

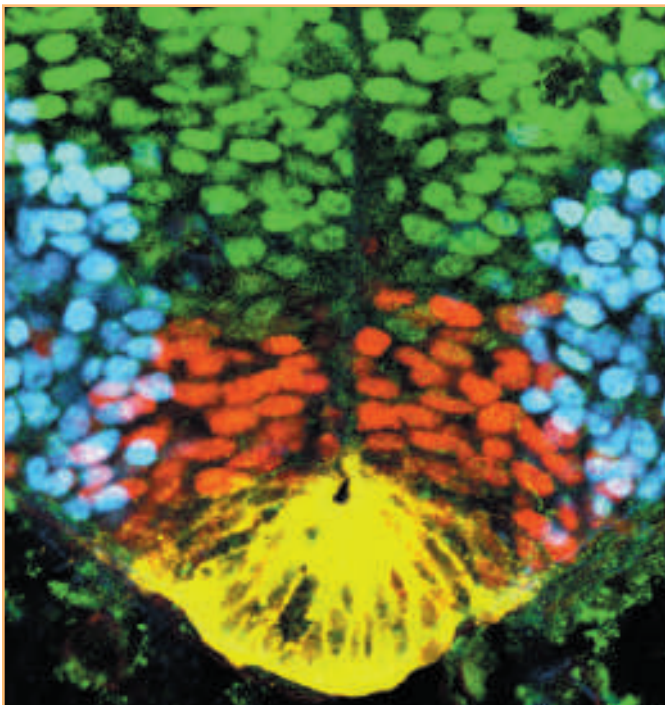
Gene Control in Development

The ultimate goal of developmental biologists is to unravel the mystery of how a fertilized egg is transformed into a complex multicellular organism. This process requires execution of a complex developmental program whereby specific genes are activated in a precise time sequence and in the correct location, generating different types of tissues and the specific cell types composing them. Classical and molecular geneticists have discovered numerous genes that participate in the highly regulated programs that result in the development of plants and animals. Understanding the molecular basis for the action of such genes is one of the most actively studied areas in all of biology.

In Chapters 10 and 11, we examined various mechanisms for regulating gene expression. As a result of these controls, cells can respond to changes in their environment and different cell types can produce characteristic proteins. By far the most prominent mechanism for regulating the production of different proteins in different cells entails a vast array of DNA-binding proteins that act, often in various combinations, to either activate or repress gene transcription.

In this chapter, we take a closer look at the spatial and temporal control of gene expression during development. Genetic and molecular studies show that different cells express different sets of genes based on their developmental history, their patterns of cell division, their position in the developing organism, and their interactions with other cells. We focus on several

well-studied cases of differential gene transcription to specify different cell types in yeast, animals, and plants to illustrate the general transcription-control strategies that regulate development. The interrelationships between transcriptional programs in development and signaling between cells are considered in Chapter 23.



Different proteins are expressed in different cells in the developing spinal cord.

OUTLINE

- 14.1 Cell-Type Specification and Mating-Type Conversion in Yeast 538**
- 14.2 Cell-Type Specification in Animals 543**
- 14.3 Anterioposterior Specification during Embryogenesis 553**
- 14.4 Specification of Floral-Organ Identity in *Arabidopsis* 571**

MEDIA CONNECTIONS



Overview: Gene Control in Embryonic Development

14.1 Cell-Type Specification and Mating-Type Conversion in Yeast

We begin our discussion of cell-type specification with the yeast *S. cerevisiae*. There are three different cell types in this unicellular eukaryote: haploid *a* and α cells, and diploid *a*/ α cells (see Figure 10-54). Because of the simplicity of yeast and the ease of studying it, our understanding of the transcription-control mechanisms specifying its three cell types is much more complete than our understanding of similar processes in animals and plants. It is likely that the mechanisms generating different cell types in the various organs and tissues in higher organisms evolved from mechanisms leading to diversification of cell types in simple unicellular organisms such as yeast.

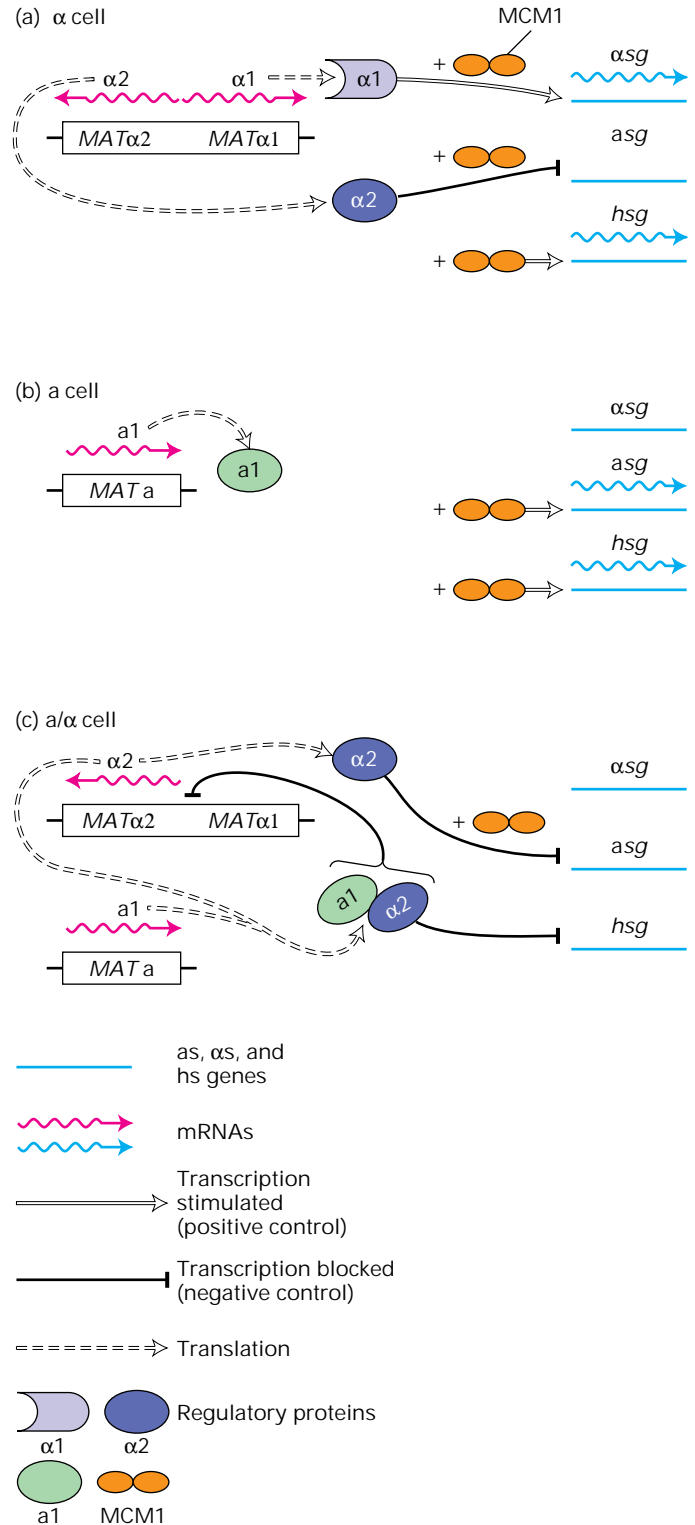
Combinations of DNA-Binding Proteins Regulate Cell-Type Specification in Yeast

Each of the three *S. cerevisiae* cell types expresses a unique set of genes. All haploid cells express certain haploid-specific genes; in addition, *a* cells express *a*-specific genes, and α cells express α -specific genes. In *a*/ α diploid cells, diploid-specific genes are expressed, whereas haploid-specific, *a*-specific, and α -specific genes are not. As illustrated in Figure 14-1, regulation of this cell type-specific transcription is mediated by three cell type-specific transcription factors ($\alpha 1$, $\alpha 2$, and *a1*) encoded at the *MAT* locus in combination with a general transcription factor called *MCM1*, which is expressed in all three cell types.

MCM1 was the first member of the MADS family of transcription factors to be discovered. The DNA-binding proteins in this family dimerize and contain a common homologous N-terminal MADS domain. (MADS is an acronym for the initial four factors identified in this family.) In later sections we will encounter other MADS transcription factors that participate in development of skeletal muscle and floral organs. As outlined in Figure 14-2, *MCM1* exhibits different activity in haploid *a* and α cells due to its association with $\alpha 1$ or $\alpha 2$ protein in α cells.

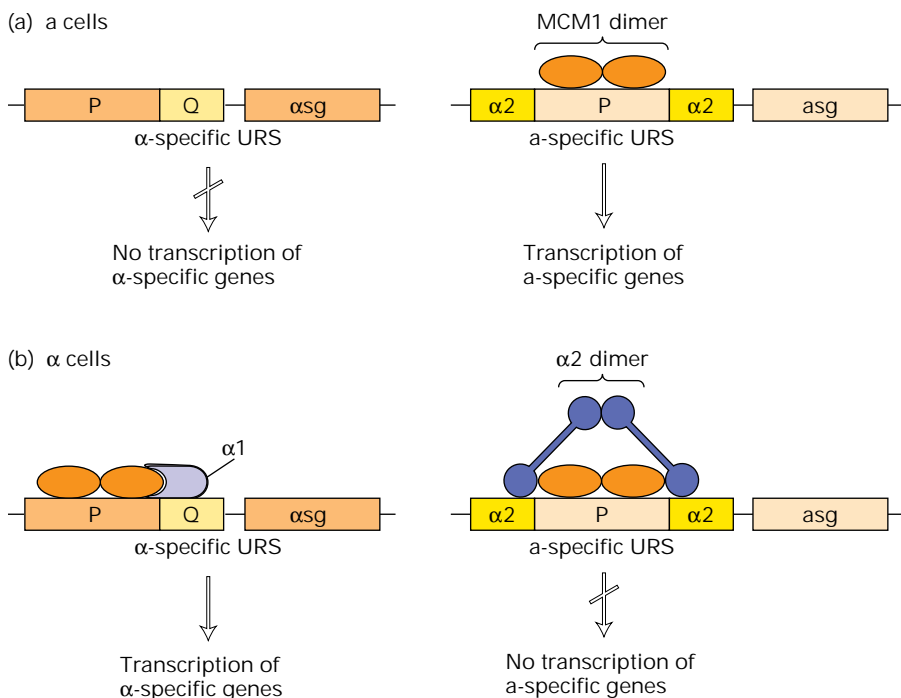
Gene Activation by *MCM1* and $\alpha 1$ -*MCM1* Complex In *a* cells, homodimeric *MCM1* binds to the P box in *a*-specific upstream regulatory sequences (URs), thereby stimulating transcription of the associated *a*-specific genes. The URs associated with α -specific genes contain two adjacent DNA sequences, the so-called P box and Q box. Although *MCM1* alone binds to the P box in *a*-specific URs, it does not bind to the P box in α -specific URs. Similarly, $\alpha 1$ does not bind alone to the Q box in an α -specific UR. The simultaneous binding of these proteins, however, occurs with high affinity and turns on transcription from the PQ site.

Gene Repression by $\alpha 2$ -*MCM1* and $\alpha 2$ -*a1* Complexes Flanking the P box in each *a*-specific UR are two $\alpha 2$ -binding sites. Both *MCM1* and $\alpha 2$ can bind independently



▲ FIGURE 14-1 Regulation of cell type-specific genes in *S. cerevisiae* by regulatory proteins encoded at the *MAT* locus together with *MCM1*, a constitutive transcription factor produced by all three cell types. As a result of this regulation, each cell type exhibits a distinctive pattern of gene expression: *as* = *a*-specific genes/mRNAs; αs = α -specific genes/mRNAs; *hs* = haploid-specific genes/mRNAs.

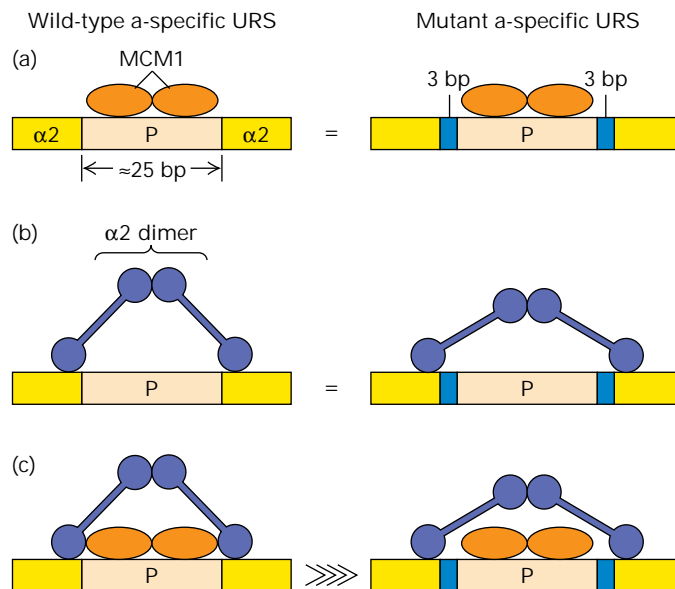
► **FIGURE 14-2 Activity of MCM1 in α and a yeast cells.** MCM1 binds as a dimer to the P site in α -specific and a -specific upstream regulatory sequences (URSs). These sequences lie upstream of and control transcription of α -specific genes and a -specific genes, respectively. The consensus sequences of the P box in α - and a -specific URSs differ slightly. (a) In a cells, MCM1 stimulates transcription of a -specific genes. MCM1 does not bind efficiently to the P site in α -specific URSs in the absence of $\alpha 1$ protein. (b) In α cells, the activity of MCM1 is modified by its association with $\alpha 1$ or $\alpha 2$. The $\alpha 1$ -MCM1 complex stimulates transcription of α -specific genes, whereas the $\alpha 2$ -MCM1 complex blocks transcription of a -specific genes. The $\alpha 2$ -MCM1 complex also is produced in diploid cells, where it has the same blocking effect on transcription of a -specific genes (see Figure 14-1c).



to an a -specific URS with relatively low affinity. However, the highly cooperative, simultaneous binding of both proteins occurs with high affinity. This high-affinity binding represses transcription of a -specific genes, ensuring that they are not expressed in α cells and diploid cells (see Figure 14-2b).

MCM1 promotes binding of $\alpha 2$ to an a -specific URS by orienting the two DNA-binding domains of the $\alpha 2$ dimer to the $\alpha 2$ -binding sequences in this regulatory sequence. Since a dimeric $\alpha 2$ molecule binds to both sites in an α -specific URS, each DNA site is referred to as a *half-site*. The relative position of both half-sites and their orientation are highly conserved among different a -specific URSs. Experiments with mutant URSs have shown that changing the orientation or spacing of the half-sites has little effect on the binding affinity of isolated $\alpha 2$ dimers in the absence of MCM1, suggesting that the two monomeric $\alpha 2$ subunits have considerable flexibility. However, the highly cooperative, high-affinity binding of $\alpha 2$ to an a -specific URS in the presence of MCM1 requires a precise spacing and orientation of the $\alpha 2$ half-sites (Figure 14-3). Presumably, the interaction between MCM1 and $\alpha 2$ constrains the flexibility of the $\alpha 2$ dimer, so that it binds with high affinity only to uniquely oriented and spaced $\alpha 2$ half-sites. Thus the affinity of $\alpha 2$ for sites in an a -specific URS is influenced by its association with MCM1.

The presence of numerous $\alpha 2$ -binding sites in the genome and the “relaxed” specificity of $\alpha 2$ protein may expand the range of genes that it can regulate. For instance, in a/α diploid cells, $\alpha 2$ forms a heterodimer with $\alpha 1$ that acts to repress both haploid-specific genes and the gene encoding $\alpha 1$ (see Figure 14-1c). The example of $\alpha 2$ suggests that



▲ **FIGURE 14-3 Relative binding affinities of MCM1, $\alpha 2$, and MCM1- $\alpha 2$ complex to wild-type and mutant a -specific upstream regulatory sequences (URSs).** (a,b) Insertion of three base pairs (blue) on either side of the P site does not affect the affinity of the independent binding of MCM1 to the P box or $\alpha 2$ to the flanking $\alpha 2$ -binding sites. (c) The high-affinity, cooperative binding of these two dimeric proteins requires correct spacing of the P site and $\alpha 2$ -binding site, as demonstrated by the much lower affinity of the MCM1- $\alpha 2$ complex for the mutant URS. [See D. L. Smith and A. D. Johnson, 1992, *Cell* 68:133.]

relaxed specificity may be a general strategy for increasing the regulatory range of a single transcription factor. Highly specific binding, then, occurs as a consequence of the interaction of $\alpha 2$ with other transcription factors at different sites in DNA.

As discussed in Chapter 10, eukaryotic repressor proteins exert their effects via several different mechanisms. The MCM1- $\alpha 2$ or $\alpha 1$ - $\alpha 2$ complex in yeast interacts with two proteins designated Tup1 and Ssn6, which do not themselves bind to DNA, to form a large complex that then represses transcription of many genes. Although this repressor complex clearly inhibits formation of a transcription-initiation complex at the promoter, the precise mechanism of transcriptional repression is not known.

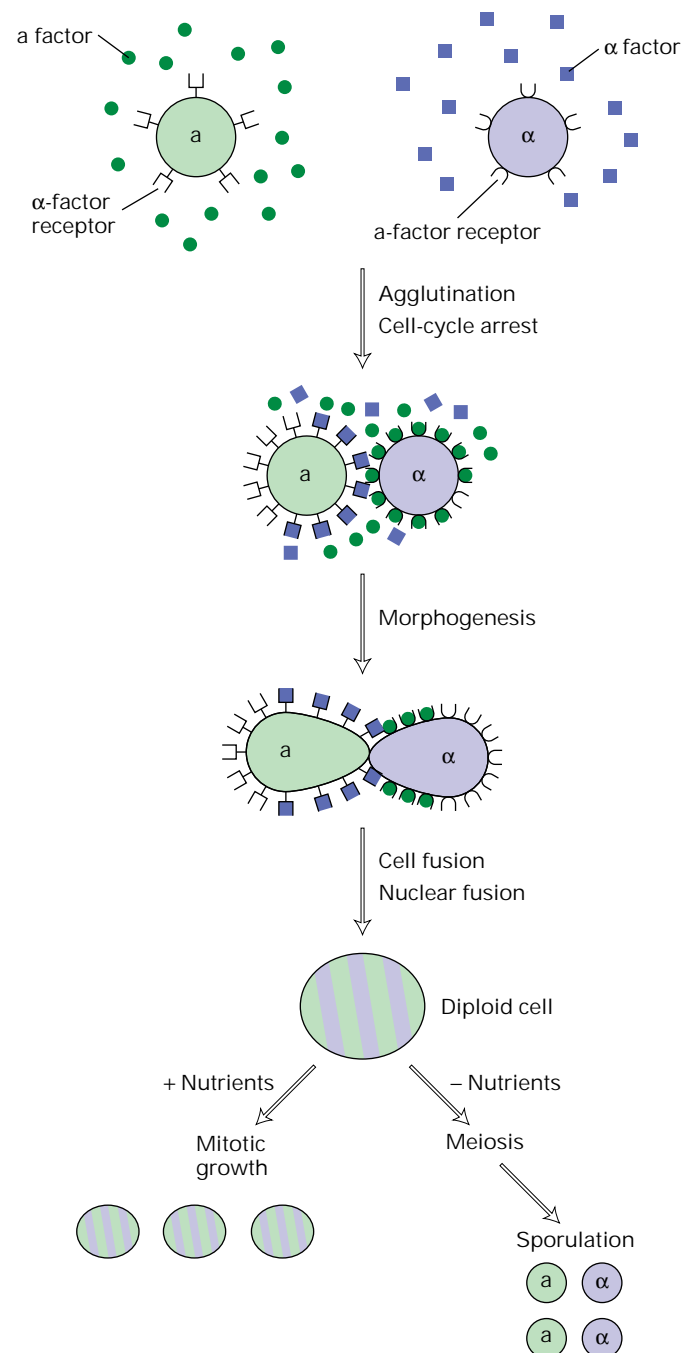
Mating of α and a Cells Is Induced by Pheromone-Stimulated Gene Expression

An important feature of the yeast life cycle is the ability of haploid a and α cells to mate, that is, attach and fuse giving rise to a diploid a/α cell. Each haploid cell type secretes a different mating factor, a small polypeptide called a *pheromone*, and expresses a cell-surface receptor that recognizes the pheromone secreted by cells of the other type. Thus a and α cells both secrete and respond to pheromones. Binding of the mating factors to their receptors leads to expression of a set of genes encoding proteins that direct arrest of the cell cycle in G_1 and promote attachment/fusion of haploid cells to form diploid cells. In the presence of sufficient nutrients, the diploid cells will continue to grow. Starvation, however, induces diploid cells to progress through meiosis, each yielding four haploid spores (Figure 14-4). If the environmental conditions become conducive to vegetative growth, the spores will germinate and undergo mitotic division.

Studies with yeast mutants have provided insights into how the a and α pheromones induce mating. For instance, haploid yeast cells carrying mutations in the sterile 12 (*STE12*) locus cannot respond to pheromones and do not mate. The *STE12* gene encodes a transcription factor that binds to a DNA sequence referred to as the *pheromone-responsive element* (PRE), which is present in many different a - and α -specific URSs. Binding of mating factors to cell-surface receptors induces a cascade of signaling events, resulting in phosphorylation of various proteins including the Ste12 protein (see Figure 20-31). This rapid phosphorylation is

correlated with an increase in the ability of Ste12 to stimulate transcription. It is not yet known, however, whether Ste12 must be phosphorylated to stimulate transcription in response to pheromone.

Interaction of Ste12 protein with DNA has been studied most extensively at the URS controlling transcription of *STE2*, an a -specific gene encoding the receptor for the α pheromone. Adjacent to this a -specific URS is a PRE that binds Ste12. When a cell is treated with α pheromone, transcription of *STE2* increases in a process that requires Ste12 protein.



► **FIGURE 14-4** Haploid yeast cells produce pheromones, or mating factors, and pheromone receptors. The α cells produce α factor and a receptor; the a cells produce a factor and α receptor. Binding of the mating factors to their cognate receptors on cells of the opposite type leads to gene activation, resulting in mating and production of diploid cells. In the presence of sufficient nutrients, these cells will grow as diploids. Without sufficient nutrients, cells will undergo meiosis and form four haploid spores.

Presumably, pheromone-induced up-regulation of the α receptor encoded by *STE2* increases the efficiency of the mating process. Ste12 protein has been found to bind most efficiently to PRE in the *STE2* URS when MCM1 is simultaneously bound to the adjacent P site. We saw previously that MCM1 can act as an activator or a repressor at different URSs depending on whether it complexes with $\alpha 1$ or $\alpha 2$. In this case, the function of MCM1 as an activator is stimulated by binding to yet another transcription factor, Ste12, whose activity is modified by extracellular signals.

As discussed in Chapter 20, surprising similarities have been uncovered between the mechanisms by which yeast respond to mating factors and higher eukaryotes respond to various extracellular factors that promote growth and differentiation.

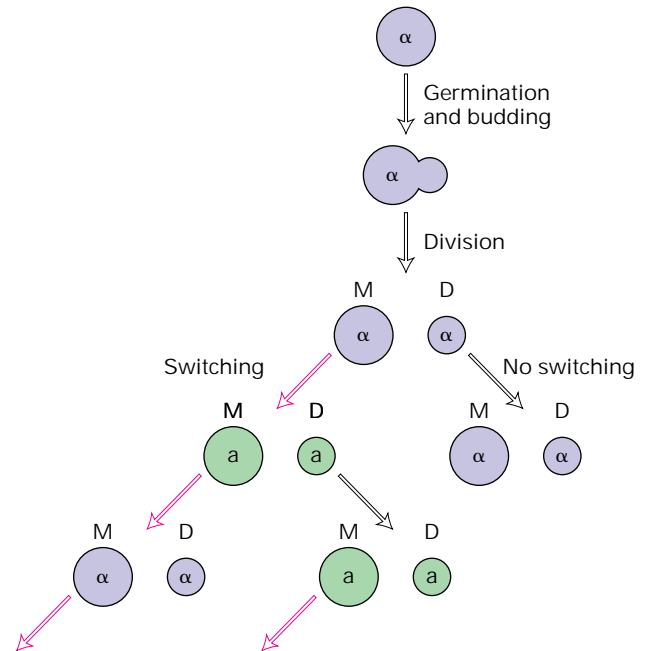
Multiple Regulation of *HO* Transcription Controls Mating-Type Conversion

Recall from Chapter 10 that two silent (nontranscribed) copies of the *MAT* locus—designated *HML* and *HMR*—are located on yeast chromosome 3 in addition to the active (transcribed) *MAT* locus. The phenotype of haploid yeast cells is determined by the mating-type sequence (*a* or α) that they carry in the central *MAT* locus. Once each generation, the sequences at *HML* or *HMR* are transferred to the central *MAT* locus, thereby converting an *a* cell to an α cell or vice versa (see Figure 10-55). This process of mating-type conversion begins with a site-specific cleavage at *MAT* by the *HO* endonuclease.

Mating-type conversion in yeast exhibits three types of specificity: it occurs only in haploid cells, during the late G_1 phase of the cell cycle, and in one of the two mitotic products, the so-called mother cell (Figure 14-5). This three-fold specificity results from the complex transcriptional regulation of the *HO* locus, which is controlled through two adjacent regulatory sequences—referred to as *URS1* and *URS2*—that lie ≈ 110 base pairs upstream of the *HO* locus. Switching occurs only in cells that express the *HO* endonuclease.

Transcription of the *HO* locus is repressed when a heterodimeric complex comprising $\alpha 2$ and $a 1$ binds to multiple sites within both *URS1* and *URS2*. These proteins are encoded by *MAT α* and *MAT a* , respectively (see Figure 14-1). Since both loci are present only in diploid cells, repression of *HO* transcription by the $\alpha 2$ - $a 1$ complex does not occur in haploid cells.

