GRADIENTS, CASCADES, AND SIGNALING PATHWAYS

6.31.1 Introduction

Key Terms

Axes are straight lines passing through an organism, around which the organism is symmetrically arranged.

The **anterior-posterior axis** is the line running from the head to the tail of an animal.

The **dorsal-ventral axis** is the line running from the back to the belly of an animal.

Development begins with a single fertilized egg, but gives rise to cells that have different developmental fates. The problem of early development is to understand how this asymmetry is introduced: how does a single initial cell give rise within a few cell divisions to progeny cells that have different properties from one another?

The means by which asymmetry is generated varies with the type of organism. The egg itself may be homogeneous, with the acquisition of asymmetry depending on the process of the initial division cycles, as in the case of mammals. Or the egg may have an initial asymmetry in the distribution of its cytoplasmic components, which in turn gives rise to further differences as development proceeds, as in the case of *Drosophila*.

Early development is defined by the formation of **axes**. By whatever means are used to develop asymmetry, the early embryo develops differences along the **anterior-posterior axis** (head-tail) and along the **dorsal-ventral axis** (top-bottom). At the stage of interpreting the axial information, a relatively restricted set of signaling pathways is employed, and essentially the same pathways are found in flies and mammals.

The paradigm for considering the molecular basis for development is to suppose that each cell type may be characterized by its pattern of gene expression, that is, by the particular gene products that it produces. The principal level for controlling gene expression is at transcription, and components of pathways regulating transcription provide an important class of developmental regulators. We may include a variety of activities within the rubric of transcriptional regulators, which could act to change the structure of a promoter region, to initiate transcription at a promoter, to regulate the activity of an enhancer, or indeed sometimes to repress the action of transcription factors. However, the regulators of transcription most often prove to be DNA-binding proteins that activate transcription at particular promoters or enhancers.

GRADIENTS, CASCADES, AND SIGNALING PATHWAYS

6.31.2 Fly development uses a cascade of transcription factors

Key Terms

A **maternal gene (maternal effect gene)** (maternal-effect gene) is usually expressed by the mother during oogenesis. Phenotypes resulting from a maternal-effect mutation depend on the genotype of the mother rather than the genotype of the embryo.

Segmentation genes are concerned with controlling the number or polarity of body segments in insects.

Homeotic genes are defined by mutations that convert one body part into another; for example, an insect leg may replace an antenna.

Key Concepts

- The genes that control the early stages of fly development code for transcription factors.
- At each stage, the factors in one area of the egg control the synthesis of further factors that will define smaller areas.
- Maternal genes are expressed during oogenesis and act in the oocyte.
- Three successive groups of segmentation genes are expressed after fertilization to control the number or polarity of segments.
- Homeotic genes control the identity of a segment.

The systematic manner in which the regulators are turned on and off to form circuits that determine body parts has been worked out in detail in *D. melanogaster*. The basic principle is that a series of events resulting from the initial asymmetry of the egg is translated into the control of gene expression so that specific regions of the egg acquire different properties. The means by which asymmetry is translated into control of gene expression differ for each of four systems that have been characterized in the insect egg. It may involve localization of factors that control transcription or translation within the egg, or localized control of the activities of such factors. But the end result is the same: spatial and temporal regulation of gene expression.

Early in development, the identities of parts of the embryo are determined: regions are defined whose descendants will form particular body parts. The genes that regulate this process are identified by loci in which mutations cause a body part to be absent, to be duplicated, or to develop as another body part. Such loci are prime candidates for genes whose function is to provide regulatory "switches." Most of these genes code for regulators of transcription. They act upon one another in a hierarchical manner, but they act also upon other genes whose products are actually

responsible for the formation of pattern. The ultimate targets are genes that code for kinases, cytoskeletal elements, secreted proteins, and transmembrane receptors.

The establishment of a specific pattern of transcription in a particular region of the embryo leads to a *cascade* of control, when regulatory events are connected so that a gene turned on (or off) at one stage itself controls expression of other genes at the next stage. Formally, such a cascade resembles those described previously for bacteriophages or for bacterial sporulation (as discussed in *Molecular Biology 3.10 The operon*), although it is more complex in the case of eukaryotic development. The common feature of regulatory proteins is that they are transcription factors that regulate the expression of other transcription factors (as well as other target proteins). As in the case of prokaryotic regulation, the basic relationship between the regulator protein and the target gene is that the regulator recognizes a short sequence in the DNA of the promoter (or an enhancer) of a target gene. All of the targets for a particular regulator are identified by their possession of a copy of the appropriate consensus sequence.

The development of an adult organism from a fertilized egg follows a predetermined pathway, in which specific genes are turned on and off at particular times. From the perspective of mechanism, we have most information about the control of transcription. However, subsequent stages of gene expression are also targets for regulation. And, of course, the cascade of gene regulation is connected to other types of signaling, including cell-cell interactions that define boundaries between groups of cells.

The mechanics of development in terms of cellular events are different in different types of species, but we assume that the principle established with *Drosophila* will hold in all cases: that a regulatory cascade determines the appropriate pattern of gene expression in cells of the embryo and ultimately of the adult. Indeed, homologous genes in distantly related organisms play related roles in development. The same pathways are found in (for example) flies and mammals, although the consequences of their employment are rather different in terms of the structures that develop.

Genes involved in regulating development are identified by mutations that are lethal early in development or that cause the development of abnormal structures. A mutation that affects the development of a particular body part attracts our attention because a single body part is a complex structure, requiring expression of a particular set of many genes. Single mutations that influence the structure of the entire body part therefore identify potential regulator genes that switch or select between developmental pathways.

In *Drosophila*, the body part that is analyzed is the segment, the basic unit that can be seen looking at the adult fly. Mutations fall into (at least) three groups, defined by their effect on the segmental structure:

- **Maternal genes** are expressed during oogenesis by the mother. They may act upon or within the maturing oocyte.
- **Segmentation genes** are expressed after fertilization. Mutations in these genes alter the number or polarity of segments. Three groups of segmentation genes act sequentially to define increasingly smaller regions of the embryo.

- **Homeotic genes** control the identity of a segment, but do not affect the number, polarity, or size of segments. Mutations in these genes cause one body part to develop the phenotype of another part.

The genes in each group act successively to define the properties of increasingly more restricted parts of the embryo. The maternal genes define broad regions in the egg; differences in the distribution of maternal gene products control the expression of segmentation genes; and the homeotic genes determine the identities of individual segments (for review see 355; 362).

Reviews

355. Mahowald, A. P. and Hardy, P. A. (1985). *Genetics of Drosophila embryogenesis*. Annu. Rev. Genet. 19, 149-177.

362. Lawrence, P. (1992). . The Making of a Fly.

GRADIENTS, CASCADES, AND SIGNALING PATHWAYS

6.31.3 A gradient must be converted into discrete compartments

Key Terms

A **denticle** is a pigmented, hardened spike of cuticle protruding from the ventral epidermis of a *Drosophila* embryo.

Key Concepts

- During the first 13 division cycles, nuclei divide in a common cytoplasm; cells form only at blastoderm.
- Gradients define the polarity of the egg along both the anterior-posterior and dorsal-ventral axes.
- The gradients consist of RNAs or proteins that are differentially distributed in the common cytoplasm.
- The location of a nucleus in the cytoplasm with regard to the two axes determines the fate of the cells that descend from it.

The basic question of *Drosophila* development is illustrated in **Figure 31.1** in terms of three stages of development: the egg; the larva; and the adult fly.

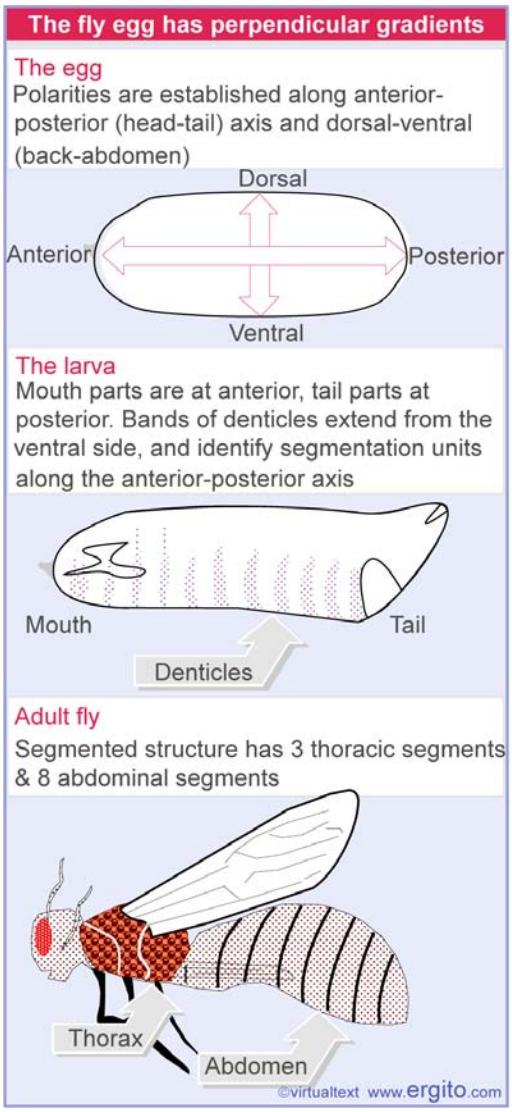

Figure 31.1 Gradients in the egg are translated into segments on the anterior-posterior axis and into specialized structures on the dorsal-ventral axis of the larva, and then into the segmented structure of the adult fly.

