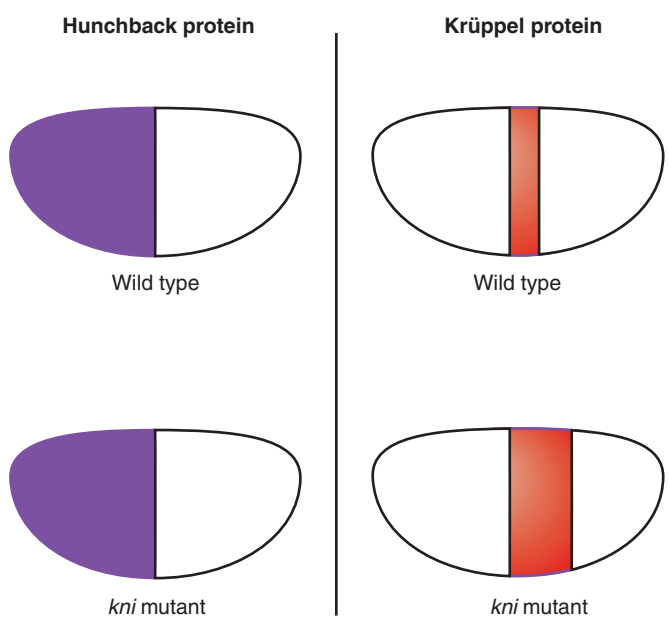
D-7 How do the segment polarity genes differ in their mode of action from the gap and pair-rule genes?

D-8 One important demonstration that Bicoid is an anterior determinant came from injection experiments analogous to those done by early embryologists. Injection experiments involve introduction of components such as cytoplasm from an egg or mRNA that is synthesized *in vitro* into the egg by direct injection. Describe injection experiments that would demonstrate that Bicoid is the anterior determinant.

D-9 In flies developing from eggs laid by a *nanos*$^-$ mother, development of the abdomen is inhibited. Flies developing from eggs that have no maternally supplied *hunchback* mRNA are normal. Flies developing from eggs laid by a *nanos*$^-$ mother that also have no maternally supplied *hunchback* mRNA are normal. If there is too much Hunchback protein in the posterior of the egg, abdominal development is prevented.

a. What do these findings say about the function of the Nanos protein and of the *hunchback* maternally supplied mRNA?
b. What do these findings say about the efficiency of evolution?

D-10 Mutant embryos lacking the gap gene *knirps* (*kni*) are stained at the syncytial blastoderm stage to examine the distributions of the Hunchback and Krüppel



proteins. The results of the *knirps*$^-$ and wild-type embryos stained for the Hunchback and Krüppel proteins is shown schematically on the previous page.

a. Based on these results, what can you conclude about the relationships among these three genes?
b. Would the pattern of Hunchback protein in embryos from a *nanos*$^-$ mutant mother differ from that shown? If yes, describe the difference and explain why. If not, explain why not.

D-11 In the analysis of eye development, mosaic *Drosophila* were constructed in which some cells were w^+ sev^+ and others were w^- sev^-. Some eye facets (made up of eight cells) contained both white and red cells. When the results from many different mosaic facets were tallied, all possible combinations of red and white cells were seen.

a. What does this indicate about the origin of cells recruited for any one facet?
b. What does this indicate about the expression of the *white* gene?

D-12 Nüsslein-Volhard and Wieschaus used 10 different mutant stocks of the form mutant/CyO where CyO is a *Drosophila* second chromosome balancer. Seven of these mutations were point mutations (mutations 1–7) and three of the mutations were cytologically detectable deletions (Del A, B, C), each removing a large number of genes. They crossed these stocks to each other and noted whether they obtained progeny without the *Cy* dominant marker on the CyO balancer chromosomes, and if so, whether such progeny females were fertile when crossed with wild-type males. The results of this analysis are tabulated here.