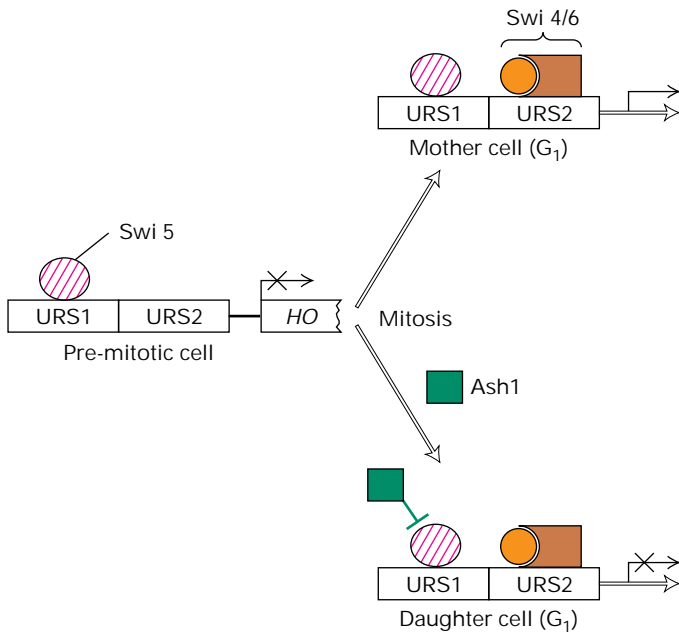
URS2 contains 10 repeats of a sequence, termed the *cell-cycle box*, that binds CCBF (cell-cycle box factor), a transcriptional activator. CCBF is composed of two proteins, designated Swi4 and Swi6, both of which are required for mating-type switching; they also help regulate the cell cycle-specific expression of other genes. Swi4 binds specifically to the cell-cycle box in the absence of Swi6, suggesting that Swi6 is necessary for the activating function of CCBF but not for its site-specific DNA-binding ability. The activity



▲ FIGURE 14-5 Specificity of mating-type conversion. Under appropriate conditions, haploid yeast spores germinate and undergo mitotic division by budding. The mother cell (M), which is larger than the daughter cell (D), can undergo a switch in the DNA sequences at the *MAT* locus before the next DNA duplication (i.e., during the G_1 phase). The first step in this gene conversion is catalyzed by *HO* endonuclease. Switching can occur in both directions ($a \rightarrow \alpha$ and $\alpha \rightarrow a$) but only in haploid mother cells in the G_1 phase of the cell cycle.

and/or expression of CCBF is responsive to Cdc28- G_1 cyclin complexes, whose protein kinase activity peaks in the late G_1 phase of the cell cycle (see Figure 13-26). Thus CCBF-mediated activation of *HO* transcription is limited to this phase of the cell cycle.

A key protein in restricting mating-type conversion to mother cells is Swi5, which binds to two short sequences within *URS1* and is required for transcription of the *HO* locus. Although *HO* transcription occurs only in late G_1 , Swi5 protein is synthesized in all stages of the cell cycle except G_1 . Presumably, Swi5 synthesized in the previous cell cycle is selectively functional in mother cells following division. The finding that both mother and daughter cells stain with antibodies to Swi5 suggests that Swi5 is inactivated in daughter cells or, alternatively, selectively activated in mother cells. Recent studies have identified an inhibitor of Swi5, called Ash1, that selectively segregates to daughter cells at mitosis (Figure 14-6). Whether Ash1 specifically interacts with Swi5, the *URS1*, or both, is not known. A set of genes have been identified that are necessary to promote Ash1 accumulation in daughter cells. In mutants lacking this machinery, neither mother nor daughter cells undergo mating-type switching because Ash1 accumulates in both cells.

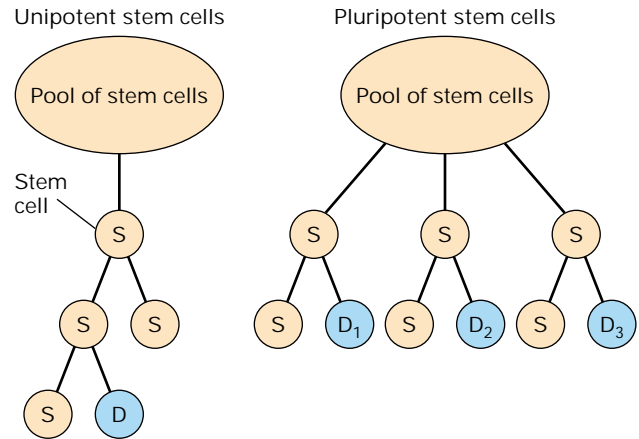


▲ FIGURE 14-6 Model for restriction of mating-type conversion to mother cells and the G₁ phase of the cell cycle. Binding of both Swi5 and the Swi4/6 complex (or CCBF) to the upstream regulatory regions (URS1 and URS2) is required for transcription of the *HO* locus, which encodes the endonuclease that initiates conversion. Swi4/6 is expressed or activated only in late G₁. Although Swi5 in pre-mitotic cells segregates to both mother and daughter cells, Ash1 (or perhaps its mRNA) segregates only to daughter cells. The Ash1-mediated inhibition of Swi5 activity prevents transcription of *HO* in daughter cells during the subsequent G₁ phase. The precise mechanisms by which this inhibition occurs are not known.

The asymmetric fate of two mitotic products, illustrated by the mother-cell specificity of mating-type conversion, is a common occurrence in developmental pathways of multicellular organisms. For example, many differentiated cells are generated from stem cells, which can divide asymmetrically to yield a stem cell plus a sibling cell that is more specialized and has lost some of its developmental potential (Figure 14-7). The mechanisms that have been shown to control mating-type conversion in yeast may provide insight into stem-cell development more generally.

Control of *HO* transcription, and hence of mating-type conversion, is even more complex than described above. Several other proteins, including the Swi1, 2, and 3 proteins, also are required for mating-type conversion. In addition, components of chromatin encoded by several *SIN* genes repress *HO* transcription, perhaps by maintaining the *HO* regulatory region in a configuration that prevents binding of the positively acting Swi proteins.

The molecular mechanisms by which these different levels of regulation converge on the *HO* locus to precisely



▲ FIGURE 14-7 The production of differentiated cells (D) from stem cells (S). Unipotent stem cells produce a single type of differentiated cell, whereas pluripotent stem cells may produce two or more types of differentiated cells.

control its transcription are not known. One model proposes that binding of Swi5 to URS1 promotes the activity of Swi1, 2, and 3. These proteins in turn somehow counteract the effect of the proteins encoded by the *SIN* genes, thereby permitting the binding of CCBF (Swi4/6) to URS2. Once Swi5 and CCBF are bound to the *HO* regulatory region in G₁, expression of the *HO* endonuclease and mating-type switching proceed (see Figure 14-6).

Silencer Elements Repress Expression at *HML* and *HMR*

As noted earlier, the *HML* and *HMR* loci on yeast chromosome 3 contain “extra” silent (nontranscribed) copies of the α and a sequences (see Figure 10-55). The location of the extra a and α sequence in *HML* and *HMR*, respectively, varies in different yeast strains. If these extra copies were transcribed during haploid growth, then haploid-specific genes would be repressed by the $\alpha 2$ - $a 1$ complex and the haploid cells could not mate (i.e., the haploid cells would be phenotypically like diploid cells). Hence, it is not only necessary to promote expression of genes required to specify one cell type, it is also necessary to repress other pathways leading to specification of different cell types.

As discussed in Chapter 10, *silencer elements* are responsible for specific repression of the a and α sequences associated with *HML* and *HMR*. Recent studies indicate that these elements, in conjunction with specific proteins, are required to assemble silencer-associated regions into higher-order chromatin structures inaccessible to the transcriptional machinery (see Figure 10-57). Similar mechanisms exist in higher eukaryotic cells, though the precise molecular mechanisms controlling this process are not as well understood as in yeast.

SUMMARY **Cell-Type Specification and Mating-Type Conversion in Yeast**

- Specification of each of the three yeast cell types—the *a* and α haploid cells and the diploid *a/\alpha* cells—is determined by a unique set of transcription factors acting in different combinations at specific cis-acting regulatory sites (see Figure 14-1).
- Some transcription factors can act as repressors or activators depending on the specific cis-acting regulatory sites they bind and the presence or absence of other transcription factors bound to neighboring sites.
- Chromatin structure can play an important role in regulating gene expression in development. In haploid cells, the opposite mating-type locus is silenced by packaging it into a higher-order chromatin structure inaccessible to transcriptional activators.
- Gene expression can be modified by specific extracellular signals through covalent modification (e.g., phosphorylation) of specific transcription factors. Binding of mating-type pheromones by haploid yeast cells activates expression of genes encoding proteins that mediate mating (see Figure 14-4).
- The asymmetric distribution of certain proteins during cell division can lead to changes in gene expression. The Ash1 protein, which is preferentially localized to daughter cells, prevents the Swi5 transcription factor from activating expression of the HO endonuclease and, hence, restricts the potential to switch mating types to mother cells (see Figure 14-6).

14.2 Cell-Type Specification in Animals

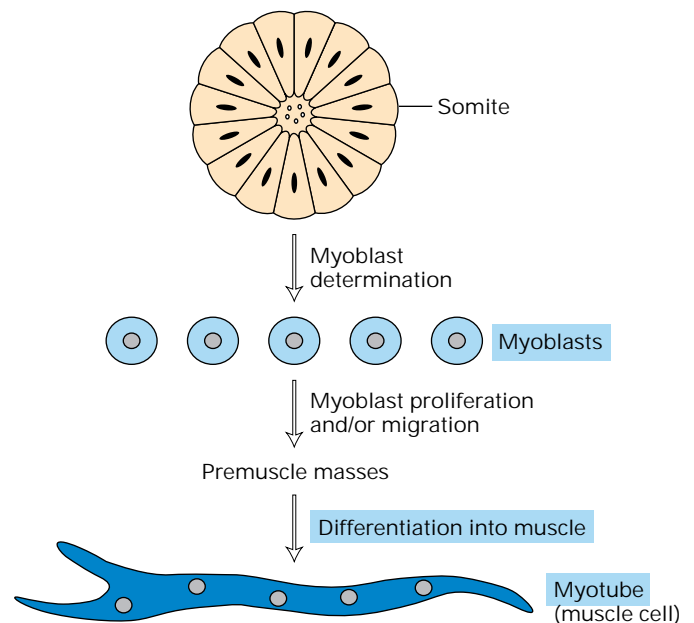
Each of the hundreds of different cell types found in multicellular organisms must be generated in the right number and in the appropriate region of the developing embryo and must be integrated into the framework of other cells to form discrete tissues. Specialized cells often have a distinctive morphology and express proteins devoted to the specific biochemical functions carried out by a particular cell type. The extensive cell specification that occurs during development of animals and plants depends on both quantitative and qualitative differences in gene expression, controlled largely at the level of transcription. An impressive array of molecular strategies, some analogous to those found in yeast cell-type specification, have evolved to carry out the complex developmental pathways that characterize multicellular organisms.

Cell biologists do not yet understand the complete set of regulatory molecules for any unique cell type in a multicellular organism that makes it different from other cells. In recent years, however, a family of related regulatory proteins have been shown to play analogous roles in the

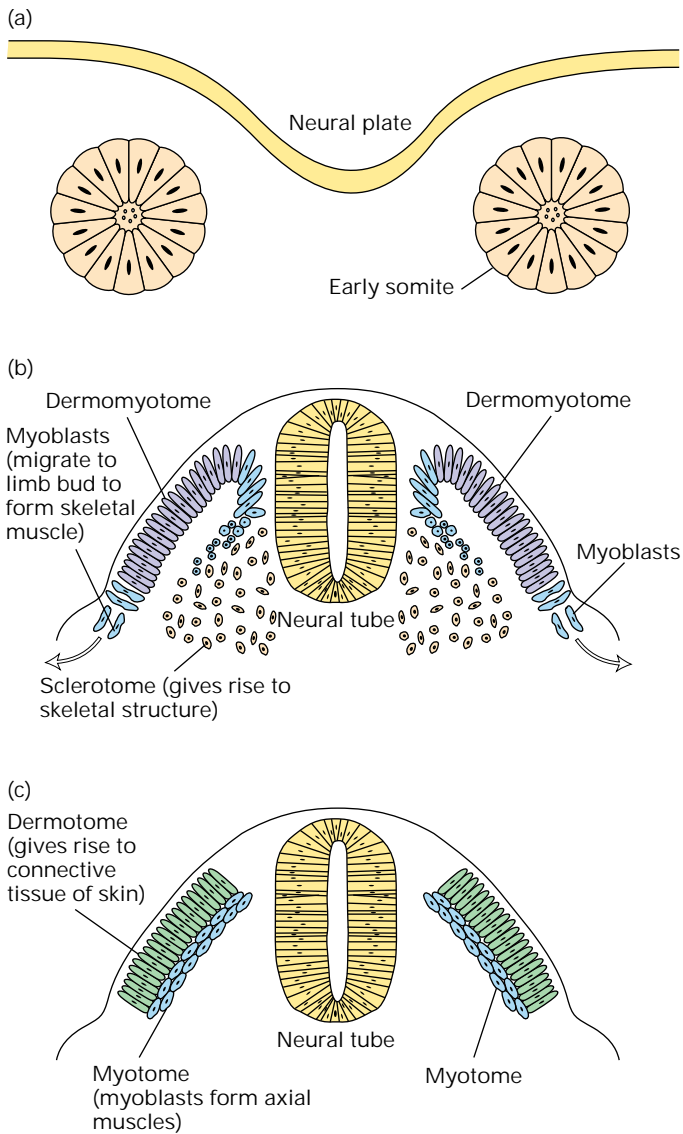
development of skeletal muscle cells (myogenesis), neural cells (neurogenesis), and perhaps other cell types. Mammalian skeletal muscle is a favorable system for investigating the role of transcription factors in controlling cell-type specification because its development can be studied both in the intact organism and in *in vitro* systems. In this section, we first describe the network of proteins that control specification of muscle cells and then consider how similar mechanisms operate during neurogenesis.

Embryonic Somites Give Rise to Myoblasts, the Precursors of Skeletal Muscle Cells

Mammalian skeletal myogenesis proceeds through three stages (Figure 14-8): determination of precursor muscle cells, called myoblasts; proliferation and in some cases migration of myoblasts; and their subsequent differentiation into mature muscle. In the first stage, myoblasts arise from blocks of mesodermal cells, called *somites*, found lateral to the neural tube in the embryo (Figure 14-9). Somites also give rise to tissues other than muscle including skeletal tissue and connective tissue in the skin. Specific signals from surrounding tissue including the neural tube and the lateral



▲ FIGURE 14-8 Schematic diagram of three stages in development of skeletal muscle in mammals. Somites are collections of embryonic mesodermal cells, some of which become determined as myoblasts. Myoblasts, which are precursor skeletal muscle cells, are distinct from other somite-derived precursor cells. A subclass of myoblasts migrates to form pre-muscle masses in the limbs and elsewhere, where they differentiate into multinucleate skeletal muscle cells, called myotubes. Other myoblasts will proliferate and differentiate to form axial muscles.



▲ **FIGURE 14-9 Embryonic determination and migration of myoblasts in mammals.** (a) Skeletal muscle is derived from embryonic structures called somites, which are blocks of mesodermal cells. (b) After formation of the neural tube, each somite forms a dermomyotome, which gives rise to skin and muscle, and a sclerotome, which develops into skeletal structures. Myoblasts form at each edge of a dermomyotome. Lateral myoblasts migrate to the limb bud. Axial myoblasts form the myotome. (c) The dermotome gives rise to skin elements (dermis), and the myotome to axial muscle. [See B. A. Williams and C. P. Ordahl, 1994, *Development* 120:785. Adapted from M. Buckingham, 1992, *Trends Genet.* 8:144.]

ectoderm play an important role in determining where myoblasts will form in the developing somite.

Myoblasts are committed to become muscle but have not yet differentiated; hence, they are referred to as *determined*. For instance, myoblasts that will form muscles in the limb migrate from the lateral side of the *myotome* to the

developing limb bud. Here the cells align, stop dividing, fuse to form a syncytium (a cell containing many nuclei but sharing a common cytoplasm), and differentiate into muscle. We refer to this multinucleate skeletal muscle cell as a *myotube*. Concomitant with cellular fusion there is a dramatic rise in the expression of genes necessary for muscle development and function. Other myoblasts from the more dorsal and medial regions of the myotome do not migrate and instead form cells of trunk muscles.

The specific extracellular signals that induce determination of each group of myoblasts are expressed only transiently. These signals trigger expression of numerous intracellular factors that can maintain the myogenic program in the absence of the inducing signals. We discuss the identification and functions of these myogenic proteins, and their interactions, in the next several sections.

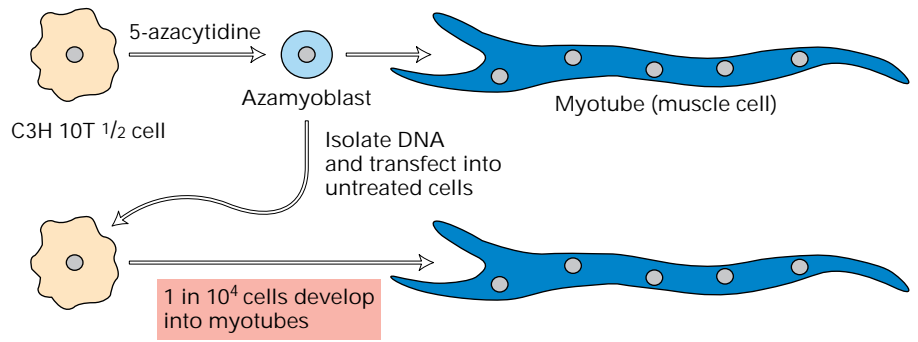
Myogenic Genes Were First Identified in Studies with Cultured Fibroblasts

In vitro studies with the fibroblast cell line designated C3H 10T have played a central role in dissecting the transcription-control mechanisms regulating skeletal myogenesis. When these cells are incubated in the presence of 5-azacytidine, a cytidine derivative that cannot be methylated, they differentiate into myotubes. Upon entry into cells, 5-azacytidine is converted to 5-azadeoxycytidine triphosphate and then is incorporated into DNA in place of deoxycytidine. As noted in Chapter 10, methylated deoxycytidine residues commonly are present in transcriptionally inactive DNA regions. Thus replacement of cytidine residues with a derivative that cannot be methylated may permit activation of genes previously repressed by methylation. The high frequency at which treated C3H 10T cells are converted into myotubes suggested to early workers that reactivation of one or a small number of closely linked genes is sufficient to drive a program of muscle development.

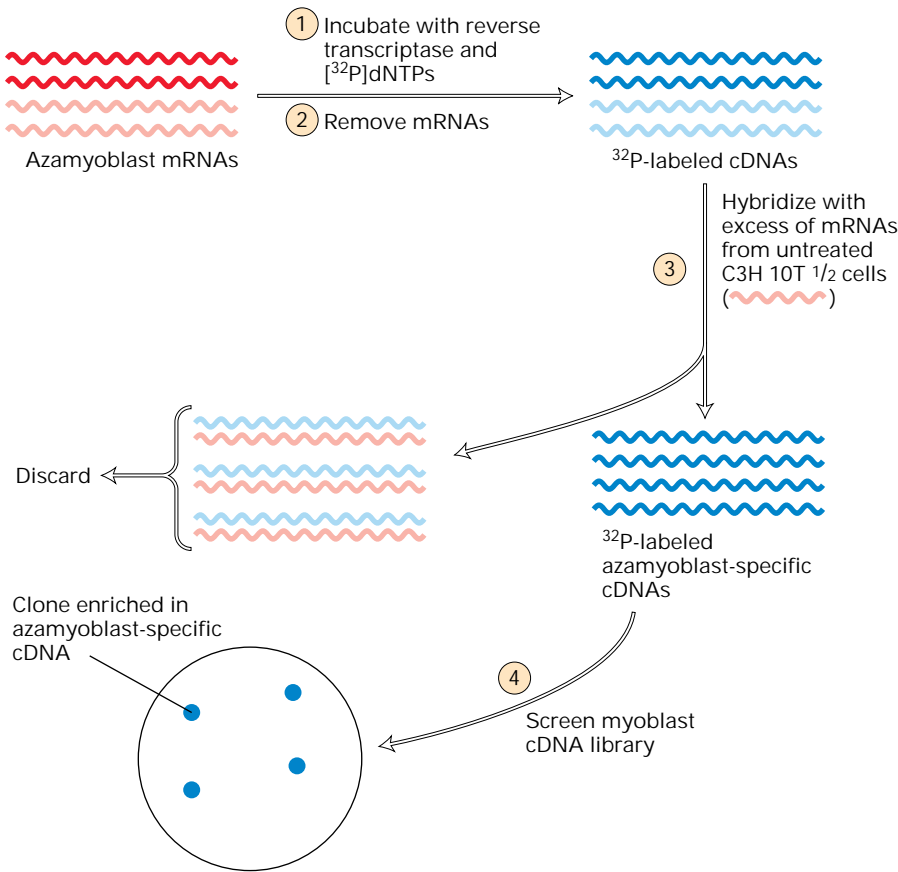
To test this hypothesis, researchers isolated DNA from C3H 10T cells grown in the presence of 5-azacytidine, so-called azamyoblasts, and transfected it into untreated cells (Figure 14-10). The observation that 1 in 10^4 cells transfected with DNA isolated from azamyoblasts was converted into myotubes was consistent with the hypothesis that one or a small set of closely linked genes is responsible for converting fibroblasts into myotubes.

Subsequent studies led to the isolation and characterization of four genes that can convert C3H 10T cells into muscle. Figure 14-11 outlines the experimental protocol for identifying and assaying one of these genes, called *myoD* for *myogenic determination gene D*. Colonies of *myoD*-transfected cells were indistinguishable from C3H 10T cells treated with 5-azacytidine, and both types of cells exhibited myotube-like properties. The *myoD* cDNA also was found to convert a number of other cell lines into muscle. Based on these findings, the *myoD* gene was proposed to play a key role in muscle development. A similar approach has identified three other genes—the *myogenin*, *myf5*, and *mrf4* genes—that also function in

► **FIGURE 14-10** Experimental system for studying mammalian myogenesis. A fibroblast cell line called C3H 10T can be converted into muscle cells by incubating them with 5-azacytidine. Under appropriate conditions, intermediate precursor cells, termed azamyoblasts, accumulate. DNA isolated from azamyoblasts can drive conversion of untreated C3H 10T cells into muscle cells. [See S. M. Taylor and P. A. Jones, 1979, *Cell* 17:771; A. B. Lasser et al., 1986, *Cell* 47:649.]

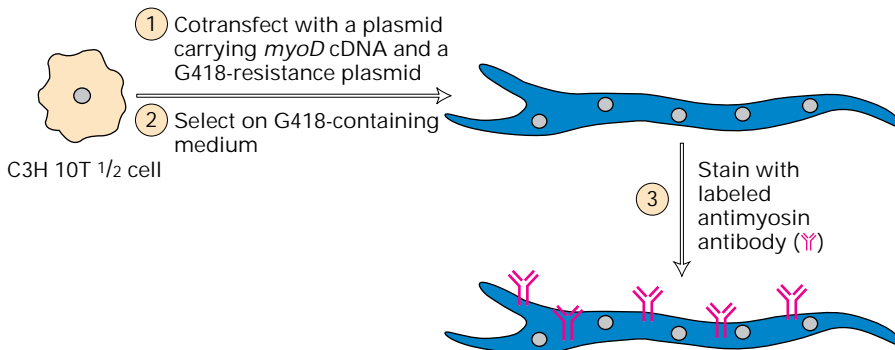


(a) Screen for azamyoblast-specific genes



◀ **FIGURE 14-11** Identification and assay of genes that drive myogenesis. (a) Azamyoblast mRNAs were isolated from cell extracts on an oligo-dT column (see Figure 7-14). Incubation with reverse transcriptase and [³²P]dNTPs yielded radiolabeled cDNAs. The cDNAs were mixed with mRNAs from *untreated* C3H 10T cells; only cDNAs derived from mRNAs (light red) produced by both azamyoblasts and untreated cells hybridized. This technique of subtractive hybridization yielded labeled azamyoblast-specific cDNAs (dark blue), at least some of which correspond to genes required for myogenesis. These cDNAs then were used as probes to screen an azamyoblast cDNA library. (b) Each of the cDNA clones identified as shown in part (a) was incorporated into a plasmid carrying a strong promoter. C3H 10T cells were cotransfected with each recombinant plasmid plus a second plasmid carrying a gene conferring resistance to an antibiotic called G418 and then selected on a medium containing G418. One of the clones, designated *myoD*, was shown to drive conversion of C3H 10T cells into muscle cells, identified by their binding of labeled antibodies against myosin, a muscle-specific protein. [See R. L. Davis et al., 1987, *Cell* 51:987.]

(b) Assay for myogenic activity of *myoD* cDNA



muscle development. As discussed in a later section, knockout experiments have demonstrated the importance of three of these genes in muscle development in the intact mouse.

Myogenic Proteins Are Transcription Factors Containing a Common bHLH Domain

The four myogenic proteins—MyoD, Myf5, myogenin, and Mrf4—are all members of the basic helix-loop-helix (bHLH) family of DNA-binding transcription factors (see Figure 10-44). Near the center of these proteins is a DNA-binding basic region adjacent to the HLH domain, which mediates dimer formation. Flanking this central DNA-binding/dimerization region are two activation domains. We refer to the four myogenic bHLH proteins collectively as muscle-regulatory factors, or MRFs (Figure 14-12a).

bHLH proteins form homo- and heterodimers that bind to a 6-bp DNA site with the consensus sequence C-A-N-N-T-G. Referred to as the *E box*, this sequence is present in many different locations within the genome (i.e., on a purely random basis the E box will be found every 256 nucleotides). Thus some mechanism(s) must ensure that MRFs specifically regulate muscle-specific genes and not other genes containing E boxes in their transcription-control regions. We will examine how this myogenic specificity may be achieved using MyoD as an example.

The affinity of MyoD for DNA is tenfold greater when it binds as a heterodimer complexed with E2A, another bHLH protein, than when it binds as a homodimer. Indeed, in developing azamyoblasts, MyoD is found as a heterodimer complexed with E2A, and E2A, as well as MyoD,

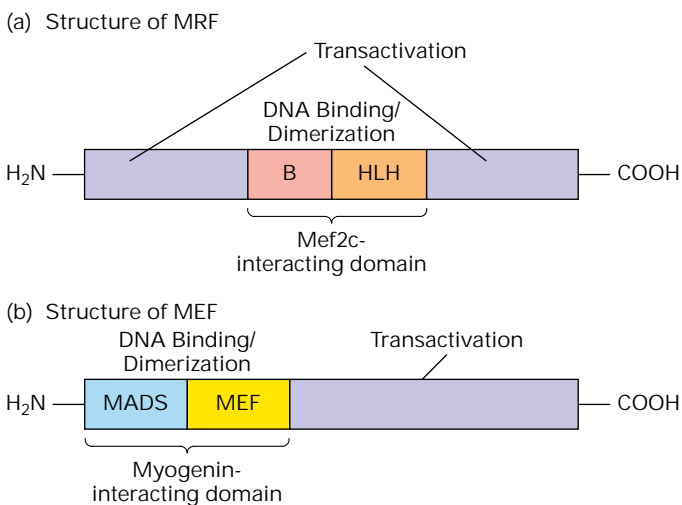
has been shown to be required for myogenesis in C3H 10T cells. The DNA-binding domains of E2A and MyoD have similar but not identical amino acid sequences, and both proteins recognize the E-box sequence in DNA. However, since E2A is expressed in other tissues, the requirement for E2A is not sufficient to confer myogenic specificity.