At the start of development, gradients are established in the egg along the anterior-posterior and dorsal-ventral axes. The anterior end of the egg becomes the head of the adult; the posterior end becomes the tail. The dorsal side is on top (looking down on a larva); the ventral side is underneath. The gradients consist of molecules (proteins or RNAs) that are differentially distributed in the cytoplasm. The gradient responsible for anterior-posterior development is established soon after fertilization; the dorsal-ventral gradient is established a little later. It is only a modest oversimplification to say that the anterior-posterior systems control positional information along the larva, while the dorsal-ventral system regulates tissue differentiation (that is, the specification of distinct embryonic tissues, including mesoderm, neuroectoderm, and dorsal ectoderm).

Insect development involves two quite different types of structures. The first part of development is concerned with elaborating the larva; then the larva metamorphoses into the fly. This means that the structure of the embryo (the larva) is distinct from the structure of the adult (the fly), in contrast with development of (for example) mammals, where the embryo develops the same body parts that are found in the adult. As the larva develops, it forms some body parts that are exclusively larval (they will not give rise to adult tissues; often they are polyploid), while other body parts are the progenitors that will metamorphose into adult structures (usually they are diploid). In spite of the differences between insect development and vertebrate development, the same general principles appear to govern both processes, and we discover relationships between *Drosophila* regulators and mammalian regulators.

Discrete regions in the embryo correspond to parts of the adult body. They are shown in terms of the superficial organization of the larva in the middle panel of **Figure 31.1**. Bands of **denticles** (small hairs) are found in a particular pattern on the surface (cuticle) of the larva. The cuticular pattern has features determined by both the anterior-posterior axis and the dorsal-ventral axis:

- Along the anterior-posterior axis, the denticles form discrete bands. Each band corresponds to a segment of the adult fly: in fact, the 11 bands of denticles correspond on a 1:1 basis with the 11 segments of the adult.
- Along the dorsal-ventral axis, the denticles that extend from the ventral surface are coarse; those that extend from the dorsal surface are much finer.

Although the cuticle represents only the surface body layer, its structure is diagnostic of the overall organization of the embryo in both axes. Much of the analysis of phenotypes of mutants in *Drosophila* development has therefore been performed in terms of the distortion of the denticle patterns along one axis or the other.

The difference in form between the gradients of the egg and the segments of the adult poses some prime questions. How are gradients established in the egg? And how is a continuous gradient converted into discrete differences that define individual cell types? How can a large number of separate compartments develop from a single gradient?

The nature of the gradients, and their ability to affect the development of a variety of cell types located throughout the embryo, depend upon some idiosyncratic features of *Drosophila* development. The early stages are summarized in **Figure 31.2**.

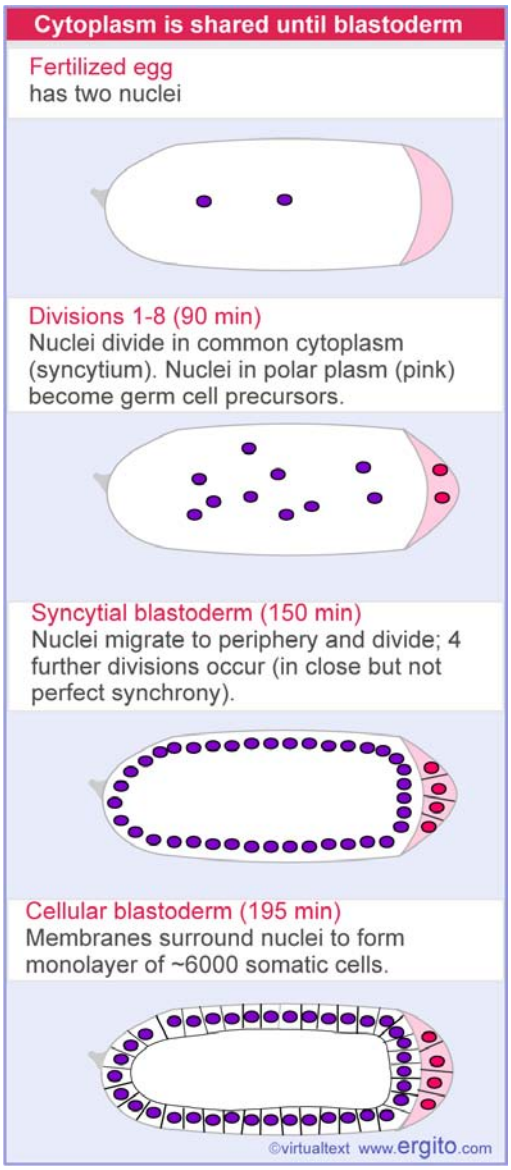

Figure 31.2 The early development of the *Drosophila* egg occurs in a common cytoplasm until the stage of cellular blastoderm.

At fertilization the egg possesses the two parental nuclei and is distinguished at the posterior end by the presence of a region called polar plasm. For the first 9 divisions, the nuclei divide in the common cytoplasm. Material can diffuse in this cytoplasm (although there are probably constraints imposed by cytoskeletal organization). At division 7, some nuclei migrate into the polar plasm, where they become precursors to germ cells. After division 9, nuclei migrate and divide to form a layer at the surface of the egg. Then they divide 4 times, after which membranes surround them to form somatic cells.

Up to the point of cellularization, the nuclei effectively reside in a common cytoplasm. At the stage of the cellular blastoderm, the first discrete compartments become evident, and at this time particular regions of the egg are *determined* to

become particular types of adult structures. (Determination is progressive and gradual; over the next few cell divisions, the fates of individual regions of the egg become increasingly restricted.) At the start of this process, nuclei migrate to the surface to form the monolayer of the blastoderm, but they do not do so in any predefined manner. It is therefore the location in which the nuclei find themselves at this stage that determines what types of cells their descendants will become. *A nucleus determines its position in the embryo by reference to the anterior-posterior and dorsal-ventral gradients, and behaves accordingly.*

GRADIENTS, CASCADES, AND SIGNALING PATHWAYS

6.31.4 Maternal gene products establish gradients in early embryogenesis

Key Terms

In *Drosophila*, a **female sterile** mutation is one in that causes sterility in the female, often because of abnormalities in oogenesis.

A **morphogen** is a factor that induces development of particular cell types in a manner that depends on its concentration.

In *Drosophila*, the **anterior system** is one of the maternal systems that establishes the polarity of the oocyte. The set of genes in the anterior system play a role in the proper formation of the head and the thorax.

In *Drosophila*, the **posterior system** is one of the maternal systems that establishes the polarity of the oocyte. The set of genes in the posterior system play a role in the proper formation of the pole plasm and the abdomen.

In *Drosophila*, the **terminal system** is one of the maternal systems that establishes the polarity of the oocyte. The set of genes in the terminal system play a role in the proper formation of the terminal structures at both ends of the fly.

Key Concepts

- Four signaling pathways are initiated outside the egg and each leads to production of a morphogen in the egg.
- The anterior system is responsible for development of head and thorax.
- The posterior system is responsible for the segments of the abdomen.
- The terminal system is responsible for producing the acron (in the head) and the telson (at the tail).
- The dorsal-ventral system determines development of tissue types (mesoderm, neuroectoderm, ectoderm).

An initial asymmetry is imposed on the *Drosophila* oocyte during oogenesis. **Figure 31.3** illustrates the structure of a follicle in the *Drosophila* ovary. A single progenitor undergoes four successive mitoses to generate 16 interconnected cells. The connections are known as "cytoplasmic bridges" or "ring canals." Individual cells have 2, 3, or 4 such connections. One of the two cells that has 4 connections undergoes meiosis to become the oocyte; the other 15 cells become "nurse cells." Cytoplasmic material, including protein and RNA, passes from the nurse cells to the oocyte; the accumulation of such material accounts for a considerable part of the volume of the egg. The cytoplasmic connections are made at one end of the oocyte, and this end becomes the anterior end of the egg.

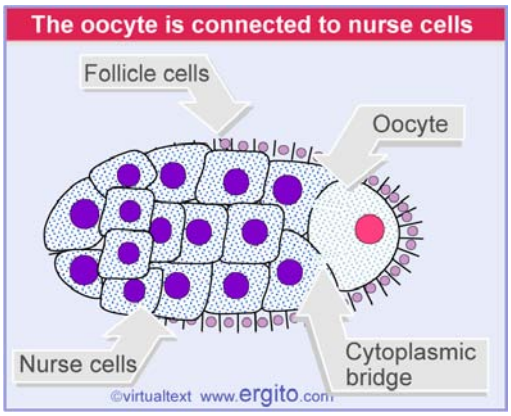

Figure 31.3 A *Drosophila* follicle contains an outer surface of follicle cells that surround nurse cells that are in close contact with the oocyte. Nurse cells are connected by cytoplasmic bridges to each other and to the anterior end of the oocyte. Follicle cells are somatic; nurse cells and the oocyte are germline in origin.

Genes that are expressed within the mother fly are important for early development. These maternal genes are identified by **female sterile** mutations. They do not affect the mother itself, but are required in order to have progeny. Females with such mutations lay eggs that fail to develop into adults; the embryos can be recognized by defects in the cuticular pattern, and they die during development.

The common feature in all maternal genes is that they are expressed prior to fertilization (although their products may act either at the time of expression or be stored for later use). The maternal genes are divided into two classes, depending on their site of expression. Genes that are expressed in somatic cells of the mother that affect egg development are called *maternal somatic genes*. For example, they may act in the follicle cells. Genes that are expressed within the germline are called *maternal germline genes*. These genes may act either in the nurse cell or the oocyte. Some genes act at both stages.