Key F: non-*Cy* progeny are obtained, and the females are fertile
S: non-*Cy* progeny are obtained, but the females are sterile
N: no non-*Cy* progeny are obtained

	Del A	Del B	Del C	1	2	3	4	5	6	7
Del A	N	N	N	N	F	S	S	F	F	F
Del B		N	F	F	N	F	S	F	F	N
Del C			N	N	F	F	F	S	S	F
1				N	F	F	F	F	F	F
2					N	F	F	F	F	N
3						S	F	F	F	F
4							S	F	F	F
5								S	S	F
6									S	F
7										N

a. Give the best genetic map of this region of the *Drosophila* second chromosome, showing the extent of the deletions with bars and positions of point mutations with an "X."
b. How many different genes are indicated by this data set? Which point mutations are in which genes?
c. How many genes needed for female fertility are indicated by this data set? Which point mutations are in which of these genes? How could you distinguish genes needed for oogenesis from maternal-effect lethal genes that make otherwise normal eggs that cannot develop into adults after fertilization?
d. How many zygotic lethal genes are indicated by this data set? Which point mutations are in which of these genes?
e. Why would it be difficult to determine the map distance between any two of the preceding point mutations that lie within the same gene?

D-13 A particular variant of a P element was constructed containing a selectable marker (the *rosy*$^+$ gene, where *rosy*$^-$ homozygotes have rose-colored eyes) and the *lacZ* gene, whose transcription is controlled by the weak P-element promoter. When the P element inserts into a gene, the gene is inactivated, producing a mutant phenotype. The insertion of these P elements was used to produce and isolate a series of female sterile mutants and study cell activities during oogenesis. Egg development occurs in an egg chamber containing a single oocyte and 15 nurse cells all derived from a germ cell precursor. The chamber is surrounded by somatic follicle cells. Follicle cells and oocyte and nurse cells interact in establishment of polarity in the egg chamber (review Fig. 19.26 in the main textbook).

a. What other information could you get about a developmental gene if the P element carrying the *lacZ* gene inserted into it?
b. How would you maintain stocks containing recessive female sterile mutations? (What would be the genotype of the stock flies and how could you study the mutants?)
c. One of the sterile mutants isolated contained occasional mispositioned oocytes within the egg chamber and some defective egg chambers. The defective gene, called *homeless,* encodes a protein that shows similarity to RNA-dependent ATPases and helicases, leading to the hypothesis that the protein may interact with, and help in the positioning of, particular RNAs. *bicoid* mRNA is a major anterior determinant for which localization is important. What experiment could you do to determine if *bicoid* mRNA is appropriately localized within the developing egg chambers in *homeless*$^-$ mutants? (Be sure to mention your control.)
d. What experiment could you perform to see whether *homeless* gene function is required either in the germ-line-derived cells (the nurse cells and/or oocyte) or is instead required in the somatic follicle cells? This experiment is made possible by the fact that various stocks are available in which the dominant ovo^D mutant gene, whose expression in germ-line-derived cells blocks

ovarian development, has been placed on different chromosome arms. Describe how you would set up this experiment and the expected results.

e. Another P-element-induced mutation isolated in the same screen was a semilethal female sterile. Inverse PCR experiments indicate that it maps to a chromosomal region where the *ras1* gene had been mapped. Complementation analysis established that this mutation is an allele of *ras1*. Other alleles of *ras1* affect viability, eggshell development, and eye morphology. Moderate alleles affect eggshell appearance and viability but not eye morphology. Different alleles also show different levels of Ras1 activity. What does this information suggest about Ras1 protein requirements in different signaling pathways?

D-14 A protein called Bruno was found to bind to the 3′ UTR of several developmental mRNAs, including *oskar* mRNA (necessary for proper formation of the posterior part of the embryo). When the *bruno* gene was cloned, the predicted protein sequence contained an RNA-binding domain consistent with the biochemical analysis of binding. The *bruno* gene mapped to a chromosomal region where the *arret* gene had previously been located. The *arret* gene, which was identified genetically in a screen for sterile mutants, is required for fertility in both sexes. How could you determine if *bruno* and *arret* are in fact the same gene that had been identified in different ways? What particular phenotypes would you examine?