MyoD stimulates transcription only when two or more molecules of MyoD are bound to multiple E boxes; this multiple binding occurs cooperatively. Although multiple copies of E-box sequences are found in most muscle-specific enhancers, they also are present in enhancers that promote expression of genes specifically expressed in other tissues. For instance, E2A is required for normal development of B cells (the white blood cells that produce antibodies), and it regulates B-cell specific genes through enhancers containing multiple E boxes. Thus neither the requirement for multiple E boxes nor the requirement for E2A is sufficient to confer myogenic specificity. Rather, as in yeast cell-type specification, the key to myogenic specificity lies in the combinatorial association of different transcription factors at different transcription-control sites.

MEFs Function in Concert with MRFs to Confer Myogenic Specificity

Some insight into how skeletal muscle cells are specified has come from in vitro mutagenesis studies in which variant E2A proteins were produced. Wild-type E2A protein cannot by itself drive C3H 10T cells to a myogenic fate, although it can bind to E boxes controlling muscle-specific genes. To identify which features of MyoD confer myogenic specificity, researchers produced altered E2A molecules in which specific amino acids present in MyoD were substituted into the equivalent positions in the E2A molecule. An E2A variant with three amino acid substitutions corresponding to residues present in MyoD was found to convert C3H 10T cells to myotubes. Two of these substitutions are in the central basic DNA-binding region of E2A, and one is just adjacent to this region. Although these substitutions allow E2A to drive myogenic conversion, they do not affect the DNA-binding specificity of E2A. This finding suggests that myogenic specificity is likely to reside in specific interactions between MyoD and other proteins. Recent studies indicate that specific amino acids in the bHLH domain of all the MRFs may confer myogenic specificity by allowing MRF-E2A complexes to bind specifically to another family of DNA-binding proteins called muscle enhancer-binding factors, or MEFs.

First identified in biochemical experiments as homodimeric proteins that bind to certain DNA sequences in a muscle-specific enhancer, MEFs later were shown to belong to the MADS family of transcription factors. In addition to containing an N-terminal MADS domain, these proteins contain a short stretch of amino acids just C-terminal to the MADS domain, called the MEF domain, and a C-terminal transcription-activation domain (Figure 14-12b). MEFs were considered excellent candidates for interaction with MRFs for two reasons: First, many muscle-specific genes contain



▲ FIGURE 14-12 Schematic diagrams of the general structures of two classes of transcription factors that participate in myogenesis. MRFs are produced only in muscle, whereas MEFs are expressed in several tissues in addition to developing muscle. The myogenic activity of MRFs is enhanced by their interaction with MEFs.

both MEF- and MRF-recognition sequences; second, although MEFs cannot induce myogenic conversion of C3H 10T cells, they enhance the ability of MRFs to do so. This enhancement requires physical interaction between a MEF and MRF-E2A heterodimer.

The interaction between these different transcription factors requires both the MADS and MEF domains in the MEF homodimer and the myogenic-specific amino acids in the bHLH domain of the MRF-E2A heterodimer. Crystallographic analysis of MRF-E2A bound to DNA indicates that these amino acids are buried within the major groove of DNA and are unable to directly contact MEFs. Hence, it seems likely that these amino acids confer a particular conformation to other regions of the MRF-E2A heterodimer that, in turn, specifically interact with MEFs. Surprisingly, although both classes of proteins can bind individually to specific DNA sequences, a single DNA site that recognizes one or the other protein is sufficient to act as a platform for assembly of a MRF-E2A-MEF complex. This finding suggests that different configurations of the DNA sites recognized by these factors may drive high levels of muscle-specific gene expression. Indeed, some muscle-specific genes contain bHLH-binding sites (i.e., E boxes); others contain MEF-binding sites; and yet others contain both types of protein-binding sites (Figure 14-13). In

contrast to MRFs, which are expressed only in developing muscle, MEFs are expressed in other tissues including the developing central nervous system.

Myogenic Stages at Which MRFs and MEFs Function in Vivo Have Been Identified

Expression of any one of the four MRFs in C3H 10T cells can induce the cells to differentiate into muscle in vitro. The functions of these proteins in the intact animal during normal myogenesis have been studied in gene-knockout experiments. In these studies, mice were prepared with gene-targeted knockout mutations in the genes encoding MyoD, Myf5, or myogenin (see Figure 8-34). By analyzing the effects of knocking out these genes, developmental biologists could determine which genes are required for myogenesis and the stage at which they act.

Mice with either the *myoD* or *myf5* gene knocked out have normal muscle, whereas those with the *myogenin* gene knocked out are missing the vast majority of skeletal muscle (Table 14-1). In mice that lack myogenin, myoblasts accumulate at sites normally occupied by skeletal muscle, indicating that myogenin is not required for formation of myoblasts but is required for their differentiation into myotubes. The simple, but erroneous, conclusion from these findings is that

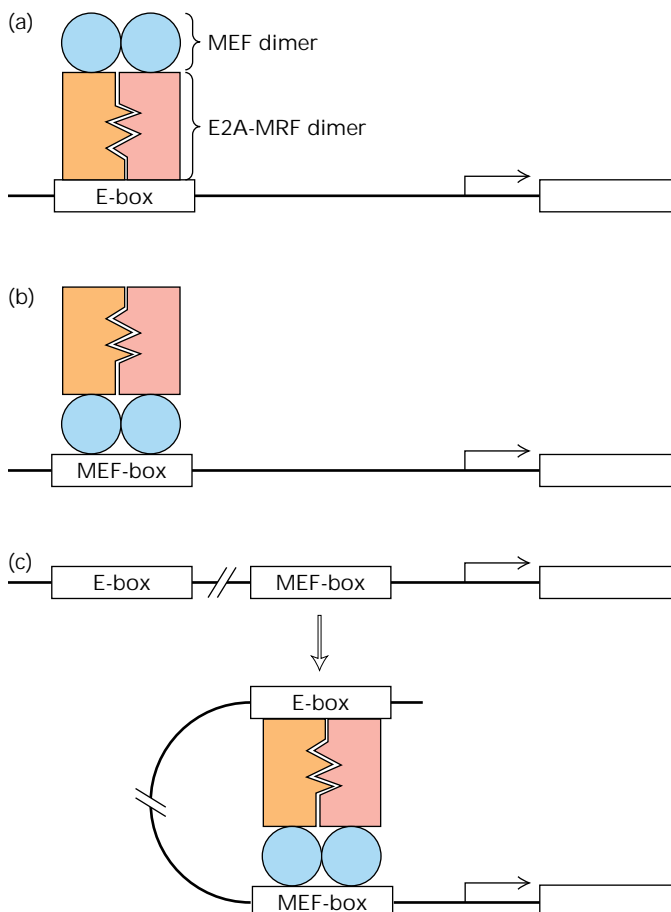


TABLE 14-1 Effect of Knockout of Myogenic Genes in Mice

Gene Knocked Out	Phenotype*			Role of Myogenic Protein
	Viable	Myoblasts	Muscle	
<i>myoD</i>	Yes	+	+	?
<i>myf5</i>	Yes	+	+	?
<i>myoD</i> ; <i>myf5</i>	No	–	–	Required for myoblast formation or survival
<i>myogenin</i>	No	+	–	Required for myoblast differentiation into muscle

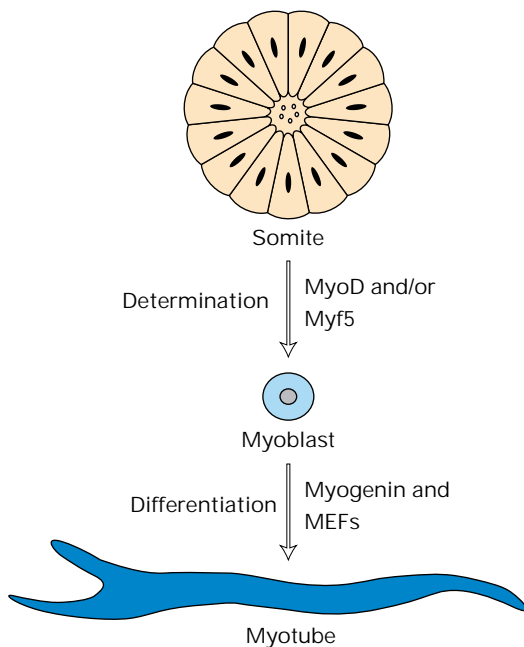
*+ sign indicates that myoblasts or mature muscle cells are found at normal sites; – sign indicates that they are not.

SOURCE: T. G. Braun et al., 1992, *Cell* 71:369; P. Hasty et al., 1993, *Nature* 364:501; M. A. Rudnicki et al., 1992, *Cell* 71:383; M. A. Rudnicki et al., 1993, *Cell* 75:1351.

◀ FIGURE 14-13 An MRF-E2A-MEF complex can assemble in the transcription-control region of muscle-specific genes containing an E box, MEF box, or both. The synergistic action of the MEF homodimer and MRF-E2A heterodimer, which directly interact, drives high-level expression of muscle-specific proteins. [Adapted from K. Yun and B. Wold, 1996, *Current Opinion in Cell Biology* 8:877.]

Myf5 and MyoD are not required for muscle development. However, since either protein can drive a myogenic program in cell culture, the loss of one gene may be compensated by the function of the other. Indeed, mice homozygous for mutations in both *myf5* and *myoD* die shortly after birth and lack skeletal muscle. In contrast to the *myogenin* mutants, myoblasts do not accumulate in the *myf5*; *myoD* double mutants, suggesting that the Myf5 and MyoD proteins are required for the formation or survival of myoblasts. Overlapping functions such as these of MyoD and Myf5 are often referred to as redundant. Redundancy provides a more robust developmental program and may allow for more flexibility in the response of cells in different regions of the developing organism to extracellular signals regulating myogenesis.

The results of these gene-knockout experiments are consistent with the observation that azacytidine-treated C3H 10T cells express Myf5 and MyoD prior to fusion but express myogenin only as they fuse to form a syncytium, which then differentiates to form a myotube. As the model in Figure 14-14 illustrates, MyoD and Myf5 are thought to have similar but overlapping functions in selecting cells from developing somites to become myoblasts; that is, they are required for myoblast determination during normal myogenesis. Myogenin, then, is required for the differentiation of myoblasts into myotubes. The fourth MRF protein, Mrf4, is expressed later in development and may play a role in the maintenance of muscle cells.



▲ FIGURE 14-14 Model of genetic control of mammalian skeletal muscle in vivo based on knockout experiments in mice and loss-of-function mutations in *Drosophila*. According to this model, MyoD and Myf5 serve a redundant function in myoblast determination, while myogenin and MEFs have distinct functions in the differentiation of myoblasts into myotubes (see text).

The role of E2A and MEFs in myogenesis have been assessed in more recent studies. Muscle development is normal in mice with a knockout mutation in the gene encoding E2A, although B-cell development is disrupted. Presumably, during muscle development, E2A-related genes may compensate for loss of E2A, much as *myoD* and *myf5* can compensate for each other. To assess this redundancy, researchers will have to knock out the E2A-related genes and generate mice lacking both E2A and its related genes.

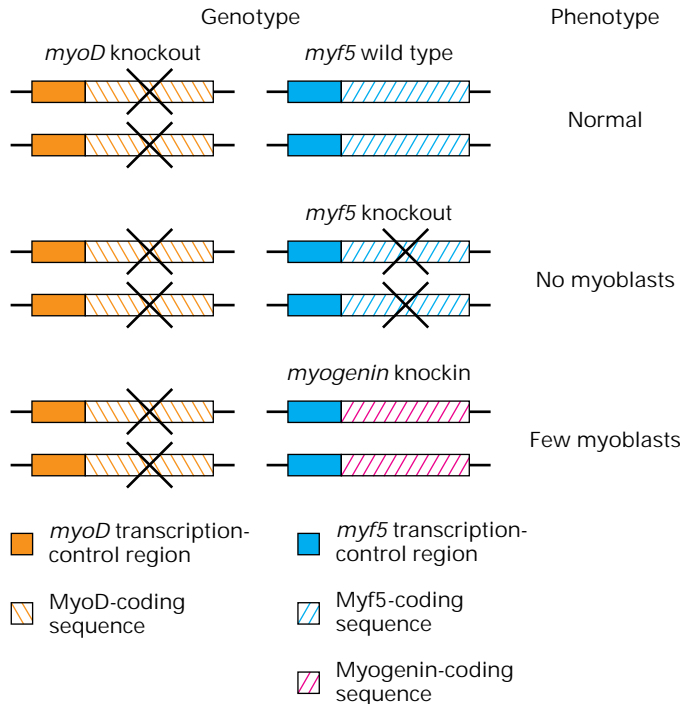
Because mice express multiple MEF proteins, scientists turned to *Drosophila*, which expresses a single MEF, to determine the function of MEFs in muscle development. In flies carrying loss-of-function mutations in the *MEF* gene, no differentiated muscle forms, although myoblasts appear to form normally. Hence, MEFs are required for differentiation, but not for determination (see Figure 14-14).

Multiple MRFs Exhibit Functional Diversity and Permit Flexibility in Regulating Development

The expression of four myogenic bHLH proteins (MRFs) in mice raises intriguing questions. Do these proteins have intrinsically different biochemical properties that correlate with distinctive roles in muscle development? That is, did functionally different MRFs evolve independently? Or have multiple MRFs evolved to facilitate the demands of gene expression in more complex organisms? That is, was duplication of an ancestral MRF gene and the subsequent evolution of divergent transcription-control elements more efficient than incorporation of different control elements into a single gene? Since many mouse genes that regulate development are found in multiple copies, understanding the role of the apparent duplication of the MRFs may provide generally applicable insights about developmental processes.

Scientists have begun to assess these possibilities using a variation of gene-knockout technology called knockin. In this technique, the coding sequences of one gene (e.g., *myf5*) are replaced by those of another (e.g., the *myogenin* gene). The experiments combining knockout and knockin technology depicted in Figure 14-15 demonstrate that myogenin and Myf5 are not functionally equivalent in vivo. Indeed, recent biochemical studies have shown that the chromatin-remodeling ability of Myf5 (and MyoD) is much greater than that of myogenin. As we discuss in more detail later, remodeling of chromatin is critical for normal development of most tissues.

As noted earlier, mice with a homozygous knockout of the *myogenin* gene accumulate myoblasts and are not viable (see Table 14-1). By creating mice homozygous for the *myogenin* knockin at the *myf5* locus (i.e., under control of the *myf5* regulatory sequences), scientists could assess the importance of the myogenin-specific transcriptional regulation. The failure of this knockin to rescue the myogenin defect indicated that the unique expression pattern conferred by the



▲ **FIGURE 14-15** Experimental demonstration that myogenin cannot substitute for Myf5 in vivo. Expression of either Myf5 or myogenin in C3H 10T cells can drive myogenesis. To test whether these proteins are functionally equivalent in vivo, researchers inserted the myogenin-coding sequences in place of the sequences encoding Myf5 within the *myf5* gene, forming a *myogenin* knockin. Myogenin will be expressed from the knockin gene in the same spatiotemporal fashion as Myf5 in a wild-type mouse. In knockout mice lacking MyoD and carrying the *myogenin* knockin in place of wild-type *myf5*, some muscle fibers formed (about 10 percent of normal), but the mice died. The inability of the *myogenin* knockin to correct the mutant phenotype resulting from the lack of both Myf5 and MyoD establishes that the functional properties of myogenin and Myf5 differ in vivo.

myogenin regulatory sequences is also critical for normal development. In summary, these studies suggest that gene duplication led to evolution of genes encoding functionally diverse MRFs whose expression is regulated by different transcription-control elements.

Terminal Differentiation of Myoblasts Is under Positive and Negative Control

The determined, yet undifferentiated, myoblast can respond to extracellular signals in the developing embryo that control proliferation (hence, the number of cells that form) and cell migration (hence, the precise location of muscle). In contrast, the differentiated muscle cell, or myotube, cannot respond to such signals. Regulation of the transition from the determined to the differentiated state thus permits the precise spatial and temporal control of cellular differentiation that is necessary to ensure normal morphogenesis in

complex multicellular organisms. The factors that regulate this critical step in various developmental pathways are still poorly understood. However, in vitro experiments have revealed several specific factors that promote or inhibit differentiation during myogenesis.

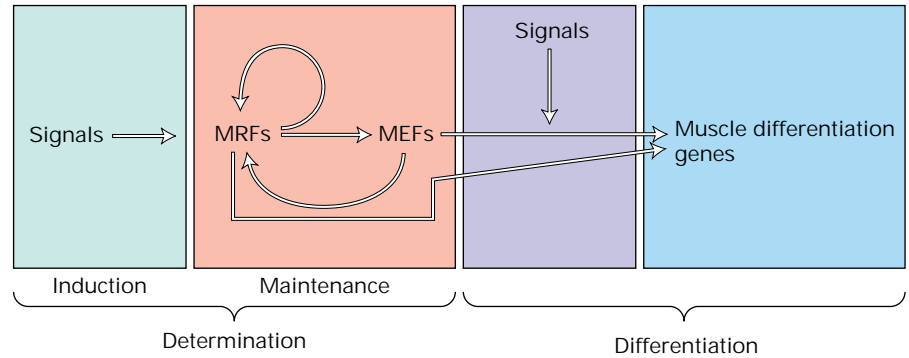
Inhibitory Proteins Screens for genes related to *myoD* led to identification of a related protein that retains the dimerization helices but lacks the DNA-binding basic region and hence is unable to bind to E-box sequences in DNA. However, this protein interacts with MyoD and E2A, thereby inhibiting formation of MyoD-E2A heterodimers and hence their high-affinity binding to DNA. Accordingly, this protein is referred to as *Id* for *inhibitor of DNA binding*. Analysis of DNA from proliferating azamyoblasts, which express MyoD, E2A, and *Id*, has shown that the MyoD-binding (or E2A-binding) site in the promoter of the muscle-specific gene encoding creatine kinase is not occupied. This finding presumably reflects the formation of inactive MyoD-*Id* or E2A-*Id* complexes and indicates that *Id* can maintain cells in a determined state during proliferative growth. When these cells are induced to differentiate into muscle (for instance by the removal of serum-containing growth factors required for proliferative growth), the *Id* concentration falls. As a result, MyoD-E2A dimers can form and bind to the promoters of target genes driving differentiation of azamyoblasts into myotubes. We can see from these results that dimerization of transcription factors with different partners not only can modulate the specificity or affinity of their binding to specific DNA sites, but also may prevent their binding entirely.

Cell-Cycle Proteins The onset of differentiation in many cell types is associated with arrest of the cell cycle, most commonly in G_1 , suggesting that cell-cycle proteins (e.g., cyclins and Cdks) may influence the transition from the determined to differentiated state. Researchers recently have found that certain inhibitors of cyclin-Cdk protein kinase activity can induce muscle differentiation in cell culture and that these inhibitors are markedly up-regulated in differentiating muscles in vivo. Conversely, differentiation of cultured myoblasts, under conditions in which they would normally differentiate, can be inhibited by transfecting the cells with DNA encoding cyclin D1 under the control of a constitutively active promoter. Expression of cyclin D1, which normally occurs only during G_1 , is up-regulated by mitogenic factors in many cell types (see Figure 13-29). The ability of cyclin D1 to prevent myoblast differentiation in vitro may mimic aspects of the in vivo signals that antagonize the differentiation pathway. The antagonism between negative and positive regulators of G_1 progression is likely to play an important role in controlling myogenesis in vivo.

A Network of Cross-Regulatory Interactions Maintains the Myogenic Program

Precursor cells in different regions of the myotome give rise to different muscles: dorsal medial precursors to epaxial

► **FIGURE 14-16 Maintenance of the myogenic program.** Transient signals from the developing spinal cord and ectoderm induce a subset of cells in the developing somite to become myoblasts. Induction is marked by the expression of MRFs and MEFs. These proteins cross-regulate each other's expression and also directly interact to control transcription of other myogenic genes. This network of interactions maintains the myogenic program after the transient inductive signals disappear. Various factors control the final decision to become postmitotic and to differentiate into muscle. See text for further discussion.

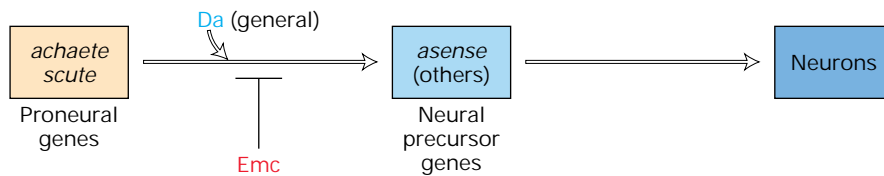


muscles, lateral precursors to hypaxial muscles, and ventrolateral precursors (after migrating) to limb muscles (see Figure 14-9). Each group of precursor cells shows a distinct pathway of myogenic gene activation induced by different signals from surrounding tissues. Once the myogenic program is activated in a region of the somite, an extensive array of cross-regulatory interactions acts to maintain it (Figure 14-16). These cross-regulatory interactions occur at two levels. First, myogenic factors, both MRFs and MEFs, positively regulate each other's expression by binding to cis-acting regulatory sites. Second, MEFs and MRFs physically interact, thereby acting synergistically to promote expression of myogenic factors that drive differentiation. Thus, although the myogenic program is induced by extracellular signals transiently expressed in tissues surrounding the somite, a network of intracellular interactions maintains the myogenic program in the absence of these signals.

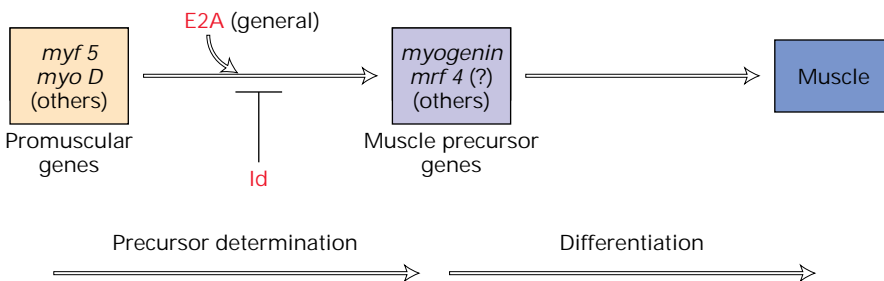
Neurogenesis Requires Regulatory Proteins Analogous to bHLH Myogenic Proteins

Four bHLH proteins that are remarkably similar to the myogenic bHLH proteins control neurogenesis in *Drosophila*. These *Drosophila* proteins are encoded by an ≈100-kb stretch of genomic DNA, termed the *achaete-scute complex* (AS-C), containing four genes designated *achaete* (*ac*), *scute* (*sc*), *lethal of scute* (*l'sc*), and *asense* (*a*). Analysis of the effects of loss-of-function mutations indicate that the Achaete (Ac) and Scute (Sc) proteins participate in determination of neural precursor cells, while the Asense (As) protein is required for neural differentiation. These functions are analogous to the roles of MyoD and Myf5 in muscle determination and of myogenin in differentiation (Figure 14-17). Two other *Drosophila* proteins, designated Da and Emc, are analogous in structure and function to mammalian E2A and Id, respectively. For example, heterodimeric complexes of

Fly neurogenesis



Vertebrate myogenesis



◀ **FIGURE 14-17 Comparison of genes that regulate *Drosophila* neurogenesis and mammalian myogenesis.** bHLH transcription factors have analogous functions in precursor determination and subsequent differentiation into mature muscle cells and neurons. In both cases, the proteins encoded by the earliest-acting genes (left) are under both positive and negative control by other related proteins (red type). [Adapted from Y. N. Jan and L. Y. Jan, 1993, *Cell* 75:827.]

Da with Ac or Sc bind to DNA better than the homodimeric forms of Ac and Sc. Emc, like Id, lacks a DNA-binding basic domain; it binds to Ac and Sc proteins, thus inhibiting their association with Da and binding to DNA.

A family of bHLH proteins related to the *Drosophila* Achaete and Scute proteins have been identified in vertebrates. One of these, called *neurogenin*, which has been isolated from the rat, mouse, and frog, appears to function in determination of neuronal precursor cells. In situ hybridization experiments have shown that neurogenin is expressed at an early stage in the developing nervous system and may induce expression of NeuroD, another bHLH protein that acts later (Figure 14-18a). Injection of large amounts of *neurogenin* mRNA into *Xenopus* embryos further demonstrated the ability of neurogenin to induce neurogenesis (Figure 14-18b). These studies suggest that the function of neurogenin is analogous to that of the Achaete and Scute in *Drosophila*; likewise, NeuroD and Asense may have analogous functions in vertebrates and *Drosophila*.

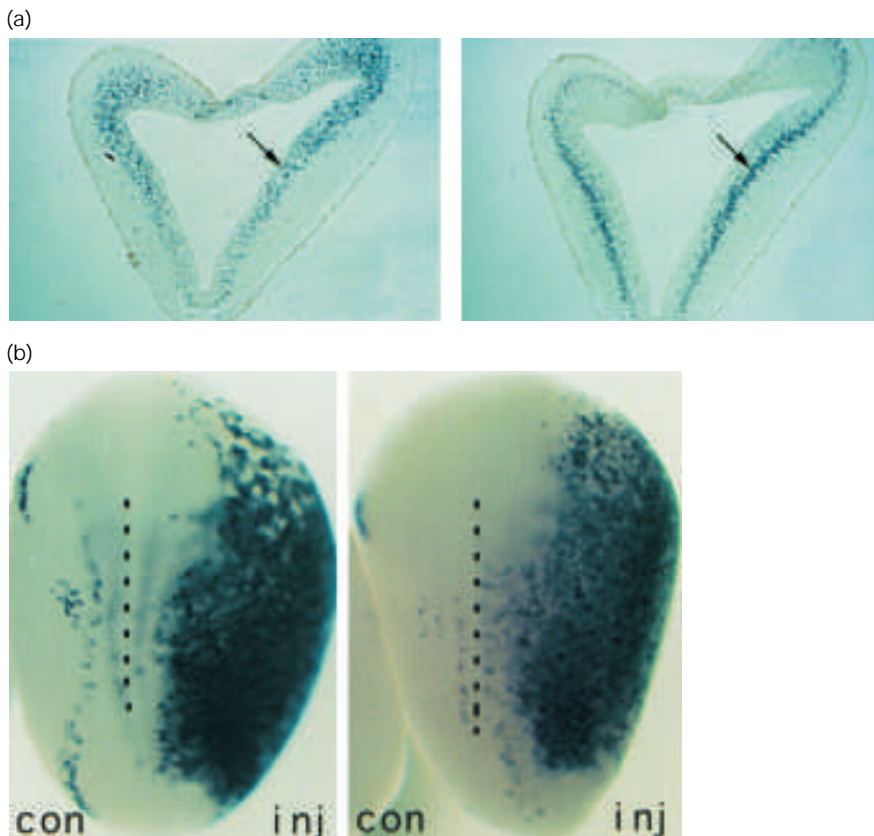
Gene knockout studies in mice, which have two *neurogenin* genes, have confirmed the essential role of neurogenin in vertebrate neurogenesis. In mice embryos that cannot express neurogenin-1, the trigeminal ganglion in the head region does not develop. However, other regions of the nervous system develop normally in *neurogenin-1* knockouts, suggesting that neurogenin-2 or other bHLH proteins regulate neurogenesis in these regions. In the region of the nervous

system affected by the loss of neurogenin-1, development is arrested before expression of NeuroD begins.