Four groups of genes concerned with the development of particular regions of the embryo can be identified by mutations in maternal genes. The genes in each group can be organized into a pathway that reflects their order of action, by conventional genetic tests (such as comparing the properties of double mutants with the individual mutants) or by biochemical assays (showing which mutants contain components that can bypass the stages that are blocked in other mutants; for review see 355; 362).

The components of these pathways are summarized in **Figure 31.4**, which shows that there is a common principle to their operation. *Each pathway is initiated by localized events outside the egg; this results in the localization of a signal within the egg.* This signal takes the form of a protein with an asymmetric distribution; this is called a **morphogen**. Formally, we may define a morphogen as a protein whose local concentration (or activity) causes the surrounding region to take up a particular structure or fate. In each of these systems, the morphogen either is a transcriptional regulator or leads to the activation of a transcription factor in the localized region. Three systems are concerned with the anterior-posterior axis, and one with the dorsal-ventral axis:

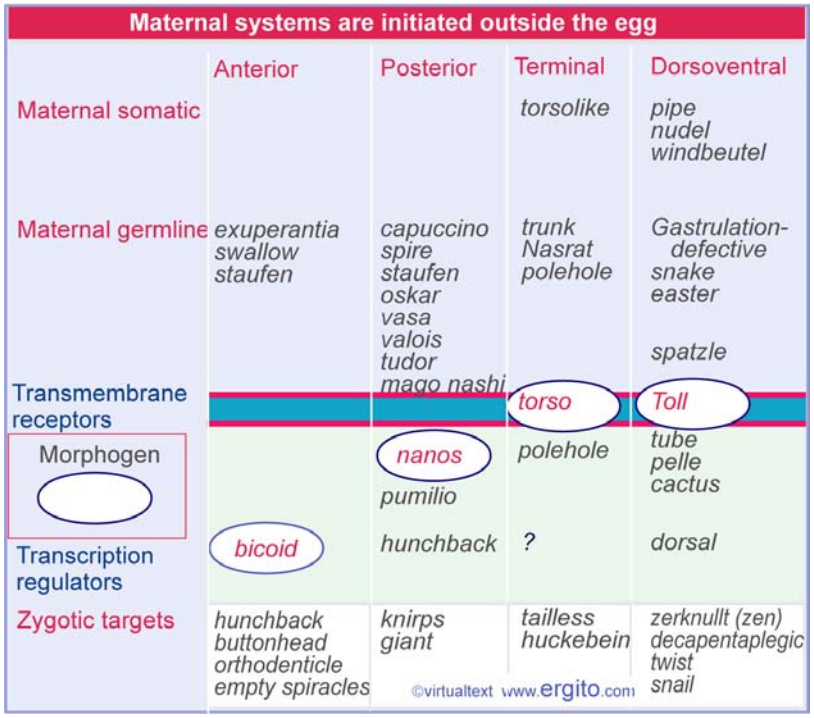

Figure 31.4 Each of the four maternal systems that functions in the egg is initiated outside the egg. The pathway is carried into the egg, where each pathway has a localized product that is the morphogen. This may be a receptor or a regulator of gene expression. The final component is a transcription factor, which acts on zygotic targets that are responsible for the next stage of development.

- The **anterior system** is responsible for development of the head and thorax. The maternal germline products are required to localize the *bicoid* product at the anterior end of the egg. In fact, *bicoid* mRNA is transcribed in nurse cells and transported into the oocyte. Bicoid protein is the morphogen: it functions as a transcriptional regulator, and controls expression of the gene *hunchback* (and probably also other segmentation and homeotic genes).

- The **posterior system** is responsible for the segments of the abdomen. The nature of the initial asymmetric event is not clear. A large number of products act to cause the localization of the product of *nanos*, which is the morphogen. This leads to localized repression of expression of *hunchback* (via control of translation of the mRNA).

- The **terminal system** is responsible for development of the specialized structures at the unsegmented ends of the egg (the acron at the head, and the telson at the tail). As indicated by the dependence on maternal somatic genes, the initial events that create asymmetry occur in the follicle cells. They lead to localized activation of the transmembrane receptor coded by *torso;* the end product of the pathway has yet to be identified.

- The fourth system is responsible for dorsal-ventral development. The pathway is initiated by a signal from a follicle cell on the ventral side of the egg. It is transmitted through the transmembrane receptor coded by *Toll*. This leads to a gradient of activation of the transcription factor produced by *dorsal* (by

controlling its localization within the cell).

About 30 maternal genes involved in pattern formation have been identified. All of the components of the four pathways are maternal, so we see that the systems for establishing the initial pattern formation all depend on events that occur prior to fertilization. The two body axes are established independently. Mutations that affect polarity cause posterior regions to develop as anterior structures, or ventral regions to develop in dorsal form. On the anterior-posterior axis, the anterior and posterior systems provide opposing gradients, with sources at the anterior and posterior ends of the embryo, respectively, that control development of the segments of the body. Defects in either system affect the body segments. The terminal and dorsal-ventral systems operate independently of the other systems (for review see 359).

Reviews

355. Mahowald, A. P. and Hardy, P. A. (1985). *Genetics of Drosophila embryogenesis*. Annu. Rev. Genet. 19, 149-177.

359. Ingham, P. W. (1988). *The molecular genetics of embryonic pattern formation in Drosophila*. Nature 335, 25-34.

362. Lawrence, P. (1992). . The Making of a Fly.

GRADIENTS, CASCADES, AND SIGNALING PATHWAYS

6.31.5 Anterior development uses localized gene regulators

Key Terms

Ectopic refers to something being out of place.

A gene or protein that plays an **instructive** role in development is one that gives a signal telling the cell what to do.

Key Concepts

- The anterior system localizes bicoid mRNA at the anterior end of the egg, generating a gradient of protein that extends along the anterior 40% of the egg.
- The concentration of bicoid protein determines the types of anterior (head) structures that are produced in each region.
- This system is instructive because it is required for development of the head structures.

Establishing asymmetry in an egg requires that some components – either RNAs or proteins – are *localized* instead of being diffused evenly through the cytosol. In anterior-posterior development in *Drosophila*, certain mRNAs are localized at the anterior or posterior end. **Figure 31.5** shows that when they are translated, their protein products diffuse away from the ends of the egg, generating a gradient along the anterior-posterior axis.

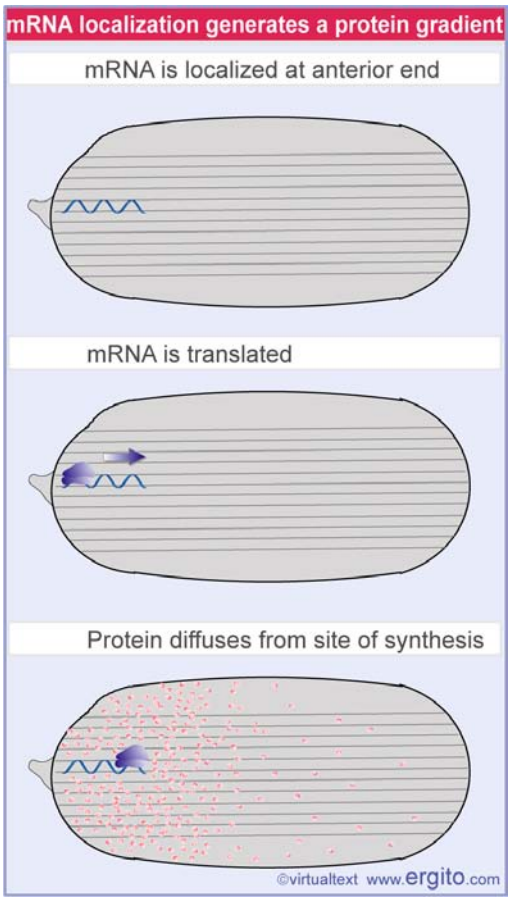

Figure 31.5 Translation of a localized mRNA generates a gradient of protein as the products diffuses away from the site of synthesis.

The existence of localized concentrations of materials needed for development can be tested by the rescue protocol summarized in **Figure 31.6**. Material is removed from a wild-type embryo and injected into the embryo of a mutant that is defective in early development. If the mutant embryo develops normally, we may conclude that the mutation causes a deficiency of material that is present in the wild-type embryo. This allows us to distinguish components that are necessary for morphogenesis, or that are upstream in the pathway, from the morphogen itself – only the morphogen has the property of localized rescue.

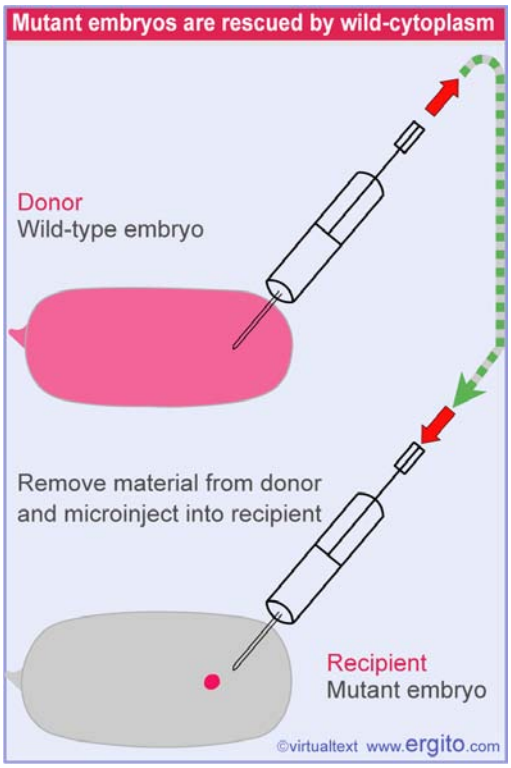

Figure 31.6 Mutant embryos that cannot develop can be rescued by injecting cytoplasm taken from a wild-type embryo. The donor can be tested for time of appearance and location of the rescuing activity; the recipient can be tested for time at which it is susceptible to rescue and the effects of injecting material at different locations.

The rescue technique identifies bicoid as the morphogen required for anterior development. *bicoid* mutants do not develop heads; but the defect can be remedied by injecting mutant eggs with cytoplasm taken from the anterior tip of a wild-type embryo. Indeed, anterior structures develop elsewhere in the mutant embryo if wild-type anterior cytoplasm is injected! (This is called **ectopic** expression.) The extent of the rescue depends on the amount of wild-type cytoplasm injected. And the efficacy of the donor cytoplasm depends on the number of wild-type *bicoid* genes carried by the donor.

These results suggest that the anterior region of a wild-type embryo contains a concentration of some product that depends on the *bicoid* gene dosage. By purifying the active component in the preparation, it is possible to show that purified *bicoid* mRNA can substitute for the anterior cytoplasm. *This implies that the components on which* bicoid *acts are ubiquitous, and all that is required to trigger formation of anterior structures is an appropriate concentration of* bicoid *product.*

The product of *bicoid* establishes a gradient with its source (and therefore the highest concentration) at the anterior end of the embryo. The RNA is localized at the anterior tip of the embryo, but it is not translated during oogenesis. Translation begins soon after fertilization. The protein then establishes a gradient along the embryo, as indicated in **Figure 31.7** (888). The gradient could be produced by diffusion of the

protein product from the localized source at the anterior tip. The gradient is established by division 7, and remains stable until after the blastoderm stage.

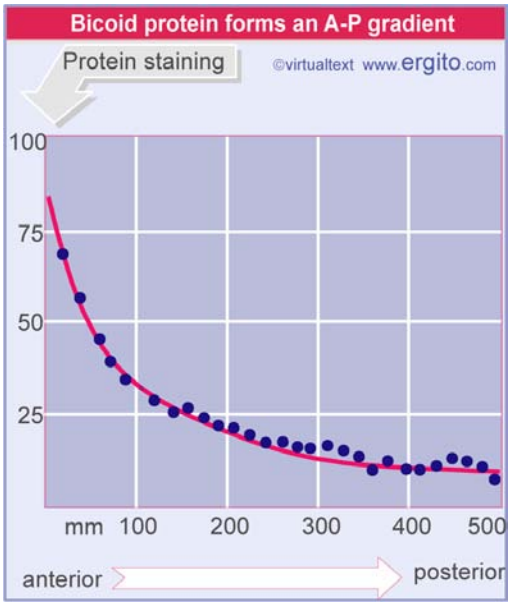

Figure 31.7 Bicoid protein forms a gradient during *D. melanogaster* development that extends for ~200 mm along the egg of 500 mm.

What is the consequence of establishing the bicoid gradient? The gradient can be increased or decreased by changing the number of functional gene copies in the mother. The concentration of bicoid protein is correlated with the development of anterior structures. Weakening the gradient causes anterior segments to develop more posterior-like characteristics; strengthening the gradient causes anterior-like structures to extend farther along the embryo. So the bicoid protein behaves as a morphogen that determines anterior-posterior position in the embryo in a concentration-dependent manner.

The fate of cells in the anterior part of the embryo is determined by the concentration of bicoid protein. The *bicoid* product is a sequence-specific DNA-binding protein that regulates transcription by binding to the promoters of its target genes. The immediate effect of *bicoid* is exercised on other genes that in turn regulate the development of yet further genes. A major target for *bicoid* is the gene *hunchback*. Transcription of *hunchback* is turned on by *bicoid* in a dose-dependent manner, that is, *hunchback* is activated above a certain threshold of bicoid protein. The effect of bicoid on hunchback is to produce a band of expression that occupies the anterior part of the embryo (see **Figure 31.21**; for review see 365).

The relationship between *bicoid* and *hunchback* establishes the principle that a gradient can provide a spatial on-off switch that affects gene expression. In this way, quantitative differences in the amount of the morphogen (bicoid protein) are transformed into qualitatively different states (cell structures) during embryonic development. *bicoid* plays an **instructive** role in anterior development, since it is a positive regulator that is *needed for expression* of genes that in turn determine the synthesis of anterior structures (for review see 3699).

Reviews

365. McGinnis, W. and Krumlauf, R. (1992). *Homeobox genes and axial patterning*. Cell 68, 283-302.

3699. Lawrence, P. A. and Struhl, G. (1996). *Morphogens, compartments, and pattern: lessons from Drosophila?* Cell 85, 951-961.

References

888. Driever, W. and Nusslein-Volhard, C. (1988). *The bicoid protein determines position in the Drosophila embryo in a concentration dependent manner*. Cell 54, 95-104.

GRADIENTS, CASCADES, AND SIGNALING PATHWAYS

6.31.6 Posterior development uses another localized regulator

Key Terms

A protein that plays a **permissive** role in development is one that sets up a situation where a certain activity can occur, but does not cause the occurrence itself.

Key Concepts

- The posterior system localizes nanos mRNA at the posterior end of the egg, generating a gradient of nanos protein that extends along the abdominal region.
- This system is permissive because its function is to repress genes whose products would interfere with posterior development.

Posterior development depends on the expression of a large group of genes. Embryos produced by females who are mutant for any one of these genes develop normal head and thoracic segments, but lack the entire abdomen. Some of these genes are concerned with exporting material from the nurse cells to the egg; others are required to transport or to localize the material within the egg.

The posterior pathway functions by a series of events in which one product is responsible for localizing the next. **Figure 31.8** correlates the order of genes in the genetic pathway with the activities of their products in the embryo. The functions *spir* and *capu* are needed for Staufen protein to be localized at the pole. Staufen protein in turn localizes *oskar* RNA; possibly a complex of Staufen protein and *oskar* RNA is assembled. These functions are needed to localize Vasa, which is an RNA-binding protein. Its specificity and targets are not known.

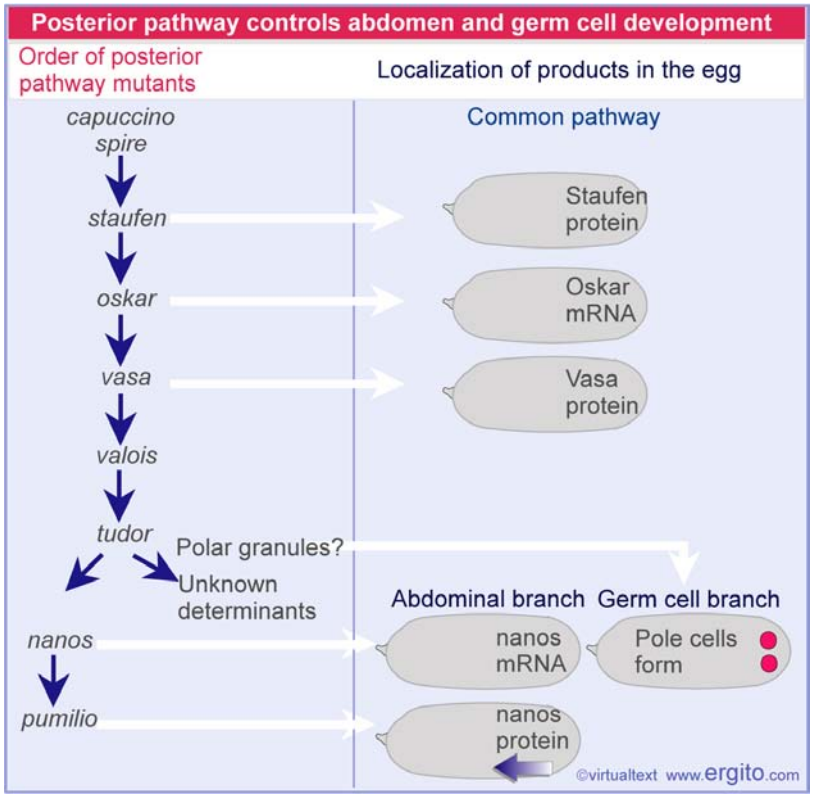

Figure 31.8 The posterior pathway has two branches, responsible for abdominal development and germ cell formation.

If *oskar* is over-expressed or mislocalized in the embryo, it induces germ cell formation at ectopic sites. It requires only the products of *vasa* and *tudor*. This implies that all of the activities that precede *oskar* in the pathway are needed only to localize *oskar* RNA. The ability both to form pole cells and to induce abdominal structures is possessed by *oskar*, in conjunction with *vasa* and *tudor* (and of course any components that are ubiquitously expressed in the egg). One effect of *oskar* function is to localize Vasa protein at the posterior end. The functions of *valois* and *tudor* are not known, but it is possible that *valois* is off the main pathway.

Two types of pattern-determining event occur at the posterior pole, and the pathway branches at *tudor*. The polar plasm contains two morphogens: the posterior determinant (nanos) controls abdominal development; and another signal controls formation of the pole cells, which will give rise to the germline (see **Figure 31.2**). All of the posterior genes except *nanos* and *pumilio* are required for both processes, that is, they are defective in both abdominal development and pole cell formation. *nanos* and *pumilio* identify the abdominal branch. We do not know whether there are additional functions representing a separate branch for germ cell formation, or whether the pathway up to *tudor* is by itself sufficient (901).

The posterior system resembles the anterior system in the basic nature of the morphogenetic event: a maternal mRNA is localized at the posterior pole. This is the product of *nanos*, and provides the morphogen. There are two important differences between the systems. Localization is more complex than in the case of the anterior system, because posterior determinants that originate in the nurse cells must be

transported the full length of the oocyte to the far pole. And nanos protein acts to *prevent* translation of a transcription factor (hunchback). Its role is said to be **permissive**, since it functions to repress genes whose products would interfere with posterior development.