Progressive Restriction of Neural Potential Requires Inhibitory HLH Proteins and Local Cell-Cell Interactions

The regulatory mechanisms responsible for restriction of particular developmental pathways to specific cells within an embryo are very complex and not thoroughly understood for any system. For instance, specific cells within somites are selected to become myoblasts, while other cells are destined to become nonmuscle tissues (see Figure 14-9). The best-understood example of such developmental restriction occurs during formation of sensory bristles in *Drosophila*. In this case, the proteins encoded by the proneural genes *achaete* and *scute* must be expressed and active in cells selected to become neural precursor cells, but not in surrounding cells.

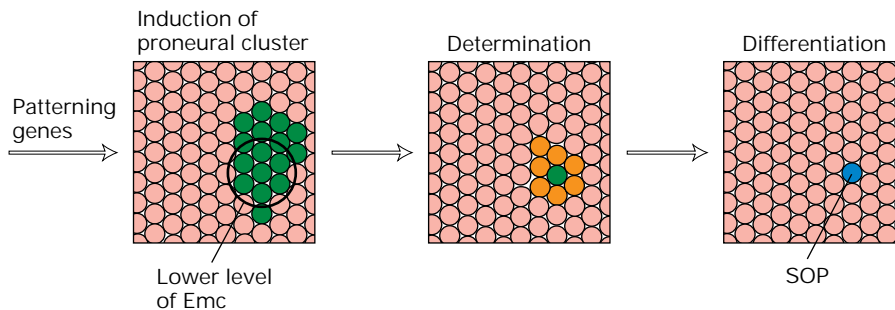
The sensory bristles located on the second thoracic segment of the adult fly arise from a monolayer of cells (a columnar epithelium) called the *wing imaginal disc*, from which the epidermis of this segment and the associated wing also are derived. Each bristle is part of a sensory organ that contains four cells: a bristle that protrudes from the epidermis; the socket into which it is inserted; a neuron that transmits the sensory information; and, finally, a cell associated



◀ **FIGURE 14-18** Experimental demonstration that neurogenin acts before NeuroD in vertebrate neurogenesis. (a) Neurogenin mRNA and neuroD mRNA were detected in the rat neural tube by in situ hybridization. Neurogenin mRNA is expressed in proliferating neuronal precursor cells in the ventricular layer, whereas neuroD mRNA is expressed in migrating neuroblasts that have left the ventricular zone. (b) One of the two cells in early *Xenopus* embryos was injected with neurogenin mRNA (inj) and then stained with a probe specific for neuron-specific β -tubulin mRNA (left) or neuroD mRNA (right). The region of the embryo derived from the uninjected cell served as a control (con). The neurogenin mRNA induced a massive increase in the number of neuroblasts expressing neuroD mRNA and neurons expressing β -tubulin mRNA in the region of the neural tube derived from the injected cell. [From Q. Ma et al., 1996, *Cell* 87:43; courtesy of D. J. Anderson.]

with the neuron referred to as a support cell. These cells are derived from a single cell by two sequential divisions. This “grandmother cell” is referred to as a sensory organ precursor (SOP), or more generally as a neural precursor cell. The pattern of SOPs in the developing imaginal disc presages the pattern of bristles in the adult, which is highly reproducible. These cells do not migrate from another location, but rather arise in distinct positions within the columnar epithelium of the imaginal disc. Each SOP emerges from a cluster of cells, the proneural cluster, that express the Achaete and Scute proteins. The pattern of proneural clusters (i.e., of Ac and Sc expression) is determined by earlier-acting patterning genes whose encoded proteins function to divide the epithelium into developmental domains. During normal development, only one cell in each proneural cluster is selected to become an SOP; the remaining cells develop into epidermal structures. Restriction of the neural program to one cell results from down-regulation of the activity and expression of the proneural proteins Achaete and Scute.

As noted earlier, *Emc* inhibits the binding of Achaete and Scute to DNA and hence their ability to determine neural precursor cells (see Figure 14-17). Loss-of-function mutations in *emc* lead to formation of multiple sensory bristles from a single proneural cluster, whereas gain-of-function mutations suppress SOP formation. In wild-type embryos, expression of *Emc* is lower in the region of each proneural cluster from which a SOP will arise than in the regions giving rise to epidermal structures. Like *achaete* and *scute*, *emc* is under complex regulation so that *Emc* is expressed in a specific pattern within the developing epithelium. The resulting variation in the relative level of proteins promoting neurogenesis (e.g., Achaete and/or Scute) and the level of proteins inhibiting it (e.g., *Emc*) in the cells of the wing imaginal disc limits SOP-forming potential to a small group of neighboring cells within each proneural cluster (Figure 14-19).



▲ **FIGURE 14-19** Formation of sensory organ precursors (SOPs) in the wing imaginal disc of *Drosophila*. A set of extracellular signaling molecules and transcription factors, encoded by so-called patterning genes, control the precise spatiotemporal pattern of expression of proneural bHLH proteins such as Achaete and Scute (green) and related proteins that antagonize their function (e.g., *Emc*). Most cells within the disc express *Emc* (light red), but only small groups of cells, the

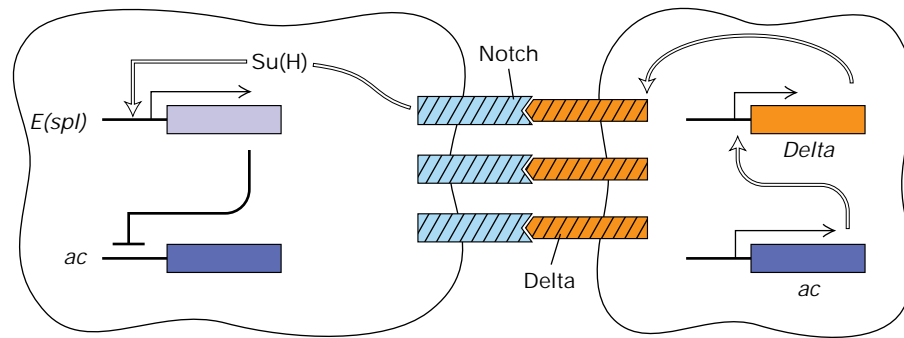
Short-range cell-cell interactions further restrict SOP formation by inhibiting the expression of Achaete and/or Scute in all but one cell of a proneural cluster. These local interactions are mediated by two cell-surface proteins: Notch, a cell-surface receptor, and Delta, its specific ligand (Chapter 23). Loss-of-function mutations in either the *Notch* or *Delta* locus result in the formation of multiple SOPs from a proneural cluster and the appearance of multiple bristles arising from a single proneural cluster, indicating that these genes act to inhibit SOP formation (Figure 14-20). Local asymmetry in the expression of Notch and Delta, reinforced by variations in the levels of Achaete, Scute, and *Emc*, permits one cell, and only one cell, in a proneural cluster to retain its neural potential and become an SOP (Chapter 23; see Figure 23-28). The selected SOP then begins to express *Asense* and other differentiation proteins, which determine the type of neuron that develops, a process termed *neuronal specification*. As production of *Asense* protein increases in an SOP, synthesis of Achaete and Scute proteins decreases.

bHLH Regulatory Circuitry May Operate to Specify Other Cell Types

The important role of bHLH proteins in myogenesis and neurogenesis is supported by discovery of similar highly conserved regulatory proteins in *C. elegans*. Moreover, considerable evidence indicates that a bHLH protein, called SCL, participates in determination of hematopoietic stem cells, which differentiate to generate the many different types of blood cells. SCL is expressed in the ventral mesoderm of the developing embryo, in a region of the mesoderm giving rise to hematopoietic stem cells; like MyoD and Myf5, SCL forms a complex with the more generally expressed E2A protein.

The specification of cell type may be an ancient function of bHLH proteins. In the cnidarian *Hydra vulgaris*, which

proneural clusters, express proneural bHLH proteins. The region of the proneural cluster from which an SOP will form expresses lower levels of *Emc*, giving these cells a bias towards SOP formation. Interactions between these cells, leading to accumulation of E(spl) repressor proteins in neighboring cells (orange), then restrict SOP formation to a single cell (see Figure 14-20). The activity of Achaete in the SOP promotes expression of *Asense* (blue), which is required for further differentiation.



▲ FIGURE 14-20 Cell-cell interactions that down-regulate proneural genes are critical in determination of a single SOP within a proneural cluster. Notch, a cell-surface receptor, and Delta, another cell-surface protein that binds to Notch, are initially expressed in all cells within the developing epithelium. Because expression of Delta is promoted by Achaete and other proneural bHLH proteins, those cells within a proneural cluster with the highest Achaete activity (*right*) will express more Delta and thus will provide the strongest signal to neighboring cells (*left*) through the Notch receptor. Interaction of Delta with Notch triggers an intracellular signaling pathway that activates a transcription factor designated Su(H), which in turn promotes

expression of *E(spl)* genes. These genes encode a family of bHLH proteins that specifically bind to and repress transcription of proneural genes, such as *achaete* (*ac*). The resulting decrease in Achaete leads to a decrease in Delta expression, thus reinforcing the direction of cell-cell signaling. As a consequence of these interactions and others, neural potential is gradually decreased in all but one cell of a proneural cluster. Some workers have proposed that at a critical level of proneural gene expression, an autoregulatory circuit becomes self-staining in one cell, leading to SOP determination. The role of Notch and Delta in developmental programs is discussed further in Chapter 23.

diverged from arthropods some 600 million years ago, a bHLH protein is specifically expressed in the nematocyte, one of some 20 different cell types found in this organism. Nevertheless, it is unlikely that all cell types will be controlled by this regulatory circuitry. Future research of diverse cell types will most likely uncover additional strategies of cell specification using different networks of transcription factors.

SUMMARY Cell-Type Specification in Animals

- In some systems, transient extracellular signals induce a cell-specification program, and an intracellular network of regulatory proteins maintains it.
- Skeletal muscle cells arise from a subclass of cells in the developing somite that are induced to express myogenic bHLH proteins, or MRFs. MyoD and Myf5 are required for commitment of mesodermal cells to myoblasts, and myogenin is required for myoblasts to differentiate into myotubes (see Figure 14-14).
- MEFs and MRFs bind to each other and act synergistically to control transcription of muscle-specific genes.
- Since myoblasts continue to proliferate and, in some cases, migrate to different regions of the developing embryo, specific mechanisms must maintain the determined state and prevent differentiation until an appropriate time. These mechanisms include inhibitory proteins (e.g., Id) that prevent the formation of bHLH dimers and proteins that promote cell-cycle progression.

- Neurogenesis in flies is controlled by a network of bHLH proteins analogous to those controlling skeletal myogenesis (see Figure 14-17). A similar network probably controls vertebrate neurogenesis.
- Neuronal precursor cells (SOPs) in *Drosophila* arise from equipotent groups of cells, called proneural clusters, which all express Achaete and/or Scute. The ability of these cells to give rise to neuronal precursors is progressively restricted (see Figure 14-19).
- Emc inhibits expression of Achaete/Scute in many cells in a proneural cluster, leaving a small number of cells competent to form SOPs. Interactions between these cells mediated by two transmembrane proteins restricts neuronal specification to a single cell (see Figure 14-20).

14.3 Anteroposterior Specification during Embryogenesis

In previous sections of this chapter, we discussed the transcription-control mechanisms that specify different cell types in yeast and animals. Each cell type expresses specific subsets of genes encoding proteins that determine biochemical and morphologic properties characteristic of that cell type. In addition to different cell types, multicellular organisms exhibit striking regional differences in their cellular organization. For instance, the tissue in hands and feet are composed of the same cells organized in very different ways.

What mechanisms determine how cells are organized in different parts of an organism? Or specify that one end of the developing embryo will become a head and the other the tail? What controls the size and position of different organs? These features of an organism are often collectively referred to as the *body plan*.

Although each phylum has a different body plan, molecular studies have revealed a striking relationship between the body plans of many phyla including the millions of different species of arthropods (e.g., insects) and the 50,000 or so chordate species (e.g., mammals). The body plan along the anteroposterior axis of these and other phyla is specified by a set of highly related transcription factors encoded by discrete clusters of *Hox genes*. The first Hox genes were discovered through classical and molecular genetic studies in *Drosophila* and have since been identified in many other organisms in different phyla through their homology to the fly genes. Mutations in the Hox genes often cause homeosis, that is, the formation of a body part having the characteristics normally found in another part at a different site. The proteins encoded by Hox genes are expressed in broad regions along the anteroposterior axis and act to specify the patterning of cells characteristic of tissues within that region in a process sometimes referred to as *regionalization*.

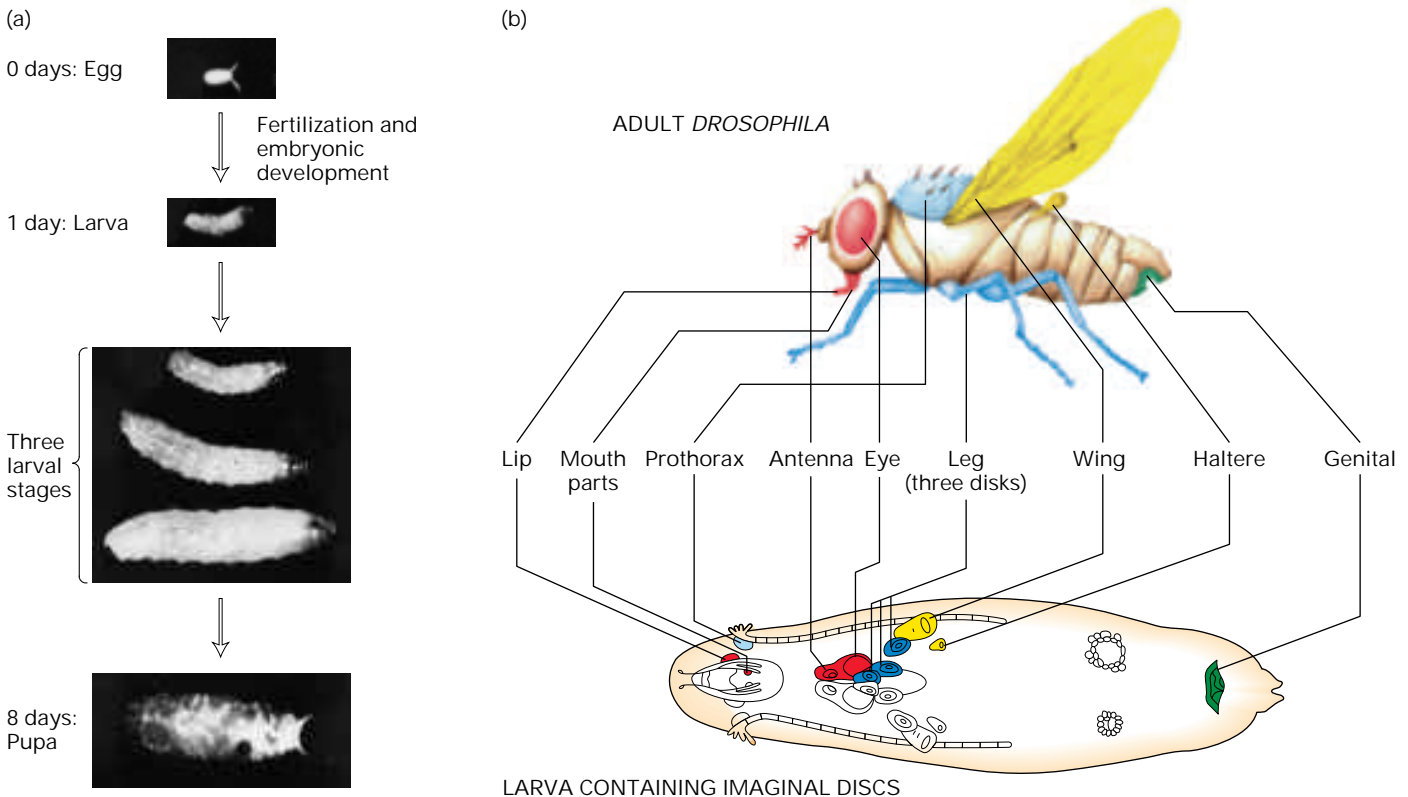
In this section, we examine the mechanisms that determine where Hox genes are expressed in the developing body and how these genes specify regional patterns of development. Because these processes are understood in most detail

in *Drosophila*, our discussion focuses on anteroposterior specification in this organism, beginning with how the fly embryo is progressively divided up into smaller and smaller domains characterized by unique patterns of expression of different combinations of transcription factors, which in turn control expression of the Hox genes. Studies of these processes in *Drosophila* not only have provided important insights into regionalization but also have uncovered widely used transcription-control mechanisms that direct various developmental pathways. After we discuss current understanding of early patterning in the fly embryo and how Hox genes function in flies, we consider the roles of Hox genes in mammalian development.

A set of conserved genes also controls regionalization along the dorsoventral axis. Because studies on dorsoventral specification have focused on the role of intercellular communication, we consider this topic in Chapter 23.

Drosophila Has Two Life Forms

The entire life cycle of *Drosophila* occurs within only 9–10 days (Figure 14-21a). The organism has two forms, a worm-like *larval* form and the adult fly form, separated by a period of metamorphosis called *pupation*. Within 1 day the fertilized egg develops into a larva; three subsequent larval stages, or *instars* as they are called, require about 4 more days. The larva is actually a separate animal from the adult. During early embryogenesis about a dozen groups of cells, termed

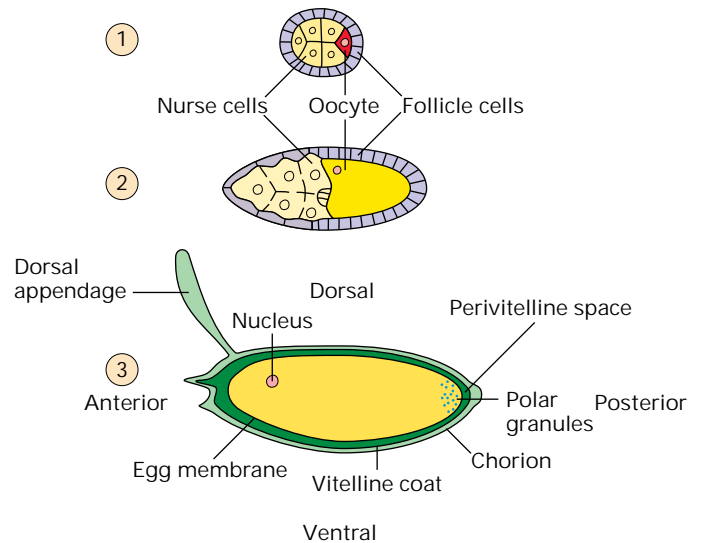


imaginal discs, are set aside and are carried inside the larva (Figure 14-21b). These groups of cells give rise to the adult epidermal structures (wings, legs, etc.). Other groups of precursor cells give rise to adult internal organs such as portions of the gut, the vast majority of the musculature, and the central nervous system. Some larval cells also are conserved in the adult. After the last larval stage, an outer shell is formed. The larval cells are broken down and nutrients derived therefrom are used in the growth and development of the cells that give rise to the different body parts of the adult fly. Pupation takes another 4 days or so. At the end of pupation the shell splits and an adult fly emerges.

Patterning Information Is Generated during Oogenesis and Early Embryogenesis

The blueprint for constructing a fruit fly, including critical spatial information, is in some respects laid down in the egg before fertilization. Production of an egg (*oogenesis*) occurs in *ovarioles*, which collectively form the fly ovary. At the distal end of an ovariole, a stem cell divides asymmetrically generating a single germ cell, which divides four times to generate 16 cells. One of these cells completes meiosis, becoming an *oocyte*; the other 15 cells become *nurse cells*, which synthesize proteins and mRNAs that are transported by a series of cytoplasmic bridges into the oocyte (Figure 14-22). These molecules are necessary for maturation of the oocyte and early stages of embryogenesis. Each group of 16 cells is surrounded by a single layer of *follicle cells*, which form the egg shell. As an oocyte matures within an ovariole, new germ cells are produced from a stem cell, displacing the previously generated oocyte, nurse cells, and surrounding follicle cells. The egg is released into the oviduct, where it is fertilized by sperm from a previous mating, which are stored in the seminal vesicle. The fertilized egg, or zygote, then is laid through the vulva.

Embryogenesis is activated by fertilization. Interestingly, the polarity of the early embryo is presaged in the mature oocyte, which has distinct anterior and posterior ends. In



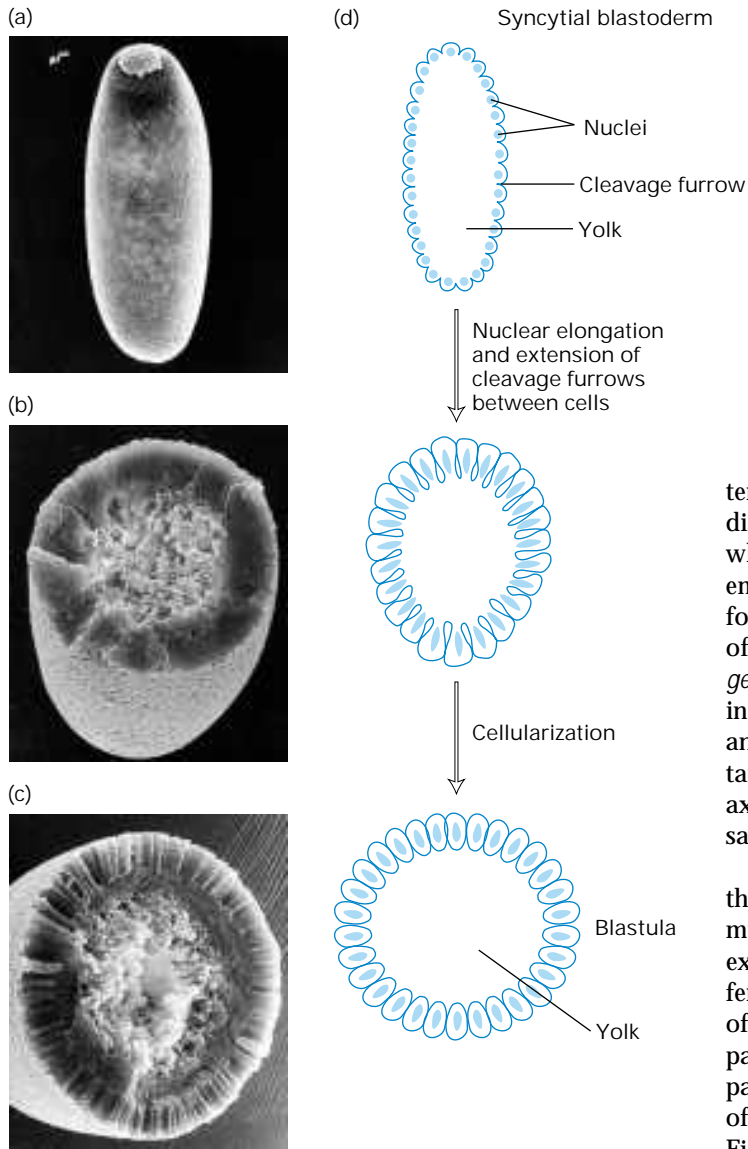
▲ FIGURE 14-22 Structure of the developing *Drosophila* oocyte at three stages in its maturation. Each developing unit, or follicle, consists of a developing oocyte, nurse cells, and a layer of somatic cells called follicle cells. Stage ①: Early in oogenesis, the oocyte is about the same size as the neighboring nurse cells. Stage ②: The nurse cells begin to synthesize mRNAs and proteins necessary for oocyte maturation, and the follicle cells begin to form the egg shell. Midway through oogenesis, the oocyte has increased in size considerably. Stage ③: The mature egg is surrounded by the vitelline coat and chorion, which compose the egg shell. The nurse cells and follicle cells have been discarded, but some of the mRNAs synthesized by nurse cells, which become localized in discrete spatial domains of the oocyte, function in early patterning of the embryo. Polar granules are distinct cytoplasmic structures located in the posterior region of the egg. This is the region in which germ cells arise. [Adapted from A. J. F. Griffiths et al., 1993, *An Introduction to Genetic Analysis*, 5th ed., W. H. Freeman and Company, p. 643.]

◀ FIGURE 14-21 The development of *D. melanogaster*.

(a) The fertilized egg develops into a blastoderm and undergoes cellularization in a few hours. The larva, a segmented form, appears in about 1 day and passes through three stages (instars) over a 4-day period, developing into a prepupa. Pupation takes ≈4–5 days ending with the emergence of the adult fly from the pupal case. (b) Groups of ectodermal cells called imaginal discs are set aside at specific sites in the larval body cavity. From these the various body parts indicated develop during pupation. Other precursor cells give rise to adult muscle, the nervous system, and other internal structures. [Part (a) from M. W. Strickberger, 1985, *Genetics*, 3d ed., Macmillan, p. 38; reprinted with permission of Macmillan Publishing Company. Part (b) adapted from same source and J. W. Fristrom et al., 1969, in E. W. Hanly, ed., *Park City Symposium on Problems in Biology*, University of Utah Press, p. 381.]

addition, some of the mRNAs produced by nurse cells become localized in very discrete spatial domains of the oocyte. The first 13 nuclear divisions of the fertilized ovum are synchronous and rapid, each division occurring about every 10 minutes. Because nuclear division is not accompanied by cell division, a syncytium forms. As the nuclei divide, they begin to migrate outward toward the plasma membrane of the embryo. By 3 hours after fertilization, the nuclei have reached the surface of the embryo, which at this stage is referred to as the *syncytial blastoderm* (Figure 14-23). Cell membranes then form around the nuclei, generating the blastula.

Early-patterning events in *Drosophila* occur before formation of the blastula, that is, before cellularization of the embryo. The mechanisms regulating spatial organization in the fly embryo rely in large part on the diffusion of developmentally important proteins, largely transcription factors, within the developing syncytium. This regulatory strategy is very different from that of other organisms, including



◀ **FIGURE 14-23** Formation of the blastula during early embryogenesis in *Drosophila*. Nuclear division is not accompanied by cell division until about 2000–4000 nuclei have formed. Electron micrographs of embryos before cellularization show surface bulges overlying individual nuclei (a) and absence of cell membranes (b), which are evident after cellularization (c). Note separation of the nuclei of so-called pole cells, which give rise to germ cells, at the posterior end (*top*) of the embryo in (a). Change of syncytial blastoderm into a blastula is illustrated in corresponding diagrams (d), in which pole cells are not shown. [See R. R. Turner and A. P. Mahowald, 1976, *Devel. Biol.* 50:95; photographs courtesy of A. P. Mahowald.]

Mutations in maternal genes necessary for early patterning were identified and classified on the basis of their disruption of the outer (cuticular) structures of the embryo, which occur in a highly reproducible pattern in wild-type embryos (Figure 14-24). These studies led to recognition of four groups of maternal genes, each controlling development of different regions of the embryo as follows: *anterior-group genes*, the head and thorax; *posterior-group genes*, abdominal segments; *terminal-group genes*, the extreme anterior and posterior regions that give rise to the extreme head and tail regions; and *dorsoventral-group genes*, the dorsoventral axis. Some mutations in the latter group lead to loss of dorsal structures, and others, to loss of ventral structures.

In order to decipher the molecular and cellular basis of these patterning mechanisms, investigators had to (1) clone the mutation-defined genes; (2) determine the pattern of mRNA expression and the distribution of the encoded proteins in different spatial domains of the embryo; and (3) assess the effects of mutations on patterning and on the expression of other patterning genes. As just described, the effect of mutations on patterning can be assessed visually by examining the patterns of cuticular structures on the surface of dead embryos (see Figure 14-24). These cuticular structures are formed by the cells directly beneath them. Early gene products can be visualized microscopically by *in situ* hybridization to detect specific mRNAs and staining with antibody to detect specific proteins. Figure 14-25 illustrates the use of these techniques to localize several early gene products formed before infolding of the blastula to form the gastrula. Gene expression in fly embryos also can be detected by use of a reporter-gene construct in transgenic flies. In this method, the *E. coli lacZ* gene, encoding β -galactosidase, is fused to a promoter element that normally controls transcription of a *Drosophila* gene of interest. Expression of β -galactosidase, which is easily assayed, thus serves as a “stand-in” or “reporter” for expression of the fly gene product (see Figure 10-24).

mammals, in which early-patterning events are regulated by interactions between cells mediated by signaling molecules (Chapter 23). Thus, although Hox genes play a central role in specifying regional identity in both insects and mammals, the early-patterning mechanisms that control where Hox genes are expressed along the body axis are different.

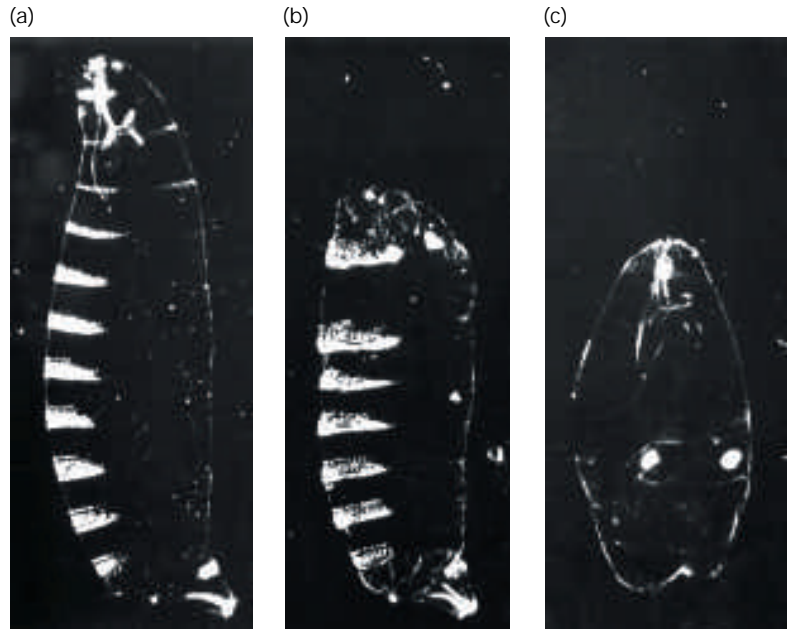
Four Maternal Gene Systems Control Early Patterning in Fly Embryos

Systematic analysis of mutations that affect early development in *Drosophila* has led to the characterization of four different systems of *maternal* genes that regulate axis determination in the embryo. These genes are transcribed in nurse cells during oogenesis, and the corresponding mRNAs are transported into the oocyte. Translation of these mRNAs yields proteins that regulate the later transcription of *zygot* genes within the early embryo.

Morphogens Regulate Development as a Function of Their Concentration

A central concept in developmental biology is that of a morphogen, a substance that specifies cell identity as a function

► **FIGURE 14-24** Abnormal patterns in the outer (cuticular) structures of the *Drosophila* embryo resulting from mutations in two of the four maternal gene systems that regulate axis determination during early embryogenesis. Because the pattern of external cuticular structures is highly reproducible, these structures serve as indicators, or markers, of regional identity along the axes in the fly embryo. Embryos derived from mothers homozygous for mutations in these early-patterning genes exhibit various types of abnormal cuticular patterns; these embryos do not survive. In these preparations, anterior is toward the top and ventral is toward the left. (a) The wild-type pattern. (b) Mutations in the *bicoid* locus (anterior system) disrupt development of the anterior abdominal segments, the thorax, and regions of the head. (c) Mutations in the posterior system lead to loss of abdominal segments. An *oskar* mutant is shown. [From D. St. Johnston and C. Nüsslein-Volhard, 1992, *Cell* 68:201.]



of its concentration. A continuous gradient of morphogen concentration can elicit a set of unique cellular responses at a finite number of *threshold* concentrations: above the threshold, one response is elicited; below it, cells respond differently. As an example, suppose that at the site in the embryo where a particular morphogen is synthesized, its concentration is high enough to establish fate A for cells in the immediate vicinity. As the distance from this site increases, the concentration of the morphogen decreases. At some distance (e.g., 10 cell diameters), a threshold concentration is reached; cells that experience morphogen levels below the threshold are consigned to cell fate B. Still farther away, the morphogen concentration may reach another threshold below which cell fate C is established, and so on.

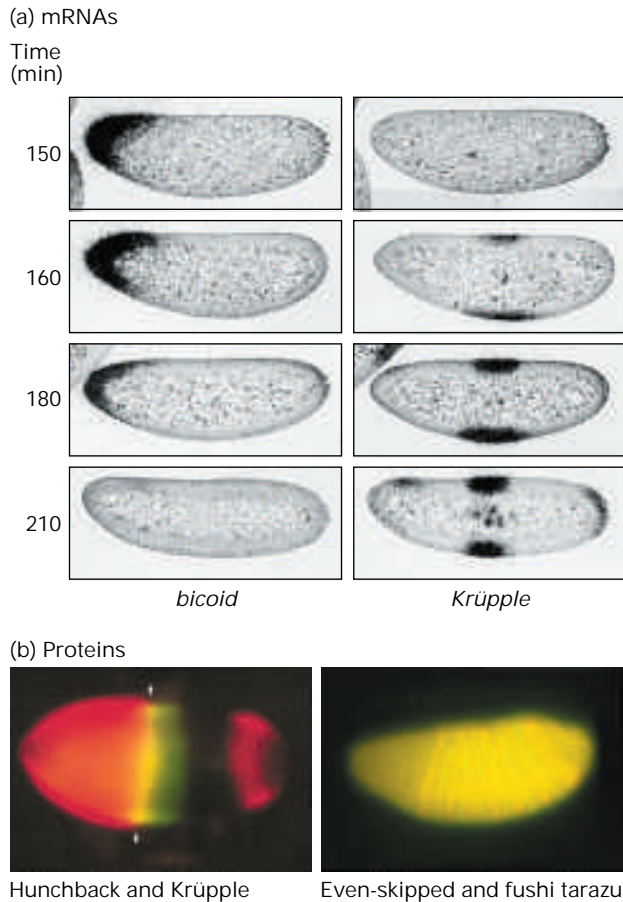
Several proteins in the developing *Drosophila* embryo have been shown to function as morphogens. Molecular analysis of mutants, for instance, has revealed that gradients of transcription factors produced from maternal mRNAs deposited in the developing oocyte control the transcription of zygotic genes in the early embryo in spatially restricted domains. The products of these zygotic genes, in turn, control the transcription of other zygotic genes, further refining specific spatial domains of gene activity in the embryo. Morphogens controlling development along the anteroposterior axis diffuse within the shared cytoplasm of the syncytial blastoderm. In contrast, patterning along the dorsoventral axis, as well as at the extreme anterior and posterior termini, occurs in response to graded extracellular signals located in the perivitelline space between the plasma membrane of the blastoderm and the vitelline membrane surrounding it. In embryos in which the cells are separated by membranes, early patterning along the anteroposterior axis also depends on

extracellular morphogens, which are considered in Chapter 23. Our discussion here focuses on how transcription activators and repressors function as morphogens within the *Drosophila* syncytial blastoderm.

Maternal *bicoid* Gene Specifies Anterior Region in *Drosophila*

The first morphogen to be described at the molecular level was the protein encoded by the *bicoid* locus in *Drosophila*. The *bicoid* mRNA, which is synthesized in nurse cells during oogenesis and transported to the maturing oocyte, is localized to the most anterior region, or anterior pole, of the early fly embryo (see Figure 14-25a). Embryos produced by female flies that are homozygous for *bicoid* mutations lack head and thoracic tissue (see Figure 14-24b). Other maternal mRNAs either are unlocalized or are localized to different regions of the embryo (e.g., *gurken* mRNA at the antero-dorsal edge and *nanos* mRNA posteriorly). The anterior localization of *bicoid* mRNA depends on its 3'-untranslated end and the products of three other maternal genes. Mutations that result in failure to localize *bicoid* mRNA produce a phenotype similar to, though less severe than, the phenotype associated with mutations in the *bicoid* gene itself.

The *bicoid* gene encodes a transcription factor whose DNA-binding region is a homeodomain (see Figure 10-40). In the early fly embryo before cellularization, Bicoid diffuses away from the anterior end where it is produced through the common cytoplasm, forming a protein gradient along the anteroposterior axis. Evidence that the Bicoid protein gradient determines anterior structures was obtained through injection of synthetic *bicoid* mRNA at different locations in



◀ **FIGURE 14-25** Localization of developmentally important gene products in early *Drosophila* embryos. All embryos are positioned with anterior to the left and dorsal at the top. (a) In situ hybridization with labeled RNA probes of whole embryo sections 2.5–3.5 hours after fertilization, which covers the period from the syncytial blastoderm to the beginning of gastrulation. The dark silver grains show the positions of the mRNAs encoded by *bicoid*, a maternal gene and *Krüpple*, a gap gene. These and other early-gene products are expressed in characteristic reproducible temporal and spatial patterns. (b) Fixed embryos stained with antibodies that are coupled to different fluorescent dyes. (Left) Hunchback protein (red) and Krüpple protein (green) in a syncytial blastoderm. Both are gap-gene products. The yellow band is a region of overlap of the two proteins. (Right) Even-skipped protein (yellow) and Fushi tarazu protein (orange) in alternating bands at beginning of gastrulation. Both are pair-rule gene products. [Part (a) from P. W. Ingham, 1988, *Nature* 335:25; photographs courtesy of P. W. Ingham. Part (b) courtesy of M. Levine.]

the embryo. This treatment led to formation of anterior structures at the site of injection with progressively more posterior structures forming at increasing distances from the injection site.

The concentration gradient of Bicoid protein, which promotes transcription of the zygotic *hunchback* gene, determines the region in which the Hunchback protein is expressed. Mutations in *hunchback* and several other zygotic genes lead to large gaps in the anteroposterior pattern of the early embryo; hence these genes are collectively called *gap genes*. Several types of evidence indicate that Bicoid protein directly regulates transcription of *hunchback*. For example, the spatial distribution of expressed Hunchback protein parallels that of the Bicoid protein gradient (Figure 14-26a–c). Moreover, analysis of the *hunchback* promoter just upstream of the transcription-start site has shown that it contains three low-affinity and three high-affinity binding sites for Bicoid protein. Studies with transgenic flies carrying reporter genes driven by synthetic promoters containing either all high-affinity or low-affinity Bicoid-binding sites have demonstrated that the affinity of the site determines the threshold concentration of Bicoid at which gene transcription is activated. That is, in response to the same Bicoid protein gradient in the embryo, expression of a reporter gene controlled by a promoter carrying high-affinity Bicoid-binding sites extends

more posteriorly than does transcription of a reporter gene carrying low-affinity sites (Figure 14-26d,e). In addition, the number of Bicoid-binding sites occupied at a given concentration has been shown to determine the amplitude, or level, of the response.

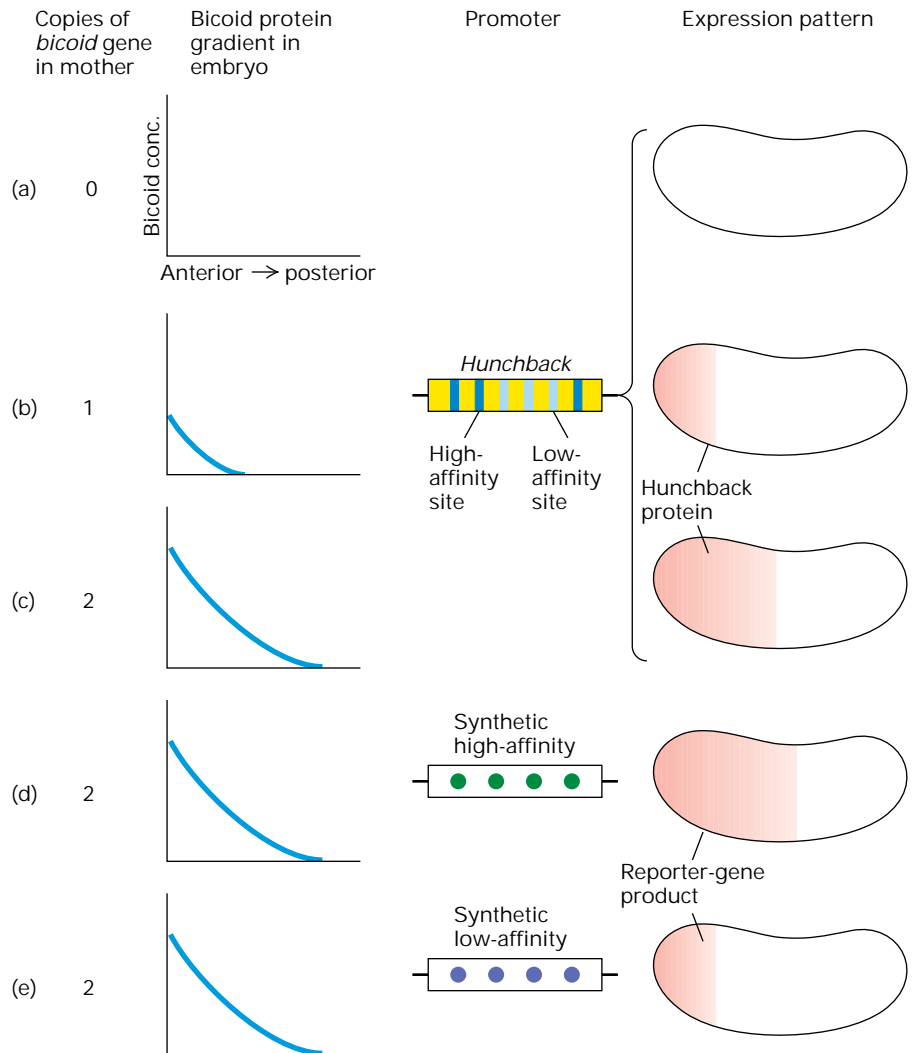
These studies on the ability of Bicoid to regulate transcription of the *hunchback* gene show that variations in the levels of transcription factors, as well as in the number and/or affinity of specific regulatory sequences controlling different target genes, contribute to generating diverse patterns of gene expression during development. Similar mechanisms are found in other developing organisms.

Maternally Derived Inhibitors of Translation Contribute to Early *Drosophila* Patterning

There are two sources of *hunchback* (*hb*) mRNA in the early fly embryo: that derived from zygotic transcription of *hunchback* under the control of Bicoid protein, which is localized anteriorly, and that derived from the mother, which is uniformly distributed. However, even though *hunchback* mRNA is present throughout the embryo, Hunchback protein is not observed in the posterior region. This exclusion of Hunchback protein from the posterior region depends on the posterior-group maternal gene called *nanos*. Maternal *nanos* mRNA, which is localized to the posterior pole, encodes a morphogen that functions in repressing translation of maternal *hunchback* mRNA in the posterior region (Figure 14-27). Other posterior-group maternal genes are necessary for the specific synthesis and localization of *nanos* mRNA to the posterior pole of the embryo. Mutations in any of these genes (e.g., *oskar*) lead to severe defects in development of the embryo (see Figure 14-24c).

Nanos protein acts in conjunction with Pumilio protein, also encoded by a posterior-group maternal gene, to repress

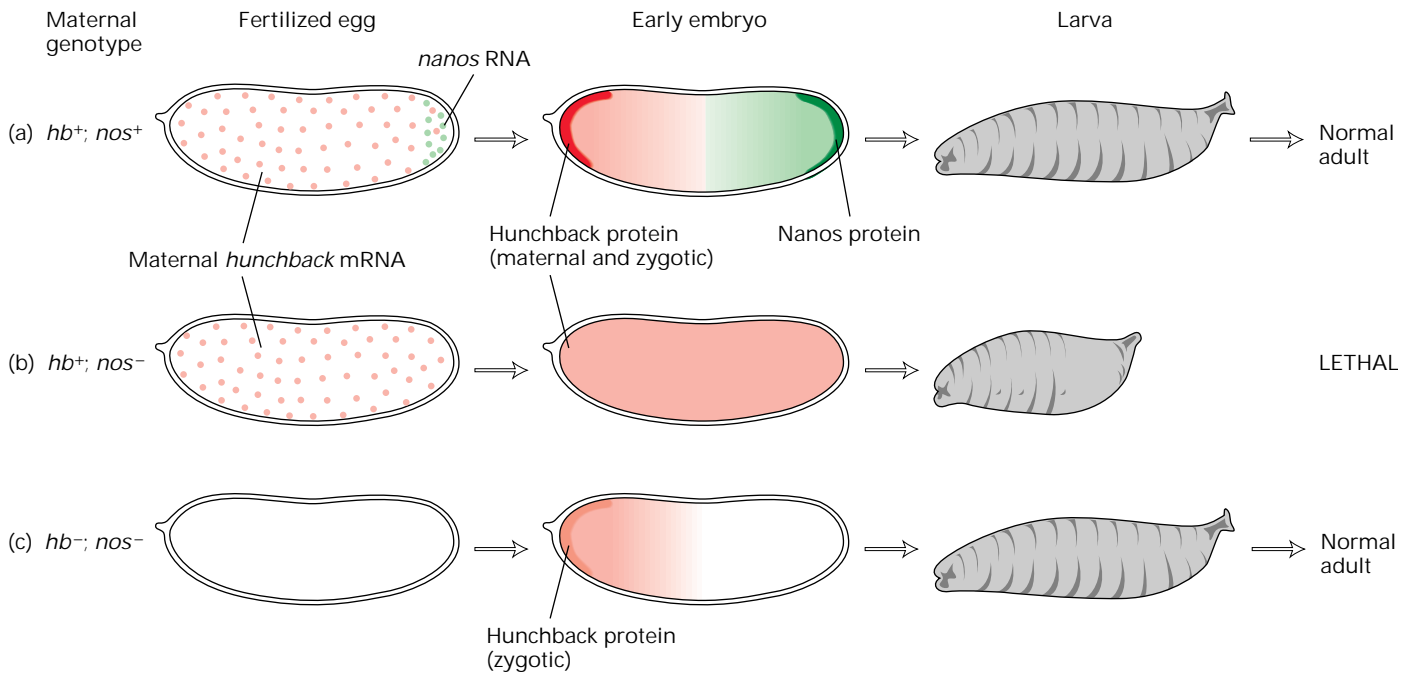
► **FIGURE 14-26** Experimental demonstration that transcription of target genes, such as *hunchback*, directly regulated by Bicoid protein reflects the concentration of Bicoid and its affinity for protein-binding sites in the promoter of the target gene. (a–c) Increasing the number of *bicoid* genes in mother flies changed the Bicoid gradient in the early embryo, leading to a corresponding change in the gradient of Hunchback protein (red) expressed from the zygotic *hunchback* gene. (d,e) The *hunchback* promoter has been shown to contain three high-affinity and three low-affinity Bicoid-binding sites. Transgenic flies carrying a reporter gene linked to a synthetic promoter containing either four high-affinity sites (d) or four low-affinity sites (e) were prepared. Expression of the reporter-gene product was dependent on the affinity of the Bicoid-binding sites in the promoter. [Adapted from D. St. Johnston and C. Nüsslein-Volhard, 1992, *Cell* 68:201.]



translation of maternal *hunchback* mRNA. Repression also depends on specific sequences in the 3'-untranslated region of *hunchback* mRNA, called Nanos-response elements (NREs). Although the precise mechanism by which repression is achieved is not known, it appears to be correlated with the length of the poly(A) tail in *hunchback* mRNA. In wild-type embryos, the length of the poly(A) tail increases immediately prior to translation of *hunchback* mRNA. The length of the poly(A) tail reflects the activities of antagonistic processes of polyadenylation and deadenylation. Recent genetic and molecular studies suggest that Nanos promotes deadenylation of *hunchback* mRNA and thereby decreases its translation. Figure 14-28 illustrates how this translational regulation helps to establish the Hunchback gradient needed for normal development.

Recent findings raise the prospect that translational control may be a widely used strategy for regulating development. For instance, similar mechanisms have been shown to occur during development of *C. elegans*, and a Nanos-

related protein has been identified in *Xenopus* embryos. Even more intriguing is the discovery that Bicoid, which promotes zygotic transcription of the *hunchback* gene, also functions to regulate translation of another *Drosophila* early-patterning gene called *caudal*. This gene plays a crucial role in patterning the posterior region of the embryo; like *bicoid*, it encodes a homeodomain-containing transcription factor. Maternal *caudal* mRNA, like maternal *hunchback* mRNA, is uniformly distributed in the early embryo. Biochemical and genetic studies have shown that Bicoid binds through its homeodomain to a specific sequence in the 3'-untranslated region of *caudal* mRNA. This binding specifically inhibits translation and hence expression of Caudal protein. Since the Bicoid concentration gradient is highest at the anterior end, its repression of *caudal* mRNA translation generates a Caudal concentration gradient that is highest at the posterior end. Future research may well identify other homeodomain proteins that influence gene expression by controlling both transcription and translation.



▲ **FIGURE 14-27** Role of maternally derived Nanos protein in excluding Hunchback protein from the posterior region of *Drosophila* embryos. mRNAs are indicated by colored dots and relative protein concentrations by color shading. (a) In embryos produced by wild-type female flies, maternal *nanos* mRNA is localized posteriorly. Once this mRNA is translated, the resulting Nanos protein inhibits translation of maternal *hunchback* mRNA in the posterior region of the embryo. (b) In the absence of Nanos, translation of maternal *hunchback* mRNA in the posterior region leads to a failure of the posterior structures to form normally and the embryo dies. (c) Because zygotic transcription of *hunchback* in the anterior region is required for normal development, viable

flies homozygous for mutations in both *hunchback* (hb^-) and *nanos* (nos^-) cannot be produced. However, pole cells (germ-line precursors) that are homozygous for both mutations can be transplanted from an early embryo into a surrogate embryo. Female germ-line cells that carry mutations in both the *hunchback* and *nanos* genes produce embryos that develop normally; zygotic expression of Hunchback in the anterior is controlled by Bicoid. This finding shows that Nanos functions solely to prevent translation of maternal *hunchback* mRNA in the posterior region. [Adapted from P. Lawrence, 1992, *The Making of a Fly: The Genetics of Animal Design*, Blackwell Scientific Publications.]

Graded Expression of Several Gap Genes Further Subdivides the Fly Embryo into Unique Spatial Domains

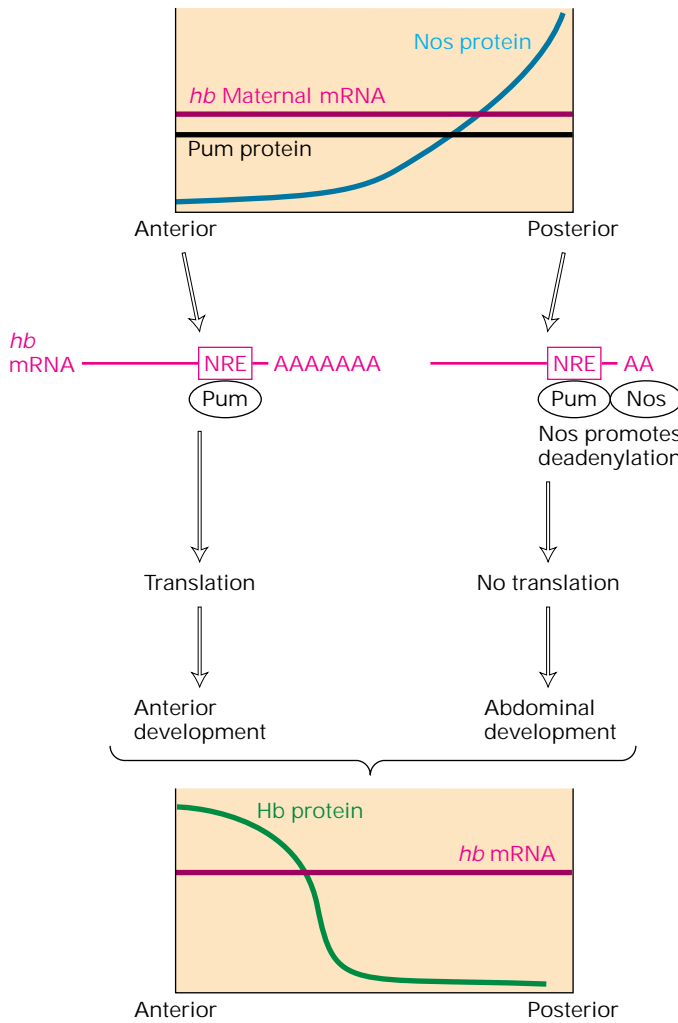
As noted previously, gap genes participate in early patterning along the anteroposterior axis of *Drosophila* embryos. All of these zygotic genes—including *hunchback*, *Krüppel*, *knirps*, and *giant*—are expressed in specific spatial domains within 2 hours following fertilization and just before cellularization of the embryo (Figure 14-29; see also Figure 14-25). Maternally derived Bicoid plays a key role in activating expression of gap genes (e.g., *hunchback*) in the anterior region, whereas both Bicoid and Caudal control gap-gene expression in more posterior regions. Although a loss-of-function mutation in either *bicoid* or *caudal* has little effect on posterior segmentation, posterior segmentation does not occur in embryos with mutations in both genes.