How do we know that *nanos* is the morphogen at the end of the pathway? Rescue experiments (along the lines shown previously in **Figure 31.6**) with the mutants in the posterior group showed that in all but one case the cytoplasm of the nurse cell contained the posterior determinant (although it was absent from the posterior end of the oocyte itself). This indicates that these mutants all act in some subsidiary role, most probably concerned with transporting or localizing the morphogen in the egg. The exception was *nanos*, whose mutants did not contain any posterior-rescuing activity. Purified *nanos* RNA can rescue mutants in any of the other posterior genes, indicating that it is the last, or most downstream, component in the pathway. Indeed, injection of *nanos* RNA into ectopic locations in embryos can induce the formation of abdominal structures, showing that it provides the morphogen.

The upper part of **Figure 31.9** shows the localization of *nanos* mRNA at the posterior end of an early embryo. But the localization poses a dilemma: *nanos* activity is required for development of abdominal segments, that is, for structures occupying approximately the posterior half of the embryo. How does *nanos* RNA at the pole control abdominal development? The lower part of **Figure 31.9** shows that nanos protein diffuses from the site of translation to form a gradient that extends along the abdominal region.

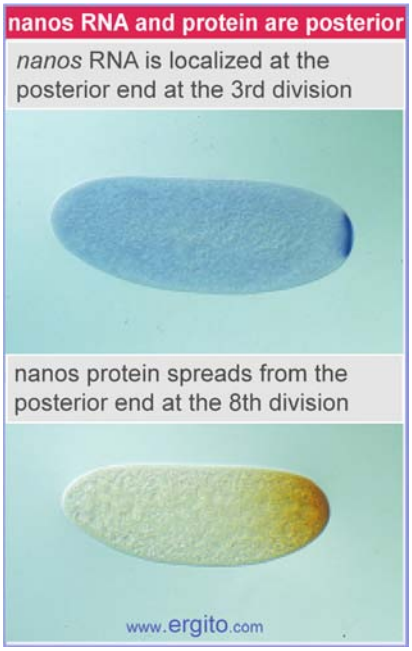

Figure 31.9 nanos products are localized at the posterior end of a *Drosophila* embryo. The upper photograph shows the tightly localized RNA in the very early embryo (at the time of the 3rd nuclear division). The lower photograph shows the spreading of nanos protein at the 8th nuclear division. Photographs kindly provided by Ruth Lehmann.

Both *bicoid* and *nanos* act on the expression of the *hunchback* gene. *hunchback* codes for a repressor of transcription: its presence is needed for formation of anterior structures (in the region of the thorax), and its absence is required for development of posterior structures. It has a complex pattern of expression. It is transcribed during oogenesis to give an mRNA that is uniformly distributed in the egg. After fertilization, the *hunchback* pattern is changed in two ways. The bicoid gradient activates synthesis of *hunchback* RNA in the anterior region. And nanos prevents translation of *hunchback* mRNA in the posterior region; a result of this inhibition is that the mRNA is degraded.

The anterior and posterior systems together therefore enhance hunchback levels in the anterior half of the egg, and remove it from the posterior half. The significance of this distribution lies with the genes that hunchback regulates. It represses the genes *knirps* and (probably) *giant*, which are needed to form abdominal structures. So *the basic role of* hunchback *is to repress formation of abdominal structures by preventing the expression of* knirps *and* giant *in more anterior regions* (894).

References

894. Struhl, G., Johnston, P., and Lawrence, P. A. (1992). *Control of Drosophila body pattern by the hunchback morphogen gradient*. Cell 69, 237-249.

901. Sprenger, F. and Nusslein-Volhard, C. (1992). *Cellular terminal regions of the Drosophila egg*. Cell 71, 987-1001.

GRADIENTS, CASCADES, AND SIGNALING PATHWAYS

6.31.7 How are mRNAs and proteins transported and localized?

Key Concepts

- The mRNAs that establish the anterior and posterior systems are transcribed in nurse cells and transported through cytoplasmic bridges into oocytes.
- *bicoid* mRNA is localized close to the point of entry, but *oskar* and *nanos* mRNAs are transported the length of the oocyte to the posterior end.
- Movement is accomplished by a motor attached to microtubules.

Anterior and posterior development both depend on the localization of an mRNA at one end of the egg. How does the mRNA reach the appropriate location and what is responsible for maintaining it there? Similar processes are involved for the anterior-posterior axis and for the dorsal-ventral axis. On the antero-posterior axis, *bicoid* and *oskar* mRNAs are localized at opposite ends of the egg. On the dorsal-ventral axis, *gurken* mRNA is initially localized at the posterior end and then becomes localized on the dorsal side of the anterior end. The principle is that the sites of transcription are distinct from the sites where the mRNAs are localized and translated, and an active transport process is required to localize the mRNAs.

bicoid, *oskar*, and *nanos* all are transcribed in nurse cells. **Figure 31.10** shows that mRNA is transported through the cytoplasmic bridges into the oocyte. Within the oocyte, *bicoid* mRNA then remains at the anterior end, but *oskar* mRNA is transported the length of the oocyte to the posterior end. The typical means by which an mRNA is transported to a specific location in a cell involves movement along "tracks", which in principle can be either actin filaments or microtubules. This basically means that the mRNA is attached to the tracks by a motor protein that uses hydrolysis of ATP to drive movement (see **Figure 31.10**). In the example of the *Drosophila* egg, microtubules are the tracks used to transport these and other mRNAs (960). In fact, the microtubules form a continuous network that connects the oocyte to the nurse cells through the ring canals (2233).

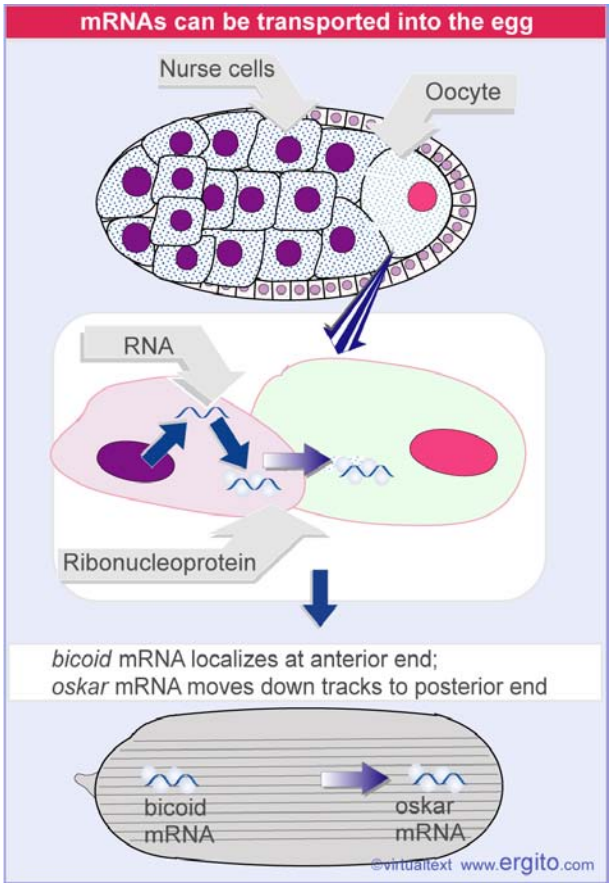

Figure 31.10 Some mRNAs are transported into the *Drosophila* egg as ribonucleoprotein particles. They move to their final sites of localization by association with microtubule.

Genes whose products are needed to transport these mRNAs are identified by mutants in which the mRNAs are not properly localized. The most typical disruption of the pattern is for the mRNAs simply to be distributed throughout the egg. The best characterized of these transport genes are *exuperantia* (*exu*) and *swallow* (*swa*). Exu protein is part of a large ribonucleoprotein complex (961). This complex assembles in the nurse cell, where it uses microtubule tracks to move to the cytoplasmic bridge (962). Then it passes across the bridge into the oocyte in a way that is independent of microtubules. In the oocyte, it attaches to microtubules to move to its location.

The properties of *exu* and *swa* mutants show that there are common components for the transport and localization of different mRNAs. We do not yet know what differences exist between the complexes involved in transporting different mRNAs. However, we assume that there must be a component of each complex that is responsible for targeting it to the right location. By following different mRNAs, it seems that in each case the complex is transported to the anterior end of the oocyte, where it aggregates. Then a decision is made on further localization, and the complex is transported to the appropriate site (for review see 2295; 2304).

Similar events occur at a later stage of development, in the syncytial blastoderm, when some mRNAs become localized on the apical side of the embryo. The same

apparatus seems to be involved as in development of the oocyte. Usually it is responsible for apical localization of the products of several pair rule and segmentation genes. However, if the maternal transcripts of *gurken*, *bicoid*, or *nanos* are injected into the syncytial blastoderm, the apparatus localizes them on the apical side of the embryo (2235). This suggests that the RNAs that are localized at early and at later times have the same set of signals to identify themselves as substrates to the localization apparatus.

mRNAs that are localized in the syncytial blastoderm are found in particles that are connected to microtubules by the motor dynein (2234). The parallels between the transport systems of the oocyte and the blastoderm suggest that dynein also connects the maternal mRNAs to the microtubules. The proteins Egl (egalitarian) and BicD (Bicaudal-D) associate with localizing transcripts and bind to dynein (2235). In their absence, maternal transcripts do not localize properly in the oocyte. This suggests that an Egl/BicD complex may be the means of connecting the mRNA to the motor.