Genetic and molecular experiments suggest that the boundaries of expression of *Krüppel*, *Knirps*, and *Giant* reflect a balance of transcriptional activation and repression involving maternally derived Bicoid and Caudal as well as Hunchback, the first gap-gene protein to be expressed. For instance, high concentrations of Hunchback protein repress

transcription of *Krüppel*, but below a critical threshold concentration, Hunchback acts as a transcriptional activator of *Krüppel*. This threshold establishes the anterior boundary of the *Krüppel* protein domain. More posteriorly, the Hunchback concentration falls below the threshold at which it activates transcription of *Krüppel*, thereby setting the posterior *Krüppel* boundary (see Figure 14-29). The *Knirps* and *Giant* proteins are each located in two domains. Expression in their anterior domains is activated by Bicoid, while expression in their posterior domains is activated by the combined action of Bicoid and Caudal. The anterior boundaries of the posterior *Knirps* and *Giant* domains are determined by Hunchback-mediated transcriptional repression, and the posterior boundaries are determined by the product of another gap gene, *tailless*.

Expression of Three Groups of Zygotic Genes Completes Early Patterning in *Drosophila*

Through the combined action of the anterior- and posterior-group maternal genes, as well as the zygotic gap genes, the early *Drosophila* embryo becomes divided into broad



▲ FIGURE 14-28 Contribution of translational repression to formation of the anterior → posterior Hunchback gradient in the early *Drosophila* embryo. Maternal *hunchback* (*hb*) and *pumilio* (*pum*) mRNAs are uniformly distributed in the early fly embryo, whereas *nanos* (*nos*) mRNA is localized to the posterior. As *nos* mRNA is translated, Nanos (Nos) protein diffuses through the embryo, establishing a posterior → anterior concentration gradient. Simultaneous binding of Nos and Pum proteins to the NRE sequence of *hb* mRNA promotes deadenylation, which inhibits its translation. As a consequence, maternally derived Hunchback (Hb) protein is expressed in a graded fashion that parallels and reinforces the Hb protein gradient resulting from zygotic transcription of *hb* controlled by Bicoid (see Figure 14-26). In mutants lacking either Nos or Pum, this translational repression is decreased; as a result Hb accumulates posteriorly. [See C. Wreden et al., 1997, *Development* 124:3015.]

expression domains characterized by different combinations of transcription factors. These factors are expressed at different levels and act in various combinations to further subdivide the embryo into specific developmental domains. They do so by activating transcription of three additional groups of zygotic genes, termed *pair-rule*, *segment-polarity*,

and *selector* genes. The pair-rule and segment-polarity genes, like the maternal and gap genes discussed previously, are expressed transiently and act to establish spatial domains where selector genes are expressed. Selector genes, which are expressed continuously from the embryo into the adult, are required to specify and maintain regional identity along the anteroposterior axis throughout the remainder of the developmental process.

Before describing the final stages in patterning along the anteroposterior axis of *Drosophila*, we need to define the terms *segment* and *parasegment* (Figure 14-30). The segments of the adult fly originally were named based on visual examination of the adult fly; thus each segment corresponds to a visually distinct unit, not a developmental unit. In contrast, parasegments correspond to the actual spatial domain along the anteroposterior axis over which a specific set of selector genes exert their patterning function. Imagine trying to divide any repetitive structure aligned head to tail; regardless of the specific borders chosen, the pattern will remain repetitive (with the exception of the two ends). During early development, the embryo is divided into 14 parasegments with the anterior border of each parasegment demarcated by a sharp band of cells expressing a particular

Protein visualized

Hunchback

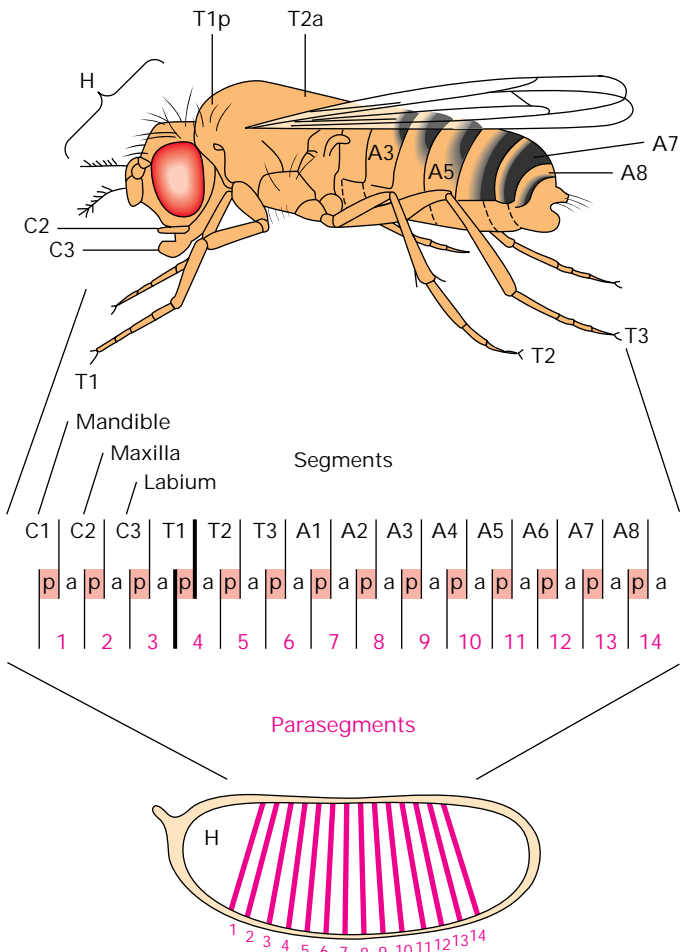
Krüppel

Knirps

Giant



▲ FIGURE 14-29 Localization of gap-gene products in early *Drosophila* embryos visualized by staining with specific antibodies against Hunchback, Krüppel, Knirps, and Giant proteins. Anterior is to the left. Hunchback protein functions as both a transcriptional activator and repressor. Transcription of the *Krüppel*, *knirps*, and *giant* genes is regulated by Hunchback, Bicoid, and Caudal. See text for discussion. [Adapted from G. Struhl et al., 1992, *Cell* 69:237.]



◀ **FIGURE 14-30** The relationship between segments in the adult fly and parasegments, which are developmental units corresponding to the domains of activity of selector genes. The head segments are designated C1–C3; thoracic segments, T1–T3; and abdominal segments, A1–A8. The anterior border of each parasegment is marked by a sharp band of cells expressing Engrailed protein (red). Each segment is divided into an anterior (a) and posterior (p) compartment. Engrailed expression marks cells in the posterior compartment. There is no mixing of cells across the anterior/posterior boundary of each segment. Each parasegment roughly corresponds to the posterior compartment of one segment and the anterior compartment of the adjacent segment to the rear. [From P. A. Lawrence, 1992, *The Making of a Fly: The Genetics of Animal Design*, Blackwell Scientific Publications.]

segment-polarity gene called *engrailed*. Each parasegment roughly contains the posterior portion of one segment and the anterior portion of the segment located just posterior to it. A segment, then, is roughly divided in half by expression of the Engrailed protein. This boundary defines a developmental unit within each segment called a *compartment*: cells posterior to this boundary will become part of the so-called posterior compartment and those anterior to it part of the anterior compartment. Once the border between these compartments is established, there is no mixing of cells across it.

Pair-Rule Genes The pair-rule genes, which include *fushi tarazu*, *hairy*, and *even-skipped*, encode transcription factors that are expressed in stripes of cells, corresponding to the parasegments, covering the central part of the embryo (see Figure 14-25b). Each pair-rule gene product is expressed in seven parasegments, either the even or odd ones. Gene expression in each stripe appears to be controlled independently by the action of different transcription factors encoded by gap and maternal genes. The results of various studies suggest that expression of pair-rule genes in discrete

stripes results from transcriptional activation in a broad domain followed by delineation of sharp borders by spatially restricted repressors.

To see how this strategy works, we consider expression of *Even-skipped* (*Eve*) stripe 2, which is controlled by the maternally derived Bicoid protein and the gap proteins Hunchback, Krüppel, and Giant. These proteins exert their effect by binding to a clustered set of regulatory sites located upstream of the *eve* promoter (Figure 14-31a). Hunchback and Bicoid activate transcription of *eve* in a broad spatial domain, whereas Krüppel and Giant repress *eve* transcription at sharp posterior and anterior boundaries. The segment of DNA containing these regulatory sites will drive expression of a *lacZ* reporter gene specifically in stripe 2. The coordinated effect of these proteins, each of which has a unique concentration gradient along the anteroposterior axis, precisely regulates the boundaries of stripe 2 expression (Figure 14-31b).

Segment-Polarity Genes Next, each parasegment is further subdivided through the action of segment-polarity genes such as *wingless* and *engrailed*. By this stage in embryogenesis, cellularization is complete, and all cells are surrounded by a plasma membrane. Further patterning depends on intercellular signaling, and several segment-polarity genes encode secreted proteins that signal developmental events in neighboring cells. Segment-polarity genes are responsible for generating the patterns of cells within each segment. Patterning of the embryonic cuticle is defined by cuticular protrusions (denticles) and regions devoid of them, which are smooth. Each cuticular pattern reflects a different combination of

pair-rule and segment-polarity genes. These genes also are frequently responsible for determining the polarity of the pattern within a parasegment; the orientation of the denticles also can be altered by the segment-polarity genes. The *engrailed* gene is expressed in the most anterior band of cells of each of the well-defined parasegments (see Figure 14-30).

Selector Genes The next genes to be transcribed in the regulatory hierarchy that controls regionalization of the *Drosophila* embryo are the selector genes. These genes correspond to the Hox genes mentioned earlier, and their encoded proteins regulate development within parasegmental domains. By this stage in embryogenesis, each band of cells along the anteroposterior axis expresses a unique combination of transcription factors, which control subsequent cell development (Figure 14-32).

The patterns of selector-gene expression are determined in the early embryo and maintained into the adult fly. Continuous expression of selector genes is required to determine the structures of the various body parts along the anteroposterior axis. Two of the best-studied selector genes in *Drosophila* are *Antennapedia* (*Antp*), which dominates development of the fourth parasegment, and *Ultrabithorax* (*Ubx*), which largely controls the sixth parasegment. Transcription of the *Antp* and *Ubx* genes begins during the third hour of embryogenesis in domains determined in part by various gap-gene products.

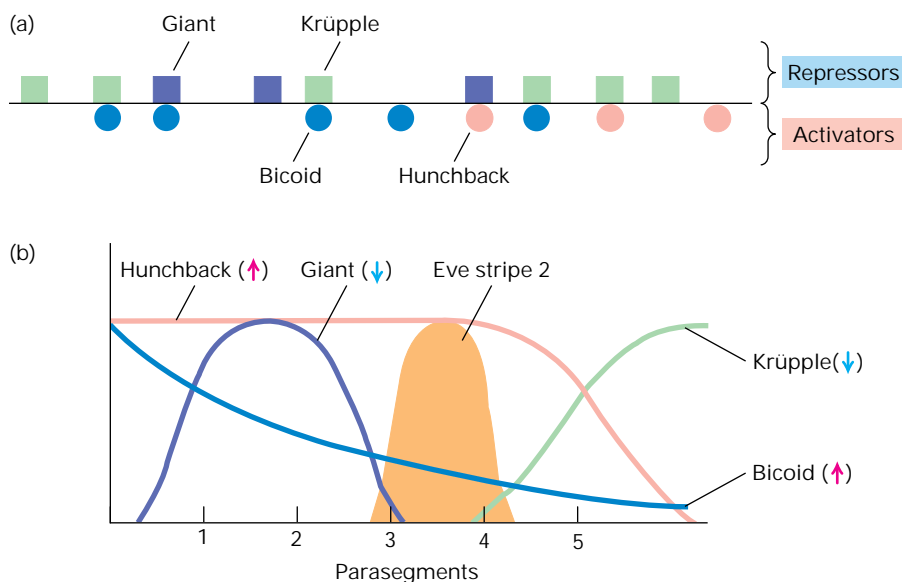
Mutations in selector genes often (but not always) cause transformation of one body part into another (homeosis), providing particularly vivid evidence of the role of genes in regional specification. For instance, misexpression of *Antp* in the primordium of the antenna results in its development into a leg rather than an antenna (Figure 14-33). Both struc-

tures are appendages covered with sensory structures although the type and distribution of these structures are quite different. Conversely, loss of *Ubx* leads to development of a pair of wings in place of balancing organs, called halteres, on the third thoracic segment (see Figure 8-8b). Genes in which mutations have such effects are referred to as homeotic genes.

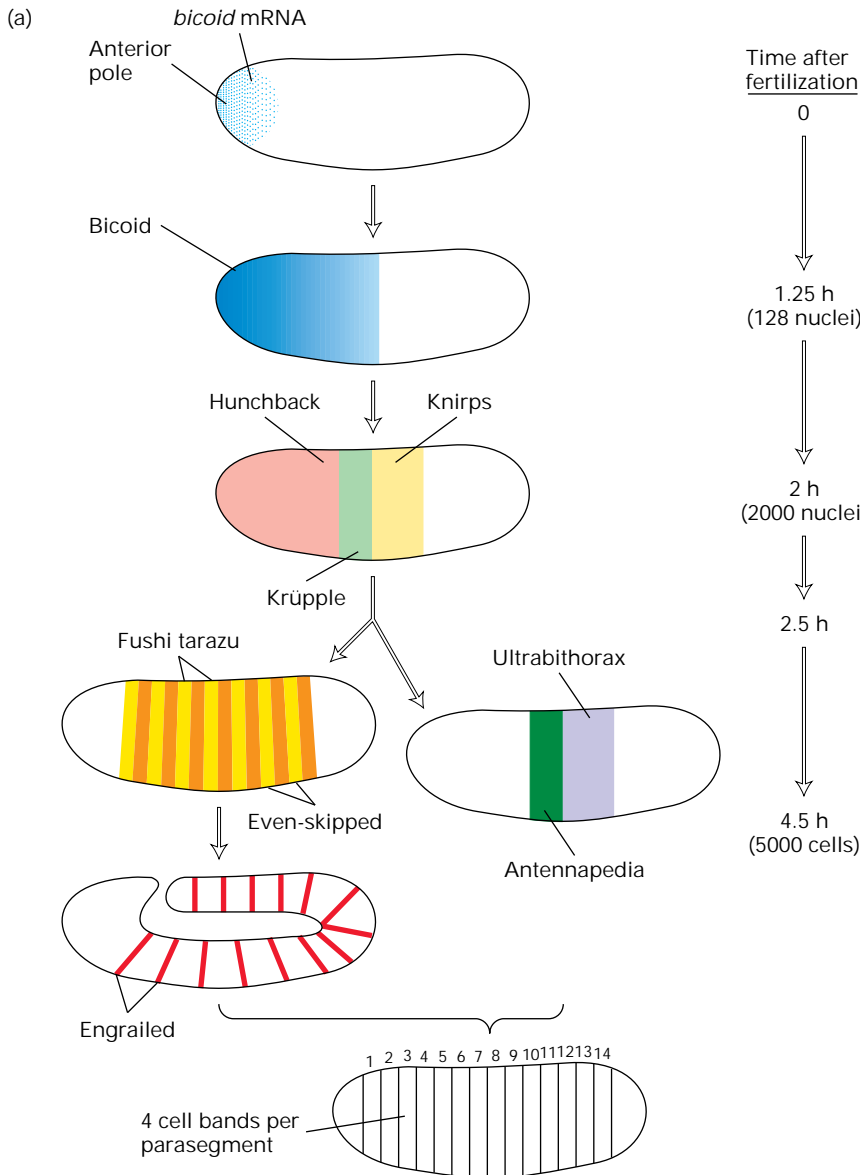
Selector genes act in specification of epidermal structures (the external surface), the musculature, neural tissue, and gut tissue along the anteroposterior axis. Because the role of selector genes in regulating development of the epidermis is understood best, we discuss these genes in some detail. These genes were first identified in *Drosophila*, but similar genes have analogous functions in mammalian development.

Selector (Hox) Genes Occur in Clusters in the Genome

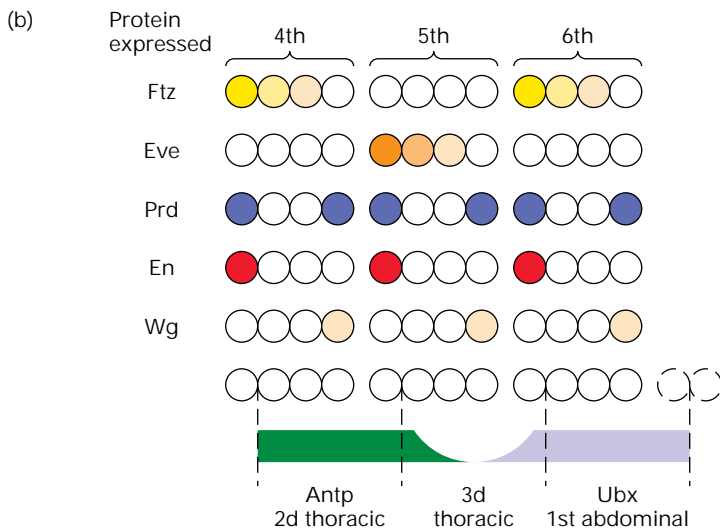
Two clusters of selector genes — the bithorax complex (BX-C) and the antennapedia complex (ANT-C) — play a central role in controlling regionalization of external structures along the anteroposterior axis in *Drosophila*. The BX-C contains three structural genes, called *Ultrabithorax* (*Ubx*), *abdominal A* (*abdA*), and *Abdominal B* (*AbdB*), each of which encodes a homeodomain-containing transcription factor. All three genes have extensive noncoding sequences (e.g., introns) that are critical in regulating their individual expression patterns within different parasegments (Figure 14-34). The linear order of the genes in the BX-C parallels their pattern of expression along the body axis. The same phenomenon is true of the ANT-C, which contains five genes called *labial* (*lab*), *proboscipedia* (*pb*), *Deformed* (*Dfd*), *Sex combs reduced* (*Scr*), and *Antennapedia* (*Antp*). Both gene clusters are located on chromosome III, separated by some 10 cM. The proteins



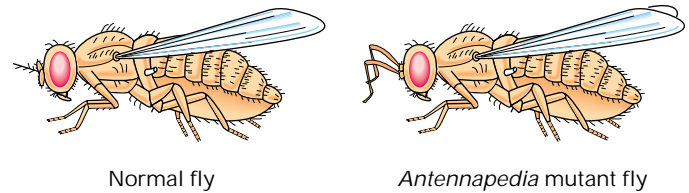
◀ **FIGURE 14-31** Expression of the Even-skipped (*Eve*) stripe 2 in the *Drosophila* embryo. (a) Diagram of the 815-bp regulatory region controlling transcription of the pair-rule gene *eve*. This region contains binding sites for Bicoid and Hunchback proteins, which activate transcription of *eve*, and for Giant and Krüpple proteins, which repress transcription. (b) Concentration gradients of *Eve* stripe 2 and of the four proteins that regulate its expression. The coordinated effect of the two repressors (↓) and two activators (↑) determine the precise boundaries of the second anterior *Eve* stripe. Expression of other stripes is regulated independently by other combinations of transcription factors encoded by maternal and gap genes. [See S. Small et al., 1991, *Genes & Devel.* 8:827.]



◀ **FIGURE 14-32** Summary of sequential expression of various genes during early development of the *Drosophila* embryo and localization of their gene products within the embryo. (a) Maternal *bicoid* mRNA is localized at the anterior pole of the egg, but Bicoid protein, which is synthesized soon after fertilization, diffuses to form a gradient. In most cases, an mRNA and its corresponding protein are present in the same regions of the embryo. The expression domains of three gap-gene products—Hunchback, Krüpple, and Knirps—are shown. Not shown is Nanos protein, which represses translation of *hunchback* mRNA in the posterior region (see Figure 14-27a). Expression of the proteins encoded by the pair-rule genes *fushi tarazu* (*ftz*), *even-skipped* (*eve*), and *paired* (*prd*) is determined by specific combinations of Bicoid and various gap-gene products (see Figure 14-31). These proteins demarcate 14 stripes corresponding to the parasegments. The segment-polarity gene *engrailed* (*en*) is expressed at the anterior end of each parasegment; it plays an important role along with other segment-polarity genes such as *wingless* (*wg*) in patterning of each parasegment. Cellularization occurs after 2.5 hours, and gastrulation (including folding of the embryo) occurs at about 4.5 hours. By this time, each parasegment consists of four bands of cells. (b) Within a parasegment, each band of cells (represented by a circle) is characterized by expression of a unique set of proteins encoded by pair-rule and segment-polarity genes. Shown here are several expression patterns in three parasegments (4–6). These expression patterns act as positional values distinguishing each cell band in a parasegment. The gap-gene products largely determine the expression domains of selector genes such as *Antennapedia* (*Antp*) and *Ultrabithorax* (*Ubx*). The proteins encoded by selector genes specify the organization of larval and adult structures within the context of the positional identity of cells within each parasegment.



► **FIGURE 14-33** Misexpression of *Antp* protein in the developing antenna leads to its transformation into a leg, a structure whose development normally is controlled by *Antp* in thoracic segments. This is an example of a homeotic transformation. Such misexpression can result from a regulatory mutation or induced expression of an *Antp* transgene. See also Figure 8-8b. [From W. McGinnis and M. Kuziora, 1994, *Sci. Am.* 270 (2):58.]



encoded by the ANT-C and BX-C control development of parasegments 0 – 5 and parasegments 5 – 14, respectively.

The organization of genes within both ANT-C and BX-C is not likely to be serendipitous, as it is observed in homologous gene clusters in organisms in different phyla. As we discuss later, mammalian homologs of the ANT-C and BX-C genes occur in four gene clusters, located on different chromosomes, that are collectively referred to as the Hox complex (Hox-C). The interrelated functions of the genes in the ANT-C and BX-C and their expression patterns suggest that these gene complexes were initially adjacent to each other in the genome and became separated during evolution. For this reason, these two *Drosophila* gene clusters are often collectively referred to as the *Hom* complex (Hom-C). For simplicity, we will refer to these complexes of selector genes in both flies and mammals as Hox complexes. The development of the brain in both flies and mammals is controlled by another set of homologous selector genes. The function of these genes is not as well understood as the Hox genes and will not be discussed here.

Combinations of Different Hox Proteins Contribute to Specifying Parasegment Identity in *Drosophila*

To illustrate the function of Hox genes in *Drosophila*, we focus our discussion on the genes in the bithorax complex

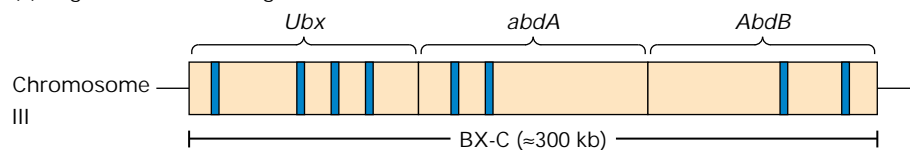
(BX-C). Specification of the identity of parasegments 5–14 in the fly embryo requires the BX-C, and removal of the entire complex leads to embryonic death. However, by analyzing the cuticle of dead mutant larvae, investigators have assessed the role of the three BX-C gene products in specifying various parasegments along the anteroposterior axis. Deletion of the entire BX-C causes transformation of parasegments 5–13 into parasegment 4 (Figure 14-35a). Since other genes also contribute to the specificity of parasegment 14 identity, it develops abnormally but does not assume a parasegment 4 identity. In a sense, then, the BX-C represses parasegment 4 identity and allows the more posterior parasegments to be specified.

Analysis of various double and single mutants permits the contribution of the individual BX-C genes to be assessed. For instance, if both *abdA* and *AbdB* functions are removed by mutation leaving only *Ubx*, parasegments 4–6 develop normally, whereas parasegments 7–13 are transformed into parasegment 6 (Figure 14-35b). In contrast, loss-of-function mutations in *AbdB* permit normal development of parasegments 4–9 but cause parasegments 10–13 to assume parasegment 9 identity (Figure 14-35c).

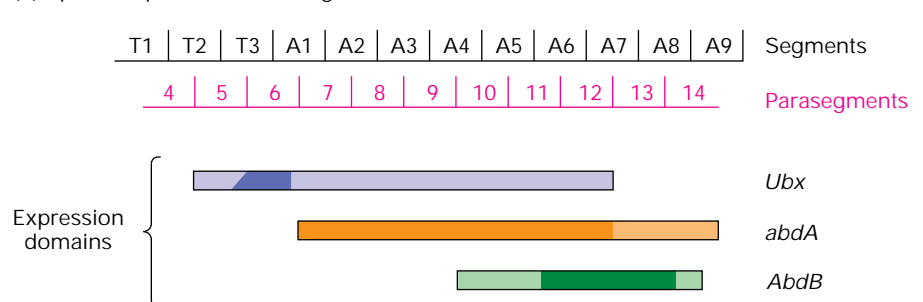
As noted already, the *Ubx*, *abdA*, and *AbdB* genes are transcribed in the same direction and their order of expression along the body axis corresponds to their order within the complex itself (see Figure 14-34b). Although the significance

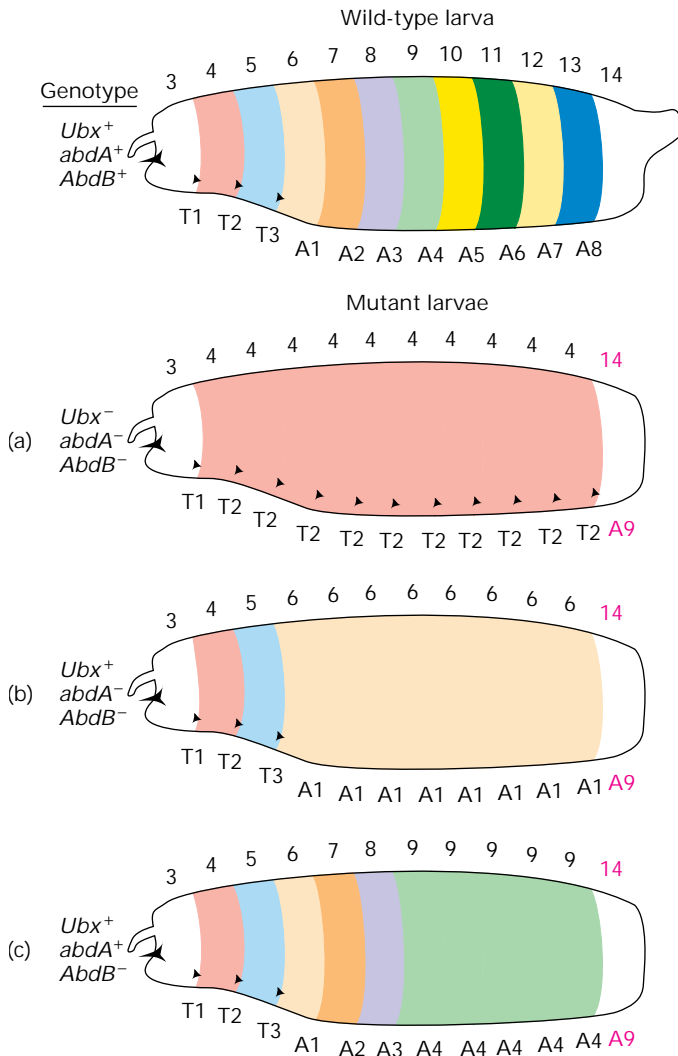
► **FIGURE 14-34** (a) Organization of genes within the bithorax complex (BX-C) on *Drosophila* chromosome III. Transcription of all three genes is from right to left. The exons (blue) make up a relatively small part of the BX-C. The large introns (tan) play an important role in regulating the specific spatial and temporal patterns of transcription of the BX-C genes along the anteroposterior axis. The intron/exon structure of *AbdB* is uncertain. (b) Expression patterns of the products (mRNAs and/or proteins) of the *Ultrabithorax* (*Ubx*), *abdominal A* (*abdA*), and *Abdominal B* (*AbdB*) genes in parasegments 4–14. Darker levels of shading indicate higher concentrations of gene products. [Part (a) adapted from M. Peifer et al., 1987, *Genes Devel.* 1:891; part (b) adapted from T. A. Kaufman et al., 1990, *Adv. Genet.* 27:309.]