We know that the localization of the *bicoid* RNA to the anterior end of the oocyte depends upon sequences in the 3′ untranslated region (887; 888). This is a common theme, and localization of *oskar* and *nanos* mRNAs is controlled in the same way (963). We assume that the 3′ sequences provide binding sites for specific protein(s) that are involved in localization. Corresponding sequences in each mRNA will provide binding sites for the proteins that target the RNA to the appropriate sites in the oocyte. However, we are still missing the identification of the crucial protein that binds to the localizing sequence in the mRNA.

Localization of RNAs is not sufficient to ensure the pattern of expression. Translation is also controlled. The production of oskar and nanos proteins is controlled by repression of the mRNAs outside of the posterior region. In each case, translation is repressed by a protein that binds to the 3′ region (964). In the case of nanos, there is overlap between the elements required for localization and repression (965; 966; 1434). The consequence of this overlap is to make localization and repression mutually exclusive, so that when a nanos mRNA is localized to the posterior end, it cannot be repressed (1433).

Last updated on January 26, 2004

Reviews

2295. Johnstone, O. and Lasko, P. (2001). *Translational regulation and RNA localization in Drosophila oocytes and embryos.* Annu. Rev. Genet. 35, 365-406.

2304. Palacios, I. M. and Johnston, D. S. (2001). *Getting the message across: the intracellular localization of mRNAs in higher eukaryotes.* Annu. Rev. Cell Dev. Biol. 17, 569-614.

References

887. Driever, W. and Nusslein-Volhard, C. (1988). *A gradient of bicoid protein in Drosophila embryos*. Cell 54, 83-93.

888. Driever, W. and Nusslein-Volhard, C. (1988). *The bicoid protein determines position in the Drosophila embryo in a concentration dependent manner*. Cell 54, 95-104.

960. Pokrywka, N. J. and Stephenson, E. C. (1995). *Microtubules are a general component of mRNA localization systems in Drosophila oocytes*. Dev. Biol. 167, 363-370.

961. Wilhelm, J. E. et al. (2000). *Isolation of a ribonucleoprotein complex involved in mRNA localization in Drosophila oocytes.* J. Cell Biol. 148, 427-439.

962. Theurkauf, W. E. and Hazelrigg, T. I. (1998). *In vivo analyses of cytoplasmic transport and cytoskeletal organization during Drosophila oogenesis: characterization of a multi-step anterior localization pathway.* Development 125, 3655-3666.

963. Gavis, E. R., Curtis, D., and Lehmann, R. (1996). *Identification of cis-acting sequences that control nanos localization*. Dev. Biol. 176, 36-50.

964. Kim-Ha, J., Kerr, K., and MacDonald, P. M. (1995). *Translational regulation of oskar mRNA by bruno, an ovarian RNA-binding protein, is essential.* Cell 81, 403-412.

965. Gavis, E. R., Curtis, D., and Lehmann, R. (1996). *A conserved 90 nucleotide element mediates translational repression of nanos RNA.* Development 122, 2791-2800.

966. Dahanukar, A. and Wharton, R. (1966). *The nanos gradient in Drosophila embryos is generated by translation regulation.* Genes Dev. 10, 2610-2620.

1433. Bergsten, S. E. and Gavis, E. R. (1999). *Role for mRNA localization in translational activation but not spatial restriction of nanos RNA.* Development 126, 659-669.

1434. Crucs, S., Chatterjee, S., and Gavis, E. R. (2000). *Overlapping but distinct RNA elements control repression and activation of nanos translation.* Mol. Cell 5, 457-467.

2233. Theurkauf, W. E., Alberts, B. M., Jan, Y. N., and Jongens, T. A. (1993). *A central role for microtubules in the differentiation of Drosophila oocytes.* Development 118, 1169-1180.

2234. Wilkie, G. S. and Davis, I. (2001). *Drosophila wingless and pair-rule transcripts localize apically by dynein-mediated transport of RNA particles.* Cell 105, 209-219.

2235. Bullock, S. L. and Ish-Horowicz, D. (2001). *Conserved signals and machinery for RNA transport in Drosophila oogenesis and embryogenesis.* Nature 414, 611-616.

GRADIENTS, CASCADES, AND SIGNALING PATHWAYS

6.31.8 How are gradients propagated?

Key Concepts

- Many morphogens form gradients that control differential expression of genes.
- A gradient in an egg or in the early *Drosophila* embryo is propagated by passive diffusion from a localized source.
- A gradient in a cellular tissue may be propagated by passive diffusion in the intercellular spaces or by an active process in which cells transmit the morphogen to other cells.
- A gradient may also be influenced by degradation of morphogen within cells.

The cytosol of an egg forms a single compartment, as indeed does the syncytium of the early *Drosophila* embryo. A protein may form a gradient simply by diffusing away from a localized source (see **Figure 31.5**). In the case of *Drosophila*, such sources are provided by localized mRNAs at either the anterior end (**Figure 31.7**) or posterior end (**Figure 31.9**).

Gradients are also important in development of tissues consisting of cells. We know several cases in which a morphogen forms a gradient, and the cells in that gradient respond differently depending upon the local concentration of the morphogen. The differential response of cells can be seen by placing them in tissue culture on a medium that contains a gradient of morphogen. In the tissue, however, there may be more constraints upon the movement of morphogen. For a gradient to form, material must move in intercellular spaces between the cells or must be transported through the cells. **Figure 31.11** distinguishes some possible models.

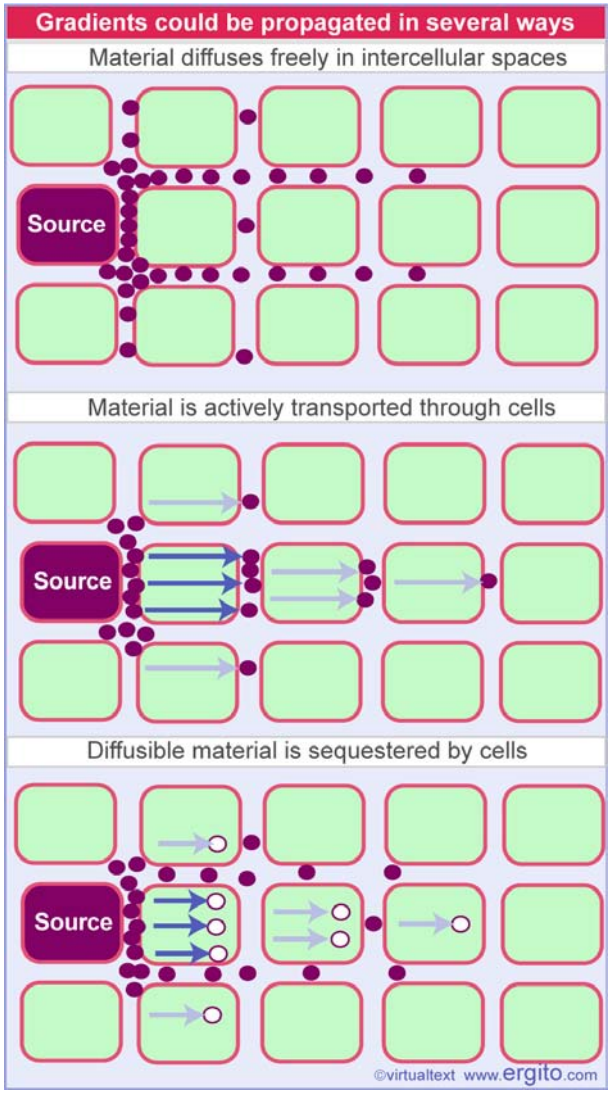

Figure 31.11 A gradient can form by passive diffusion or by active transport or may be affected by removing the diffusing material.

The simplest situation is for passive diffusion. This appears to be responsible for the gradient of activin (a TGF β homologue) that is secreted by cells in the amphibian embryo and induces formation of the mesoderm tissue layer (see *Molecular Biology 6.31.13 TGF β/BMPs are diffusible morphogens*). **Figure 31.12** shows that the critical experiment is to create a tissue (*in vitro*) whose continuity is interrupted by a layer of cells that can neither respond to activin nor synthesize it. If the gradient is not stopped by these cells, the activin must be able either to pass freely through them or (more likely) diffuse past them. This turns out to be the case (2169).

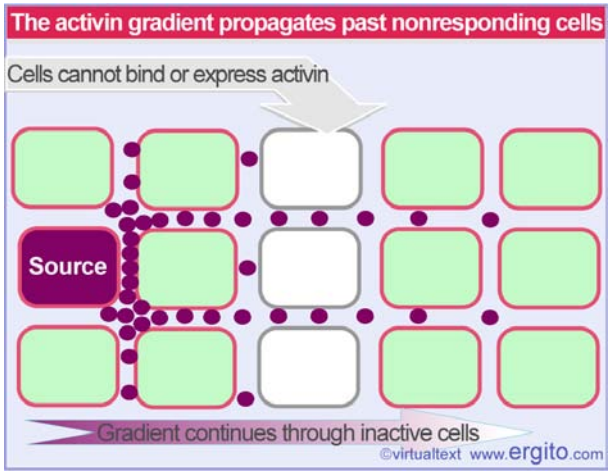

Figure 31.12 Insertion of a layer of cells that cannot interact with activin does not prevent propagation of the gradient.

Formation of the anterior-posterior axis of the *Drosophila* wing is determined by a gradient of the TGF β homologue, Dpp (2172). Gradient formation may involve both of the mechanisms shown in **Figure 31.11** for controlling distribution. The gradient cannot be propagated unless the cells in the issue have an active receptor for TGF β and also the protein dynamin, which is involved in endocytosis (internalization) of Dpp (2171). The basic means of propagating the gradient appears to be transcytosis, in which Dpp is taken up at one face of the cell, transported across the cell, and secreted through the membrane at the other side. Some of the Dpp may be degraded instead of being passed on to the next cell. **Figure 31.13** implies that the shape of the gradient may be influenced by the balance between these two processes.