(a) Organization of BX-C genes



(b) Spatial expression of BX-C genes





▲ FIGURE 14-35 Contribution of BX-C genes—*Ubx*, *abdA*, and *AbdB*—to determination of parasegment identity. The numbers above each larva indicate the parasegments; those below, the corresponding segments. The cuticular pattern of larvae is used to assign an identity to each parasegment (PS), which is indicated by color, as depicted in the wild type at the top. Red PS and segment labels indicate abnormal patterns that do not correspond exactly to any found in wild-type larvae. See text for discussion. [Adapted from P. A. Lawrence, 1992, *The Making of a Fly: Genetics of Animal Design*, Blackwell Scientific Publications.]

of this correlation is unclear, the conservation of this order in vertebrate Hox complexes argues that it plays an important role in controlling patterning. Indeed, the “out-of-order” expression that occurs in single mutants lacking a functional *abdA* leads to marked defects in parasegments 10–14; that is, they do not correspond morphologically to any wild-type parasegment. This finding suggests that the products of *Ubx* and *AbdB*, in the absence of the *abdA* gene product, do not provide recognizable patterning informa-

tion in parasegments 10–14. Presumably, during normal development of the epidermis, *Ubx* and *AbdB* are never expressed together without *abdA*. Indeed, analysis of other “out-of-order” mutants indicates that the BC-X genes must be expressed along the body axis in the order *Ubx*, *abdA*, and *AbdB*—their order in the genome—for normal patterning information to be generated.

The cuticular patterns illustrated in Figure 14-35 result from loss-of-function mutations in BX-C genes. Researchers also have assessed the effect of gain-of-function mutations in transgenic flies. For instance, transgenic embryos carrying *Ubx* under control of a heat-shock promoter express the *Ubx* protein uniformly along the anteroposterior axis, whereas in wild-type embryos *Ubx* expression is concentrated in parasegments 5 and 6 (see Figure 14-34b). In these transgenic embryos, parasegments 6–14 form normally, but parasegments 1–5 are transformed into parasegment 6. During normal development, the identity of these anterior parasegments is controlled by the ANT-C.

These gain- and loss-of-function studies reveal a consistent relation between the *Drosophila* selector genes: Genes that are expressed more posteriorly suppress the action of genes that are expressed more anteriorly. Thus ectopic expression of a Hox gene in a region more anterior to its normal expression domain results in an anterior → posterior transformation in morphology. The conservation of this phenomenon in the mouse indicates that it must have evolutionary significance.

Although it is clear that BX-C genes and other Hox genes control specification of tissue along the anteroposterior axis, the mechanisms by which this occurs are poorly understood. Recent studies indicate that the different combinations of Hox proteins in various cells of the embryo contribute to this regionalization. Since Hox genes are transcribed after cell membranes have formed, these genes contribute to patterning in the context of signaling between cells. Patterning of tissues, then, is a dynamic process in which gene expression is tightly linked to the context of cells within tissues (Chapter 23).

Specificity of *Drosophila* Hox-Protein Function Is Mediated by *Exd* Protein

Similar to transcription factors expressed earlier in embryogenesis, the Hox proteins encoded by BX-C and ANT-C genes most likely control different developmental pathways by regulating the expression of different sets of target genes. However, the discovery that different Hox proteins bind with high affinity to the same short DNA sequences, which are found on average once every kilobase, seemed incompatible with this mechanism. Recent genetic and molecular experiments show that the ability of Hox proteins to control expression of different genes depends on the product of the *extradenticle* (*exd*) gene.

In *Drosophila* embryos with loss-of-function mutations in *exd*, the Hox genes are expressed as in wild-type embryos,

but the structures controlled by them do not develop normally. This finding suggested that the homeodomain protein encoded by *exd* may act in combination with Hox proteins to control transcription of specific target genes. Exd protein has been shown to dimerize with different Hox proteins, forming heterodimers that exhibit different DNA-binding specificities. For example, the Hox proteins Labial (Lab) and Deformed (Dfd), which are encoded in the ANT-C, both bind to the same DNA sites, but Lab-Exd and Dfd-Exd heterodimers each bind selectively to a specific unique sequence (Figure 14-36). Thus, specific interactions between different Hox proteins and Exd may lead to conformational changes unique to each heterodimer, leading to different DNA-binding specificities. Recent studies have shown that Exd may also contribute to the specificity of Hox function by converting bound Hox proteins from repressors to activators. In addition, Exd may act in a Hox-independent fashion to repress yet other genes.

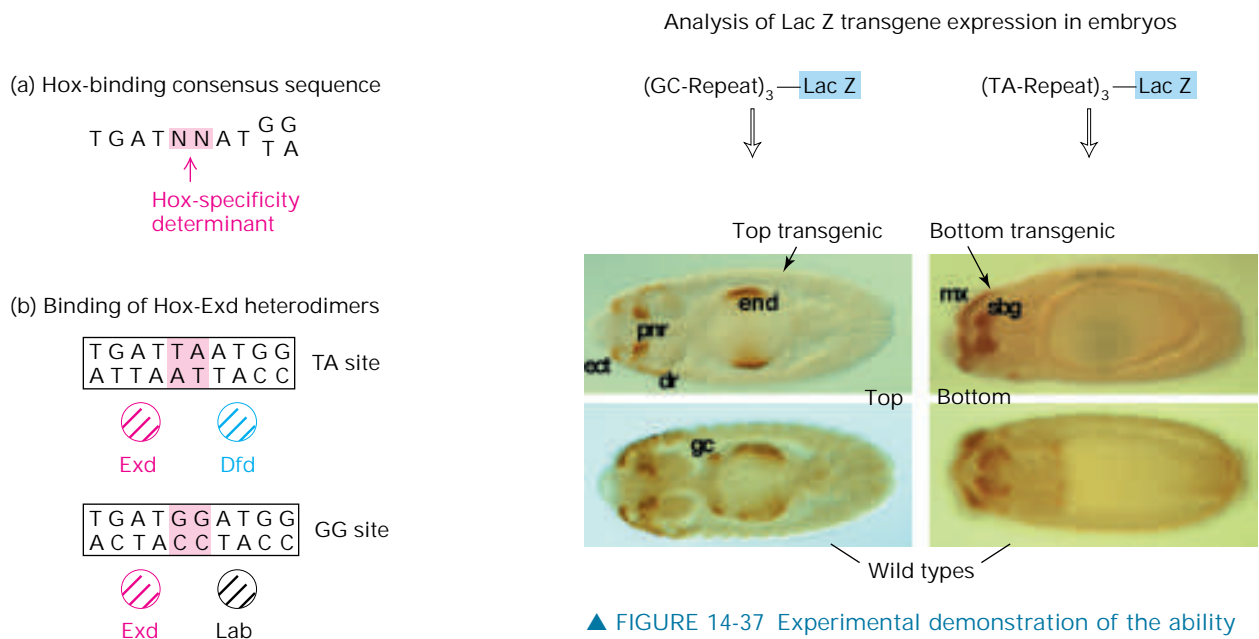
Hox-Gene Expression Is Maintained by Autoregulation and Changes in Chromatin Structure

As mentioned earlier, many of the early-patterning proteins necessary for establishment of specific patterns of Hox-gene expression are produced only transiently during embryogenesis. However, once Hox genes are turned on, they must con-

tinue to be transcribed in specific regions, as Hox proteins are required throughout development and into adult stages.

The transcription-control regions of some Hox genes contain binding sites for their encoded proteins. Thus these Hox proteins help maintain their own expression through an autoregulatory loop. The *lab* and *Dfd* genes discussed in the previous section provide examples of this phenomenon. The DNA site specific for Lab-Exd heterodimer (GG site) is present in the transcription-control region of the *lab* gene; likewise, the Dfd-Exd-binding site (TA site) is present in the *Dfd* control region. To demonstrate autoregulation of *lab* and *Dfd* in vivo, transgenes carrying the *lacZ* gene linked to tandem arrays of either the GG site or TA site were constructed. When these transgenes were introduced into flies, β -galactosidase was expressed in the same regions of fly embryos as the Lab and Dfd proteins (Figure 14-37).

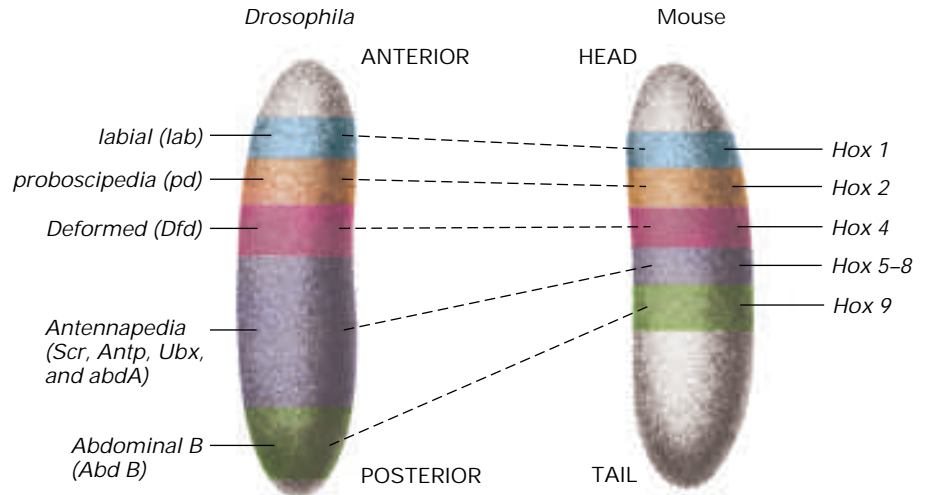
Another mechanism for maintaining normal patterns of Hox-gene expression requires proteins that modulate chromatin structure. These proteins are encoded by two classes of genes referred to as the *trithorax* group and *polycomb*



▲ FIGURE 14-36 Role of Exd protein in conferring DNA-binding specificity on *Drosophila* Hox proteins. (a) Various Hox proteins, including Dfd and Lab, bind to a 10-bp consensus sequence that differs in the nucleotides (N) at the two central positions. (b) Exd-Dfd and Exd-Lab heterodimers specifically recognize Hox-binding sites in which the central dinucleotide is TA or GG, respectively.

▲ FIGURE 14-37 Experimental demonstration of the ability of Hox-binding sites to direct gene expression in specific regions of *Drosophila* embryos. Transgenes containing the *lacZ* gene linked to three copies of the Lab-specific GG site or Dfd-specific TA site from *Drosophila* were introduced into flies. The pattern of expression of β -galactosidase in the transgenic fly embryos mimicked the Lab and Dfd expression patterns, respectively, in wild-type embryos. [From S.-K. Chan et al., 1997, *Development* 124(1):2007; courtesy of R. S. Mann.]

► FIGURE 14-39 Schematic diagram depicting expression domains of the indicated Hox genes in *Drosophila* and mouse embryos. Note that in the domain labeled *Antennapedia*, four genes are expressed; these are analogous to *Hox5–Hox8*. Dashed lines connect corresponding regions in *Drosophila* and mouse. [Adapted from M. McGinnis and M. Kuziora, 1994, *Sci. Am.* 270(2):58; drawing by Tomo Narashima, © 1994, Scientific American, Inc. All rights reserved.]



However, the discovery of mammalian homologs of *exd* (called *Pbx* in mammals) and of *polycomb*-group and *trithorax*-group genes suggests that the mechanisms for controlling the DNA-binding specificity of Hox proteins and for maintaining their expression may be similar in *Drosophila* and higher organisms.

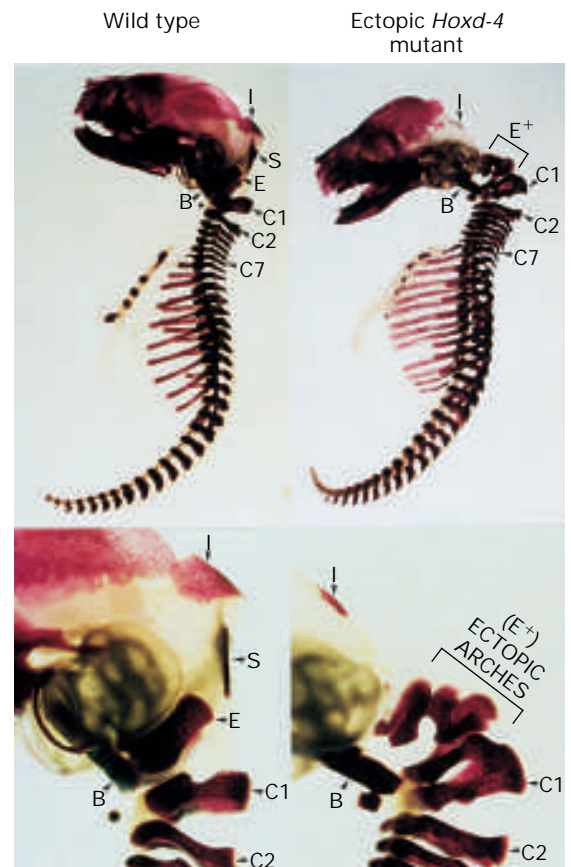
Mutations in Hox Genes Result in Homeotic Transformations in the Developing Mouse

The genes composing the *Drosophila* Hox complexes were isolated based on developmental defects observed in flies carrying mutations within these genes. In contrast, the mammalian Hox genes were isolated based on their homology with the fly genes not on their function. If mammalian Hox genes are functionally equivalent to the fly genes, then mutations in the mammalian genes would be expected to produce homeotic transformations along the body axis. Gene-knockout and transgenic technology have been used to assess the functional role of several Hox genes in controlling regional identity in the mouse. The results of these studies support the view that the Hox genes in mammals and flies play qualitatively similar roles in controlling regional identity along the anteroposterior axis.

Gain-of-Function Mutations in *Hoxd-4* Gene The *Hoxd-4* protein is a homolog of the *Drosophila* *Dfd* protein, which is encoded within the ANT-C (see Figure 14-38). The anterior border of *Hoxd-4* expression in the mouse embryo includes

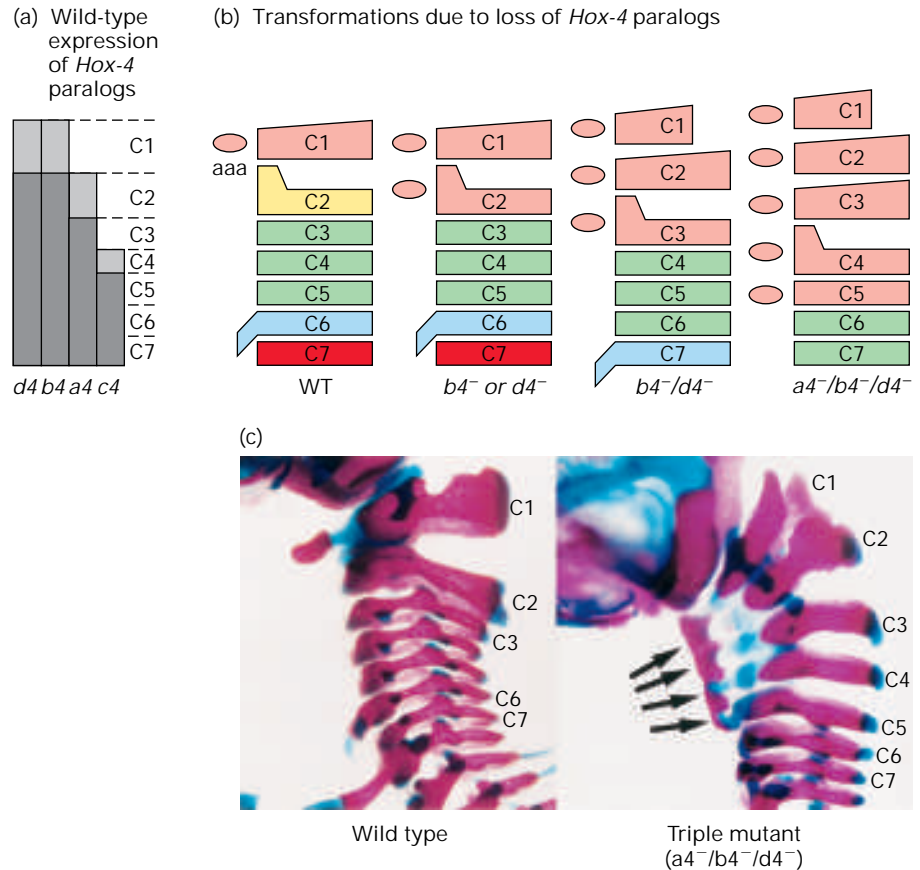
precursor cells that normally give rise to cervical vertebrae. Ectopic expression of *Hoxd-4* in a domain anterior to this region (in a region giving rise to occipital vertebrae) was achieved by fusing *Hoxd-4* cDNA to the regulatory sequences that control transcription of *Hoxa-1*, which normally is expressed more anteriorly. In newborn transgenic mice carrying this construct, which has the effect of a gain-of-function mutation, anterior (occipital) vertebrae are morphologically similar to more posterior (cervical) vertebrae (Figure 14-40).

► FIGURE 14-40 Effect of ectopic expression of *Hoxd-4* protein anterior to its usual expression domain on development of anterior structures in mouse embryos. The photographs show skeletal preparations of a normal newborn mouse and a transgenic newborn that misexpresses *Hoxd-4*; the lower photographs are higher magnifications of the upper ones. The transgenic mouse exhibits a reduction or loss of the supraoccipital (S) and extraoccipital (E) bones of the skull and the appearance of bony structures called ectopic arches (E^+), which are fused to the basioccipital (B) bone; these structures appear similar to neural arches characteristic of more posterior vertebrae. The cervical vertebrae are indicated by C followed by a number. [From T. Lufkin et al., 1992, *Nature* 359:835.]



► **FIGURE 14-41** Effect of loss-of-function mutations in *Hox-4* paralogs on development of cervical vertebrae in mice.

Double and triple mutants were obtained by interbreeding mice carrying a knockout mutation in a single paralog. (a) Expression domains of the *Hox-4* paralogous proteins in prevertebrae with respect to the cervical vertebrae (C1–C7). Gray indicates lower expression than black. (b) Schematic diagrams showing that sequential removal of *Hox-4* paralogs leads to progressive posterior → anterior transformations. In embryos with homozygous mutations in *Hoxb-4* or *Hoxd-4*, C2 is transformed into C1; in mutants lacking both these paralogs, C2 and C3 are transformed into C1. In triple mutants lacking functional *Hoxb-4*, *Hoxd-4*, and *Hoxa-4*, C2–C5 are transformed into C1. In double and triple mutants abnormalities are seen in the structure of C1. These are indicated by the shorter bar. (c) Skeletal structures of a wild-type mouse embryo and a triple mutant homozygous for loss-of-function mutations in the *Hox-4a*, *Hox-4b*, and *Hox-4d* paralogs. The mutant has multiple anterior arch-like structure (arrows) that extend more posterior than normal. aaa = anterior arch of the atlas; ex = exoccipital bone. [From G. S. B. Horan et al., 1995, *Genes & Devel.* 9:1667.]



This homeotic transformation is similar to those resulting from ectopic expression of *Drosophila* Hox genes discussed previously; that is, ectopic expression of a posteriorly localized Hox gene product in more anterior regions leads to an anterior → posterior transformation.

Loss-of-Function Mutations in *Hox-4* Genes Assessing the function of mammalian Hox genes from loss-of-function studies is complicated by the presence of multiple copies (paralogs) of these genes. For instance, to adequately assess the individual contributions of the four *Hox-4* paralogs to anteroposterior patterning in mice, scientists generated double and triple mutants carrying various combinations of mutated *Hoxa-4*, *Hoxb-4*, and *Hoxd-4*. The *Hox-4* paralogs have overlapping but not identical expression domains in the cervical region of the prevertebrae of mouse embryos (Figure 14-41a). By analyzing the morphologic characteristics of the cervical vertebrae in various mutants, researchers have found that vertebrae assume a progressively more anterior morphology as the dosage of *hox-4* paralogs decreases in a particular prevertebral region (Figure 14-41b,c). Therefore, as in flies, Hox genes repress the activity of other, more anteriorly expressed Hox genes. In mouse, this phenomenon is referred to as *posterior prevalence*. Careful analyses of the transformations seen in different single and double mutants demonstrate that different paralogs control overlapping, but

distinct, patterning functions; that is, development of some morphologic features requires two paralogs, whereas development of others requires only a single paralog.

Genetic studies also suggest that, in addition to interactions between paralogous genes, there are similarities in the phenotypes of knockouts of some nonparalogous genes. Hence, Hox genes from different paralogous groups act in a combinatorial fashion to specify different axial patterns. As in flies, the mechanism by which these genes coordinate cellular patterning by controlling the expression of specific target genes remains largely unknown.

SUMMARY **Anteroposterior Specification during Embryogenesis**

- Gradients of transcription factors are generated in the early *Drosophila* embryo through translation of spatially restricted maternal mRNAs and subsequent diffusion of the encoded proteins through the common cytoplasm of the syncytial blastoderm. These proteins, in turn, control the patterned expression of specific target genes along the anteroposterior axis.
- Different target genes contain different combinations of transcription-control sequences that bind specific

transcriptional activators and repressors. The expression patterns characteristic of different target genes can be controlled by particular combinations of protein-binding sites in cis-acting control regions, as well as by differences in the number and affinity of DNA-binding sites for the same transcription factor (see Figure 14-31).

- Early-patterning events, utilizing maternal, gap, pair-rule, and segment-polarity genes, generate a unique pattern of transcription factors expressed in different regions along the anteroposterior axis of *Drosophila* embryos (see Figure 14-32). These transcription factors are expressed transiently and play an essential role in establishing the domains in which different selector genes are expressed.
- Selector genes, which are expressed and required continuously throughout development, direct further development of tissues to form the structures and organs characteristic of each part of the body. The Hox selector genes control the unique morphologic characteristics of different regions along the anteroposterior axis of the major phyla, including arthropods (e.g., insects) and chordates (e.g., mammals).
- Misexpression of Hox genes causes homeotic transformation—the development of body parts in abnormal positions.
- The specificity of Hox proteins is controlled in part by the formation of heterodimers between individual Hox proteins and Exd protein in *Drosophila* or Pbx in mammals (see Figure 14-36). Since different Hox-Exd dimers have different DNA-binding specificities, they control expression of different sets of genes.
- Maintenance of the Hox expression patterns occurs through positive autoregulatory loops involving the Hox genes themselves, and through modulation of chromatin by proteins encoded by both *polycomb*-group and *trithorax*-group genes.

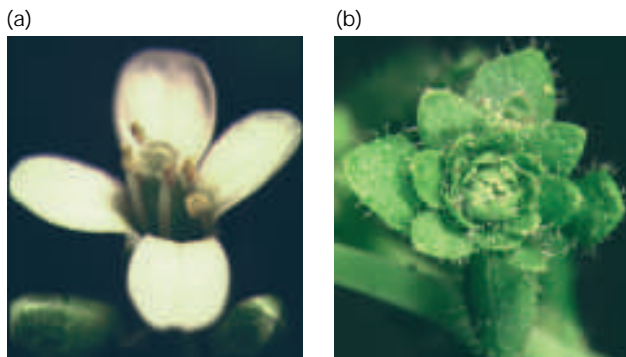
14.4 Specification of Floral-Organ Identity in *Arabidopsis*



Important strides have been made in dissecting the mechanisms controlling the development of plants. These advances have been possible largely due to the choice of *Arabidopsis thaliana* as a model plant. This plant has many of the same advantages as flies and worms for use as a model system. First, *Arabidopsis* is small and easy to maintain in the laboratory. Second, mutants can be easily produced because *Arabidopsis* has a short generation time (8 weeks) and seeds can be mutagenized by treatment with chemicals or radiation. In addition, the small size of the *Arabidopsis* genome facilitates positional cloning methods to isolate the genes defined by mutations (Chapter 8). And finally, transgenic *Arabidopsis* plants are readily made (see Figure 8-38). Although many different aspects of development are being investigated in this plant, our discussion will focus on the transcription-control mechanisms regulating the formation of flowers. As we will see, these mechanisms are strikingly similar to those controlling cell-type and anteroposterior regional specification in yeast and animals.