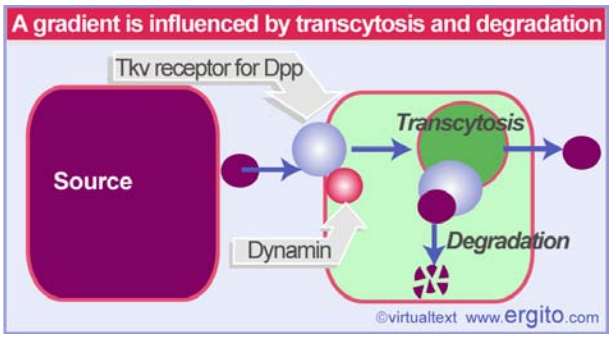

Figure 31.13 A gradient can be propagated by transcytosis, when a morphogen is endocytosed (internalized) at once face of a cell and secreted at the other face. The gradient will be sharpened if some of the morphogen is degraded within the cell.

References

2169. Gurdon, J. B., Harger, P., Mitchell, A., and Lemaire, P. (1994). *Activin signalling and response to a morphogen gradient.* Nature 371, 487-492.

2171. Entchev, E. V., Schwabedissen, A., and Gonzalez-Gaitan, M. (2000). *Gradient formation of the TGF-beta homolog Dpp.* Cell 103, 981-991.

2172. Nellen, D., Burke, R., Struhl, G., and Basler, K. (1996). *Direct and long-range action of a DPP morphogen gradient.* Cell 85, 357-368.

GRADIENTS, CASCADES, AND SIGNALING PATHWAYS

6.31.9 Dorsal-ventral development uses localized receptor-ligand interactions

Key Concepts

- Gurken mRNA is localized on the dorsal side of the oocyte.
- It is translated into a TGF α-like growth factor that interacts with the Torpedo receptor on the adjoining follicle cell.
- The Torpedo receptor triggers a Ras/MAPK pathway that prevents the follicle cell from acquiring a ventral fate.

Dorsal-ventral development displays a complex interplay between the oocyte and follicle cells, involving separate pathways that are required to develop ventral and dorsal structures. The formation of ventral pattern starts with the expression of genes in the oocyte that are needed for proper development of the follicle cells. And then expression of genes in the follicle cells transmits a signal to the oocyte that results in development of ventral structures. Another pathway is responsible for development of dorsal structures in the developing egg. Each of these systems functions by activating a localized ligand-receptor interaction that triggers a signal transduction pathway (886).

The localization of *gurken* mRNA initiates dorsal development, but also plays a role earlier in anterior-posterior patterning. These are key events that define the spatial asymmetry of the egg chamber. (The requirement of *gurken* for both pathways is the only feature that breaches their independence.)

First *gurken* mRNA is localized on the posterior side of the oocyte. This results in a signal that causes adjacent follicle cells to become posterior. The follicle cells signal back to the oocyte in a process that results in the establishment of a polarized network of microtubules. This is necessary for the localization of the maternal transcripts of *bicoid* and *oskar* to opposite poles (see *Molecular Biology 6.31.5 Anterior development uses localized gene regulators*).

Dorsal-ventral polarity is established later when *gurken* mRNA becomes localized on the dorsal side of the oocyte. **Figure 31.14** illustrates the pathway and its consequences. The products of *cornichon* and *brainiac* are needed for proper localization of the *gurken* mRNA or for activation of the protein. Of the group of loci that act earlier, the products of *K10* and *squid* are needed to localize the RNA; and *cappuccino* and *spire* mutants have an array of defects that suggest their products have a general role in organizing the cytoskeleton of the oocyte. Accordingly, *cappuccino* and *spire* are required also for the earlier localization of *gurken* mRNA involved in anterior-posterior patterning.

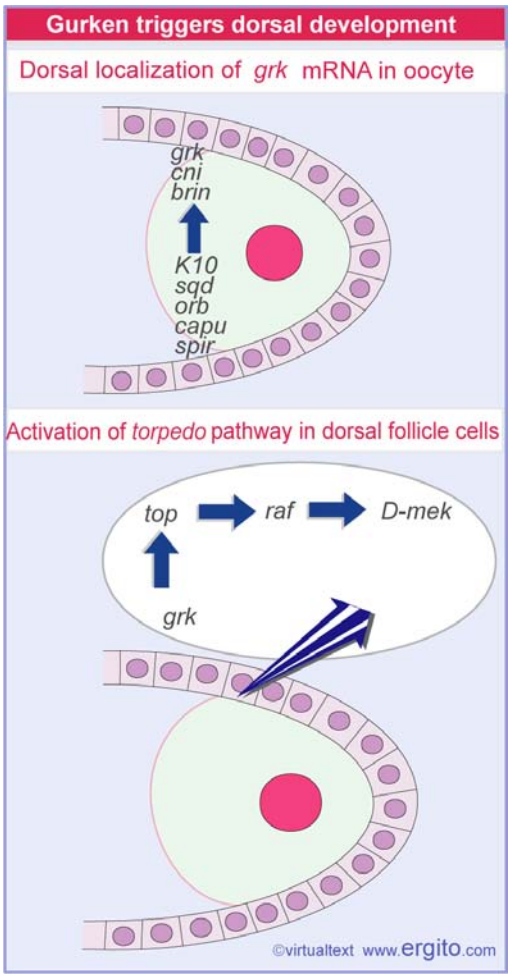

Figure 31.14 Dorsal and ventral identities are first distinguished when *grk* mRNA is localized on the dorsal side of the oocyte. Synthesis of Grk activates the receptor coded by *torpedo*, which triggers a MAPK pathway in the follicle cells.

gurken codes for a protein that resembles the growth factor TGF α (2885; 2886). The next locus in the pathway is *torpedo*, which codes for the *Drosophila* EGF receptor. It is expressed in the follicle cells. So the pathway moves from oocyte to follicle cells when the ligand (Gurken), possibly in a transmembrane form that exposes the extracellular domain on the oocyte, interacts with the receptor (Torpedo) on the plasma membrane of a follicle cell (for review see 2887).

An interesting and general principle emerges from the activation of Torpedo, which is a typical receptor tyrosine kinase. Activation of Torpedo leads to the activation of a Ras signaling pathway, which proceeds through Raf and D-mek (the equivalent of MAPKK), to activate a classic MAP kinase pathway. The ultimate readout of this pathway is not known, but its effect is to prevent activation on the dorsal side of the embryo of the ventral-determining pathway (see *Molecular Biology 6.31.10 Ventral development proceeds through Toll*).

The utilization of this pathway shows that similar pathways may be employed in

different circumstances to produce highly specific effects. The trigger to activate the pathway in the oocyte-follicle cell interaction is the specific localization of Gurken. The consequence is a change in the properties of follicle cells that prevents them from acquiring ventral fates. The basic components of the pathway, however, are the same as those employed in signal transduction of proliferation signals in vertebrate systems. The same pathway is employed again in the specific development of retinal cells in *Drosophila* itself, where another receptor-counter receptor interaction activates the Ras pathway, with specific, but very different effects on cell differentiation. So essentially the same pathway can be employed to interpret an initial signal and produce a response that is predetermined by the cell phenotype.

Last updated on 8-30-2002

Reviews

2887. Wylie, A. A., Murphy, S. K., Orton, T. C., and Jirtle, R. L. (1996). *Intercellular signaling and the polarization of body axes during Drosophila oogenesis.* Genes Dev. 10, 1711-1723.

References

886. Schupbach, T. (1987). *Germ line and soma cooperate during oogenesis to establish the Dorsal-ventral pattern of egg shell and embryo in D. melanogaster.* Cell 49, 699-707.

2885. Price, J. V., Clifford, R. J., and Schupbach, T. (1989). *The maternal ventralizing locus torpedo is allelic to faint little ball, an embryonic lethal, and encodes the Drosophila EGF receptor homolog.* Cell 56, 1085-1092.

2886. Schejter, E. D. and Shilo, B. Z. (1989). *The Drosophila EGF receptor homolog (DER) gene is allelic to faint little ball, a locus essential for embryonic development.* Cell 56, 1093-1104.

GRADIENTS, CASCADES, AND SIGNALING PATHWAYS

6.31.10 Ventral development proceeds through Toll

Key Concepts

- The follicle cell on the ventral side produces an enzyme that modifies a proteoglycan.
- The proteoglycan triggers a series of proteolytic cleavages in the perivitelline space of the oocyte that activate the spatzle ligand.
- spatzle activates the receptor Toll, which is related to IL1 receptor, and triggers a pathway leading to activation of dorsal, which is related to the vertebrate transcription factor NF- κB.

Development of ventral structures requires a group of 11 maternal genes whose products establish the dorsal-ventral axis between the time of fertilization and cellular blastoderm (see **Figure 31.4**). **Figure 31.15** shows that the dorsal-ventral pattern can be viewed from the side by the phenotype of the cuticle, and can be seen in cross-section to represent the formation of different types of tissues. The *dorsal* system is necessary for the development of ventral structures including the mesoderm and neurogenic ectoderm. (The system was named for the effects of mutations [to dorsalize], rather than for the role of the gene products [to ventralize].) Mutants in any genes of the *dorsal* group lack ventral structures, and have dorsal structures on the ventral side, as indicated in the figure. But injecting wild-type cytoplasm into mutant embryos rescues the defect and allows ventral structures to develop (for review see 367).