Flowers Contain Four Different Organs

Flowers comprise four different organs called sepals, petals, stamens and carpels, arranged in concentric circles called whorls. The number and type of floral organs and the number of whorls vary among plant species. *Arabidopsis* has a complete set of floral organs, including four sepals in whorl 1, four petals in whorl 2, six stamens in whorl 3, and two carpels containing ovaries in whorl 4 (Figure 14-42a). These organs form from a collection of undifferentiated, morphologically indistinguishable cells called the *floral meristem*. As cells within the center of the floral meristem divide, four concentric rings of primordia form sequentially. The outer ring primordia, which give rise to the sepals of whorl 1,



◀ FIGURE 14-42 (a) Flowers of wild-type *Arabidopsis thaliana* have four sepals in whorl 1, four petals in whorl 2, six stamens in whorl 3, and two carpels in whorl 4. (b) In *Arabidopsis* with mutations in all three classes of floral organ-identity genes, the four floral organs are transformed into leaf-like structures. [From D. Weigel and E. M. Meyerowitz, 1994, *Cell* 78:203. courtesy of E. M. Meyerowitz.]

form first, followed by the primordia giving rise to the petals, then the stamens and carpels primordia. Genetic studies have shown that normal flower development requires three classes of *floral organ-identity genes*. In plants with mutations in all three classes, concentric whorls of leaf-like structures replace the floral organs, indicating that these structures are modified leaves (Figure 14-42b).

Three Classes of Genes Control Floral-Organ Identity

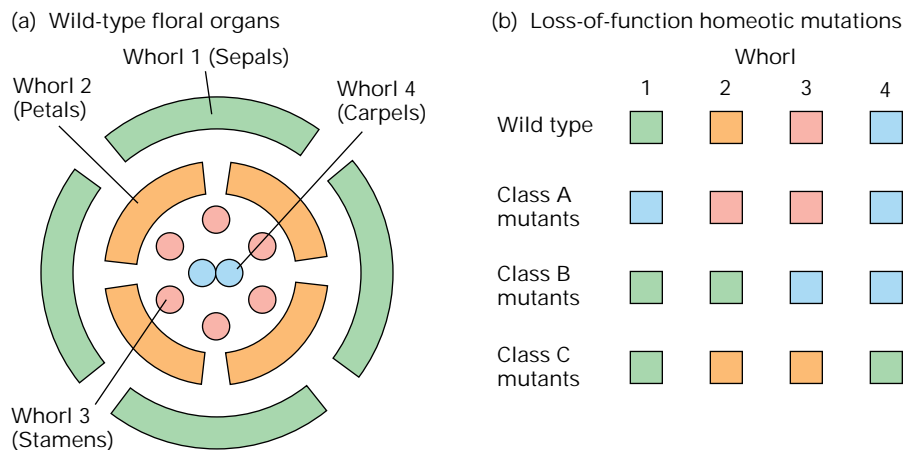
Analyses of mutations causing transformation of one floral organ into another led to identification of floral organ-identity genes. These homeotic mutations are equivalent to the homeotic mutations in flies and mammals in which one part of the body is replaced by another. In these studies, researchers mutagenized *Arabidopsis* seeds. Each seed contains two cells, each of which contributes to all the tissues in the adult plant. In a mutagenized seed, mutations are randomly dispersed in the genome of each seed. Thus, after germinating, a seed gives rise to a mosaic plant. Self-fertilization generates plants in which one-quarter are homozygous for the induced mutations. Mutant flowers in which transformation in floral-organ identity has taken place are easily recognized.

As illustrated in Figure 14-43, three classes of loss-of-function mutations (A, B, and C) cause floral-organ transformations in *Arabidopsis*. In plants lacking all A, B, and C function, the floral organs develop as leaves. Based on the phenotypes of these loss-of-function mutants, scientists have proposed a model to explain how three classes of genes control floral-organ identity. According to this ABC model, class A genes specify sepal identity in whorl 1 and do not require either class B or C genes to do so. Similarly, class C genes

specify carpel identity in whorl 4 and, again, do so independently of class B and C genes. In contrast to these structures, which are specified by only a single class of genes, the petals in whorl 2 are specified by class A and B genes, and the stamens in whorl 3 are specified by class B and C genes. To account for the observed effects of removing A genes or C genes, the model also postulates that A genes repress C in whorls 1 and 2 and, conversely, C genes repress A genes in whorls 3 and 4.

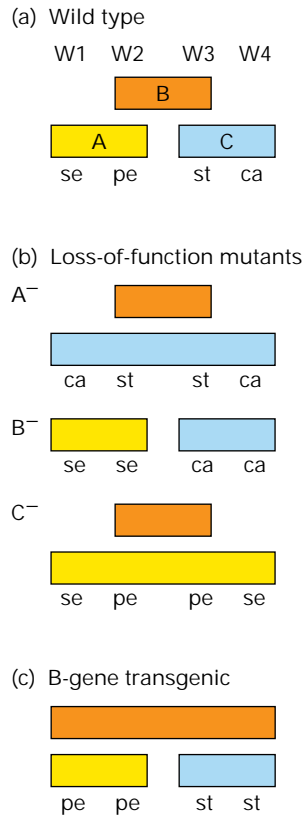
To determine if the actual expression patterns of class A, B, and C genes are consistent with this model, researchers cloned these genes and assessed the expression patterns of their mRNAs in the four whorls in wild-type *Arabidopsis* plants and in loss-of-function mutants (Figure 14-44a,b). Consistent with the ABC model, A genes are expressed in whorls 1 and 2, B genes in whorls 2 and 3, and C genes in whorls 3 and 4. Furthermore, in class A mutants, C class genes are also expressed in organ primordia of whorls 1 and 2; similarly, in class C mutants, class A genes are also expressed in whorls 3 and 4. These findings are consistent with the homeotic transformations observed in these mutants.

To test whether these patterns of expression are functionally important, scientists produced transgenic *Arabidopsis* plants in which floral organ-identity genes were expressed in inappropriate whorls. For instance, introduction of a transgene carrying class B genes linked to an A-class promoter leads to ubiquitous expression of class B genes in all whorls (Figure 14-44c). In such transgenics, whorl 1, now under the control of class A and B genes, develops into petals instead of sepals; likewise, whorl 4, under the control of both class B and C genes, gives rise to stamens instead of carpels. These results support the functional importance of the ABC model for specifying floral identity.



▲ FIGURE 14-43 Identification of three classes of genes that control specification of floral organs in *Arabidopsis*. (a) Schematic diagram of the arrangement of wild-type floral organs, which are found in concentric circles called whorls. (b) Effect of loss-of-function mutations leading to transformations of one organ into another. Class A mutations affect organ identity

in whorls 1 and 2: sepals (green) become carpels (blue) and petals (orange) become stamens (pink). Class B mutations cause transformation of whorls 2 and 3: petals become sepals and stamens become carpels. In class C mutations, whorls 3 and 4 are transformed: stamens become petals and carpels become sepals. [See D. Wiegand and E. M. Meyerowitz, 1994, *Cell* 78:203.]



▲ FIGURE 14-44 Expression patterns for three classes of floral organ-identity genes in wild-type *Arabidopsis* (a), loss-of-function mutants (b), and a transgenic that misexpresses class B genes (c). Colored bars represent the A, B, and C mRNAs in each whorl (W1, W2, W3, W4). The observed floral organ in each whorl is indicated below: sepal = se; petals = pe; stamens = st; and carpels = ca. See text for discussion. [See D. Wiegand and E. M. Meyerowitz, 1994, *Cell* 78:203; B. A. Krizek and E. M. Meyerowitz, 1996, *Development* 122:11.]

Many Floral Organ-Identity Genes Encode MADS Family Transcription Factors

Sequencing of floral organ-identity genes has revealed that many encode proteins belonging to the MADS family of transcription factors. As discussed in earlier sections, members of this class of proteins, including MCM1 in yeast and MEFs in muscle, form homo- and heterodimers. Thus floral-organ identity may be specified by a combinatorial mechanism in which differences in the activities of different homo- and heterodimeric forms of various A, B, and C proteins regulate the expression of subordinate downstream genes necessary for the formation of the different cell types in each organ. Alternatively, these proteins may act solely as homodimers with organ identity reflecting the activities of different combinations of genes regulated by these homodimers in different organ primordia.

Biochemical experiments have not yet resolved the mechanisms by which organ diversity is regulated by floral organ-

identity genes. The proteins encoded by two class B genes—*pistillata* and *apetala3*—have been shown to form heterodimers that bind to specific DNA sequences. However, neither the *Pistillata* nor *Apetala3* protein forms DNA-binding heterodimers with class A gene products (e.g., *Apetala1*) or class C gene products (e.g., *Agamous*). Hence, the differences between the identity of whorls 2 and 3 (giving rise to petals and stamens, respectively) does not appear to represent the formation of different heterodimeric combinations of the proteins encoded by the three classes of mutation-defined floral organ-identity genes.

SUMMARY Specification of Floral-Organ Identity in *Arabidopsis*

- Flowers have four different organs—petals, sepals, stamens and carpels—which develop in concentric whorls. These organs are modified leaves.
- Genetic studies have identified three classes of genes that participate in specifying floral-organ identity. The ABC model proposes how these classes of genes work together to specify organ identity in each whorl.
- Class A, B, and C genes encode transcription factors of the MADS family. The patterns of expression of these gene products are consistent with the ABC genetic model (see Figure 14-44). However, the mechanisms by which these proteins regulate development of the cell types composing each organ remain to be elucidated.

PERSPECTIVES for the Future

During the past two decades remarkable progress has been made in understanding the mechanisms regulating gene transcription during development. Through a combination of genetic and biochemical studies, the DNA sequences regulating expression of many different genes in simple and higher eukaryotes and the transcription factors that bind to them have been identified. Dissection of developmental programs in various organisms has shown that they depend not only on combinations of transcription factors expressed by specific cells but also on interaction between these cells and extracellular signals acting through membrane receptors or through their direct interaction with transcription factors. Continuation of such studies will provide additional insight into the way specific genes are regulated in different places and times in development and the strategies by which the organism utilizes the genome to specify the large number of different cells and tissues in an organism.

Two important developments are likely to change radically the nature and extent of our knowledge of transcriptional regulatory programs controlling development. First, sequencing of the entire genomes of several higher eukaryotes, including *Drosophila*, the mouse, and *Arabidopsis*, most

likely will be completed within a few years. Indeed, the sequence of *C. elegans* genome has been completed already. The entire sequence of the human genome should be determined within the next 5 years. Second, new technology will provide remarkably efficient ways to monitor expression of many genes at one time, indeed in principle, the entire transcriptional program of an individual cell.

The prototype for this level of analysis comes from studies in the yeast *S. cerevisiae*, whose genome has been sequenced in its entirety (Chapter 7). Monitoring of the transcriptional program of a vast number of genes as a function of cell type and physiologic state has already been accomplished using DNA microarray technology (Chapter 7). In this approach, robotics technology is used to attach an array of thousands of different DNA sequences to a microscope slide. cDNAs from small numbers of cells synthesized in the presence of fluorescent nucleotides then are hybridized to the DNA sequences attached to the microscope slide. The amount of fluorescent-labeled DNA bound to each sequence attached to the slide is read using a confocal scanning microscope. Gene expression in yeasts of different genotypes, of different mating types, or under different physiologic conditions can be compared by labeling the cDNAs from different populations of cells with different fluorescent tags. The ratio of these probes bound to specific sequences in the microarray will reveal differences in expression at the same loci under different conditions. The microarray technology is rapidly advancing and is already being used to assess gene expression in higher eukaryotes.

Where will these studies lead? In the not-too-distant future, researchers will be able to document the expression of perhaps all the genes in a given cell type, the putative regulatory sequences controlling their expression, and the complete set of DNA-binding transcription factors that the cell expresses. This new knowledge will provide starting points for the biochemical and genetic dissection of many hitherto unstudied transcriptional regulatory regions. It will also provide an enormous data base for eventually deciphering the genetic strategy used by an organism as a whole to generate the many different cell types characteristic of higher eukaryotes.

PERSPECTIVES in the Literature

You are working with a well-characterized in vitro system that allows you to induce myoblasts to differentiate into myotubes synchronously. Your long term goal is to describe the transcriptional regulatory networks that control muscle differentiation. An important step in your studies is to describe patterns of gene expression. You are interested in studying which genes are turned on and off and in what order during the transition from a myoblast to a differentiated myotube. Design a set of experiments that would allow you to carry out these studies. To get started, see the following references:

Brown, P. O., and D. Botstein. 1999. Exploring the new world of the genome with DNA microarrays. *Nature Genetics* (1 Suppl):33–37.

Chu, S., et al. 1998. The transcriptional program of sporulation in budding yeast. *Science* 282:699–705.

DeRisi, J., V. R. Iyer, and P. O. Brown. 1997. Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* 278:680–686.

Testing Yourself on the Concepts

1. To determine the molecular and cellular basis of patterning mechanisms, what strategies do investigators take in model organisms and in humans?
2. Describe two specific examples of syncytial cells and how transcription and translation during development are affected by these structures.
3. Describe three examples of MADS box transcription factors and their roles in development.
4. Predict the phenotypes of double mutants in *Arabidopsis* that cause a loss of (a) A and B function; (b) B and C function; and (c) A and C function.
5. Describe two specific examples of the effects of chromatin on the transcriptional control of development.
6. Discuss how inhibitory proteins such as Id and Emc function and affect development.

MCAT/GRE-Style Questions

Key Concept Please read the section titled “Multiple Regulation of *HO* Transcription Controls Mating-Type Conversion” (p. 541) and refer to Figure 14-6; then answer the following questions:

1. The fact that Cdc28–G₁ cyclin complexes are required for CCBF activity implies that
 - a. CCBF phosphorylation by the Cdc28–G₁ cyclin complex activates *HO* transcription.
 - b. CCBF dephosphorylation by the Cdc28–G₁ cyclin complex activates *HO* transcription.
 - c. CCBF phosphorylation by the Cdc28–G₁ cyclin complex induces binding of swi6 binding to the URS2.
 - d. CCBF dephosphorylation by the Cdc28–G₁ cyclin complex induces binding of swi6 binding to the URS2.
2. Ash1 accumulation in daughter cells
 - a. Requires inhibition of swi5.
 - b. Requires expression of other yeast genes.
 - c. Inhibits swi5.
 - d. Degrades swi5 protein.

3. What is the phenotype of a *swi1* mutant?
 - a. It is defective in mating.
 - b. It is defective in the mating-type switch.
 - c. The mating-type switch occurs in both progeny.
 - d. Lethality.
4. What is the phenotype of $\alpha 2$ or $\alpha 1$ mutants?
 - a. They are defective in mating.
 - b. They are defective in the mating-type switch.
 - c. The mating-type switch occurs in both progeny.
 - d. Lethality.
5. A promoter-reporter gene construct carrying the URS1 and URS2 sequences would be expressed in
 - a. Mother cells only.
 - b. Daughter cells only.
 - c. Both mother and daughter cells.
 - d. Yeast not undergoing mating-type switching.

Key Experiment Please read the section titled “Maternal *bicoid* Gene Specifies Anterior Region in *Drosophila*” (p. 557) and refer to Figure 14-26; then answer the following questions:

6. One maternal copy of a mutated *bicoid* gene containing a deletion of the 3' untranslated end would result in an embryo with *hunchback* transcribed
 - a. At the anterior end only.
 - b. Throughout the embryo.
 - c. At the posterior end only.
 - d. Not at all.
 7. What is the resulting phenotype of the embryo in Figure 14-26d?
 - a. The boundary of anterior structures is increased.
 - b. The boundary of posterior structures is increased.
 - c. Loss of anterior structures.
 - d. None.
 8. The Bicoid protein gradient as compared to the *bicoid* RNA gradient
 - a. Is confined to a smaller area of the embryo.
 - b. Extends more posteriorly.
 - c. Is less dense.
 - d. Is less stable.
 9. The boundary of *hunchback* expression increases posteriorly in an embryo from a mother with one versus two copies of *bicoid* because
 - a. More *bicoid* RNA diffuses throughout the embryo.
 - b. More Bicoid protein diffuses throughout the embryo.
 - c. More Hunchback protein diffuses throughout the embryo.
 - d. More high-affinity sites are present in the *hunchback* promoter.
 10. If multiple copies of the synthetic high-affinity promoter-reporter gene construct were placed in a maternal background containing no copies of *bicoid*, what would the expression pattern of the reporter gene look like?
 - a. A small amount localized to the anterior end.
 - b. A large amount localized to the anterior end.
 - c. A large amount extended more posteriorly.
 - d. No expression.
- Key Application** A prenatal diagnosis strategy for detecting muscular diseases that involves testing the capability of embryonic cells to form muscle after inducing MyoD expression in vitro has been suggested. Please read the sections titled “Myogenic Genes Were First Identified in Studies with Cultured Fibroblasts” (p. 544) and “Myogenic Proteins Are Transcription Factors Containing a Common bHLH Domain” (p. 546) and answer the following questions:
11. Normal embryonic cells from humans would likely form myoblasts in culture after induction of MyoD expression because
 - a. MyoD is known to convert fibroblasts to muscle in vitro.
 - b. Embryonic cells are similar in nature to azamyoblasts.
 - c. The B-cell development pathway cannot be stimulated in these cells.
 - d. E2A is not expressed in these cells.
 12. Promoters containing E boxes
 - a. Are only found in muscle-specific genes.
 - b. Are rare.
 - c. Are always occupied by E2A/MyoD heterodimers.
 - d. Always confer expression in muscle cells.
 13. Myogenin, Myf5, Mrf4, and MyoD
 - a. Have the same function.
 - b. Dimerize and bind E boxes.
 - c. Are expressed at the same time during development.
 - d. Have divergent structures.
 14. For a fibroblast to convert to an azomyoblast, which of the following must be present:
 - a. 5-Azacytidine.
 - b. MyoD and E2A.
 - c. E2A.
 - d. An MRF and E2A.

Key Terms

autoregulatory loop 567	homeosis 554
azamyoblasts 544	Hox complex 565
cell-type specification 538	knockin 548
determination 543	MAT locus 541
differentiation 543	morphogen 556

muscle enhancer-binding factors (MEFs) 546
 muscle regulatory factors (MRFs) 546
 myoblast 543
 myotube 544
 parasegment 561
 proneural cluster 552
 selector genes 563
 sensory organ precursor (SOP) 552
 syncytium 544
 whorl 571

References

Cell-Type Specification and Mating-Type Conversion in Yeast

- Bobola, N., et al. 1996. Asymmetric accumulation of Ash1p in postnaphase nuclei depends on a myosin and restricts yeast mating-type switching to mother cells. *Cell* 84:699–709.
- Herskowitz, I., J. Rine, and J. Strathern. 1992. Mating-type determination and mating-type interconversion in *Saccharomyces cerevisiae*. In *The Molecular and Cellular Biology of the Yeast Saccharomyces: Gene Expression*, Vol. II. Cold Spring Harbor Laboratory Press.
- Keleher, C. A., et al. 1992. Ssn6-Tup1 is a general repressor of transcription in yeast. *Cell* 68:709–719.
- Li, T., et al. 1995. Crystal structure of the MATa1/MAT α 2 homeodomain heterodimer bound to DNA. *Science* 270:262–269.
- MacKay, V., and T. R. Manney. 1974. Mutations affecting sexual conjugation and related processes in *Saccharomyces cerevisiae*. I. Isolation and phenotypic characterization of nonmating mutants. *Genetics* 76:255–271.
- Sil, A. and I. Herskowitz. 1996. Identification of an asymmetrically localized determinant, Ash1p, required for lineage-specific transcription of the yeast *HO* gene. *Cell* 84:711–722.
- Smith, D. L., and A. D. Johnson. 1992. A molecular mechanism for combinatorial control in yeast: MCM1 protein sets the spacing and orientation of the homeodomains of an alpha 2 dimer. *Cell* 68:133–142.
- Song, D., et al. 1991. Pheromone-dependent phosphorylation of the yeast STE12 protein correlates with transcriptional activation. *Genes & Devel.* 5:741–750.
- Takizawa, P. A., et al. 1997. Actin-dependent localization of an RNA encoding a cell-fate determinant in yeast. *Nature* 389:90–93.
- Vershon, A. K. and A. D. Johnson. 1993. A short, disordered protein region mediates interactions between the homeodomain of the yeast alpha 2 protein and the MCM1 protein. *Cell* 72:105–112.
- Wolberger, C., et al. 1991. Crystal structure of a MAT alpha 2 homeodomain-operator complex suggests a general model for homeodomain-DNA interactions. *Cell* 67:517–528.

Cell-Type Specification in Animals

- Benezra, R., et al. 1990. The protein Id: a negative regulator of helix-loop-helix DNA-binding proteins. *Cell* 61:49–59.
- Bour, B. A., et al. 1995. *Drosophila* MEF2, a transcription factor that is essential for myogenesis. *Genes & Devel.* 9:730–741.
- Davis, R. L., H. Weintraub, and A. B. Lassar. 1987. Expression of a single transfected cDNA converts fibroblasts to myoblasts. *Cell* 51:987–1000.
- Ellis, H. M., D. R. Spann, and J. W. Posakony. 1990. Extramacrochaetae, a negative regulator of sensory organ development in *Drosophila*, defines a new class of helix-loop-helix proteins. *Cell* 61:27–38.

- Lee, J. L., et al. 1995. Conversion of *Xenopus* ectoderm into neurons by NeuroD, a basic helix-loop-helix protein. *Science* 268:836–844.
- Ma, Q., C. Kintner, and D. J. Anderson. 1996. Identification of neurogenin, a vertebrate neuronal determination gene. *Cell* 87:43–52.
- Mead, P. E., et al. 1998. SCL specifies hematopoietic mesoderm in *Xenopus* embryos. *Development* 125:2611–2620.
- Molkentin, J. D., et al. 1995. Cooperative activation of muscle gene expression by MEF2 and myogenic bHLH proteins. *Cell* 83:1125–1136.
- Nabeshima, Y., et al. 1993. Myogenin gene disruption results in perinatal lethality because of severe muscle defect. *Nature* 364:532–535.
- Rawls, A., and E. N. Olson. 1997. MyoD meets its maker. *Cell* 89:5–8.
- Rudnicki, M. A., et al. 1993. MyoD or Myf-5 is required for the formation of skeletal muscle. *Cell* 75:1351–1359.
- Skeath, J. B., and S. B. Carroll. 1991. Regulation of *achaete-scute* gene expression and sensory organ pattern formation in the *Drosophila* wing. *Genes & Devel.* 5:984–995.
- Wang, Y., et al. 1996. Functional redundancy of the muscle-specific transcription factors Myf5 and myogenin. *Nature* 379:823–825.
- Yun, K., and B. Wold. 1996. Skeletal muscle determination and differentiation: story of a core regulatory network and its context. *Curr. Opin. Cell Biol.* 8:877–889.

Anteroposterior Specification during Embryogenesis

- Capecchi, M. R. 1997. The role of Hox genes in hindbrain development. In W. M. Cowan, T. Jessel, and S. L. Zipursky, eds., *Molecular and Cellular Approaches to Neural Development*, Oxford University Press.
- Castelli-Gair, J., and M. Akam. 1995. How the Hox gene *Ultrabithorax* specifies two different segments: the significance of spatial and temporal regulation within metameres. *Development* 121:2973–2982.
- Chan, S-K., et al. 1994. The DNA-binding specificity of Ultrabithorax is modulated by cooperative interactions with Extradenticle, another homeoprotein. *Cell* 78:603–615.
- Driever, W., and C. Nüsslein-Volhard. 1988. A gradient of bicoid protein in *Drosophila* embryos. *Cell* 54:83–93.
- Driever, W., and C. Nüsslein-Volhard. 1988. The bicoid protein determines position in the *Drosophila* embryo in a concentration-dependent manner. *Cell* 54:95–104.
- Dubnau, J., and G. Struhl. 1996. RNA recognition and translational regulation by a homeodomain protein. *Nature* 379:694–699.
- Gavis, E. R., and R. Lehmann. 1992. Localization of nanos RNA controls embryonic polarity. *Cell* 71:301–313.
- Lawrence, P. A. 1992. *The Making of the Fly: The Genetics of Animal Design*. Blackwell Scientific Publications.
- Lufkin, T. 1996. Transcriptional control of Hox genes in the vertebrate nervous system. *Curr. Opin. Genet. Devel.* 6:575–580.
- Murata, Y., and R. P. Wharton. 1995. Binding of Pumilio to maternal hunchback mRNA is required for posterior patterning in *Drosophila* embryos. *Cell* 80:747–756.
- Nüsslein-Volhard, C., and E. Wieschaus. 1980. Mutations affecting segment number and polarity in *Drosophila*. *Nature* 287:795–801.
- Orlando, V., and R. Paro. 1995. Chromatin multiprotein complexes involved in the maintenance of transcription patterns. *Curr. Opin. Genet. Devel.* 5:174–179.
- Peifer, M., F. Karch, and W. Bender. 1987. The bithorax complex: control of segmental identity. *Genes & Devel.* 1:891–898.
- Rivera-Pomar, R., et al. 1995. Activation of posterior gap gene expression in the *Drosophila* blastoderm. *Nature* 376:253–256.

- Small, S., et al. 1991. Transcriptional regulation of a pair-rule stripe in *Drosophila*. *Genes & Devel.* 5:827–839.
- St. Johnston, D., and C. Nüsslein-Volhard. 1992. The origin of pattern and polarity in the *Drosophila* embryo. *Cell* 68:201–219.
- Stanojevic, D., S. Small, and M. Levine. 1991. Regulation of a segmentation stripe by overlapping activators and repressors in the *Drosophila* embryo. *Science* 254:1385–1387.
- Struhl, G., P. Johnston, and P. A. Lawrence. 1992. Control of *Drosophila* body pattern by the hunchback morphogen gradient. *Cell* 69:237–249.
- Wreden, C., et al. 1997. Nanos and Pumilio establish embryonic polarity in *Drosophila* by promoting posterior deadenylation of hunchback mRNA. *Development* 124:3015–3023.
- Specification of Floral-Organ Identity in *Arabidopsis***
- Fosket, D. E. 1994. *Plant Growth and Development: A Molecular Approach*. Academic Press.
- Krizek, B. A., and E. M. Meyerowitz. 1996. Mapping the protein regions responsible for the functional specificities of the *Arabidopsis* MADS domain organ-identity proteins. *Proc. Nat'l. Acad. Sci. USA* 93:4063–4070.
- Krizek, B. A., and E. M. Meyerowitz. 1996. The *Arabidopsis* homeotic genes APETALA3 and PISTELLATA are sufficient to provide the B class organ identity function. *Development* 122:11–22.
- Reichman, J. L., and E. M. Meyerowitz. 1997. MADS domain proteins in plant development. *Biol. Chem.* 378:1079–1101.
- Sablowski, R. W., and E. M. Meyerowitz. 1998. A homolog of NO APICAL MERISTEM is an immediate target of the floral homeotic genes APETALA3/PISTILLATA. *Cell* 92:93–103.
- Weigel, D. 1995. The genetics of flower development: from floral induction to ovule morphogenesis. *Ann. Rev. Genet.* 29:19–39.
- Weigel, D., and E. M. Meyerowitz. 1994. The ABCs of floral homeotic genes. *Cell* 78:203–209.
- Wolpert, W. 1998. *Principles of Development*. Oxford University Press (Oxford, Eng.).