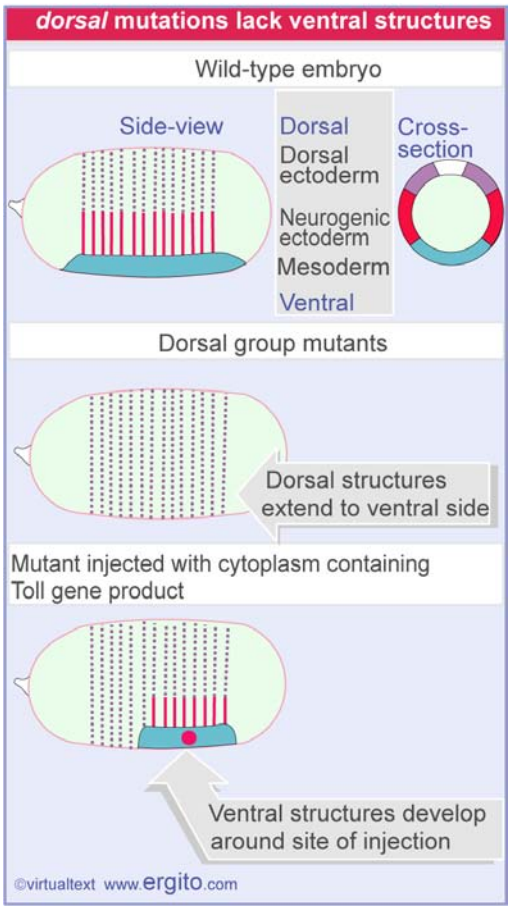

Figure 31.15 Wild-type *Drosophila* embryos have distinct dorsal and ventral structures. Mutations in genes of the *dorsal* group prevent the appearance of ventral structures, and the ventral side of the embryo is dorsalized. Ventral structures can be restored by injecting cytoplasm containing the Toll gene product.

The ventral-determining pathway also begins in the follicle cell and ends in the oocyte. The pathway is summarized in **Figure 31.16**. The initial steps are not well defined, and require the expression of three loci in the follicle cells on the ventral side (2888). These loci function before fertilization, but the egg does not receive the signal until after fertilization.

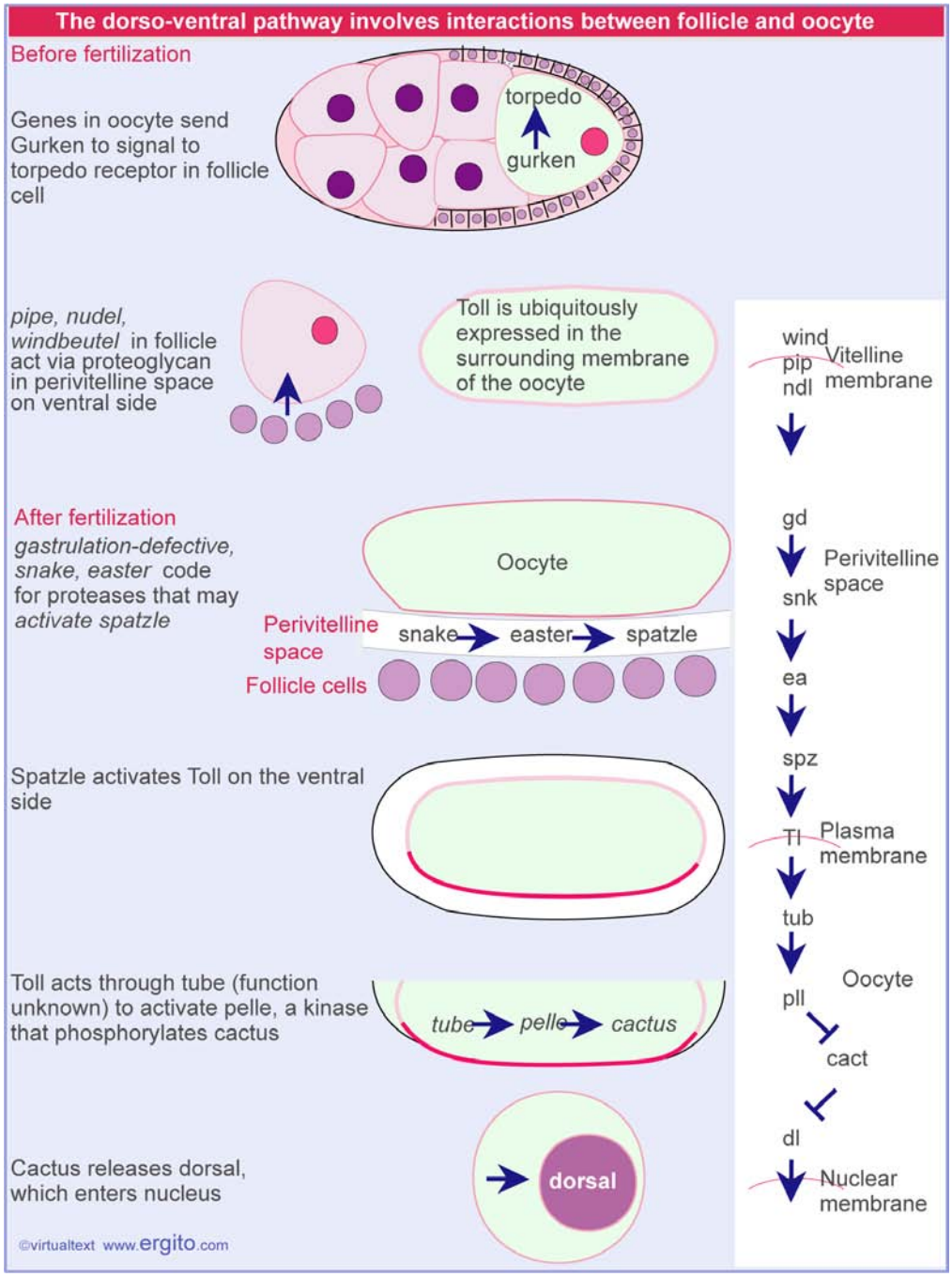

Figure 31.16 The dorsal-ventral pathway is summarized on the right and shown in detail on the left. It involves interactions between follicle cells and the oocyte. The pathway moves into the oocyte when spatzle binds to Toll and activates the morphogen. The pathway is completed by transporting the transcription factor dorsal into the nucleus.

The three loci that act in the follicle cells are *nudel* (*ndl*), *windbeutel* (*wind*), and *pipe*. The roles of ndl and wind are not known, but *pipe* plays an interesting and novel role. *pipe* codes for an enzyme whose sequence suggests that it is similar to the enzyme heparan sulfate 2-O-sulfotransferase (HSST) that is involved in the synthesis of a class of proteins called proteoglycans. They are components of the extracellular

matrix. These proteins have covalently attached carbohydrate side-chains called glycosaminoglycans that have a characteristic pattern of attached monosaccharides. HSST is an enzyme that adds sulfate to the 2-O position of certain sugar residues in the monosaccharide.

The *pipe* gene is expressed in follicle cells on the ventral side of the embryo (2883). We assume that, like HSST, it functions within the Golgi apparatus of the follicle cell. Its substrate is not known, but **Figure 31.17** shows that it is probably secreted from the follicle cell into the perivitelline space (the outermost layer of the oocyte). Because the proteoglycan is synthesized on the ventral side, this creates an asymmetry at the surface of the egg.

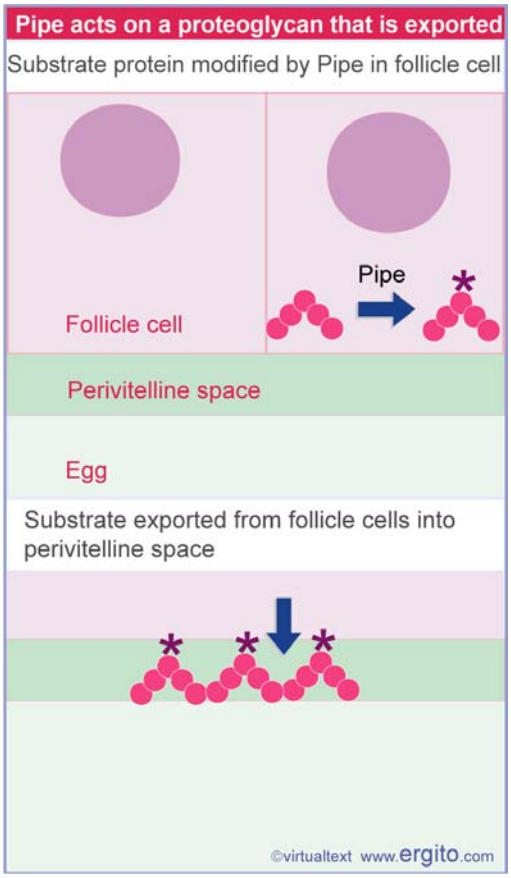

Figure 31.17 Pipe modifies a proteolglycan within the follicle cells. The modified protein is exported to the perivitelline space

The presence of the proteoglycan in some unknown way triggers a series of proteolytic cleavages that occur in the perivitelline space. Several proteases act in succession, ending with the cleavage of the *spatzle* product (2884). spatzle provides a ligand for a receptor coded by the *Toll* gene. Toll is the first component of the pathway that functions in the oocyte.

Rescue experiments identify *Toll* as the crucial gene that conveys the signal into the oocyte. *Toll*$^{-}$ mutants lack any dorsal-ventral gradient, and injection of Toll induces

the formation of dorsal-ventral structures. The other genes of the dorsal group code for products that either regulate or are required for the action of *Toll*, but they do not establish the primary polarity (896).

There is a paradox in the distribution of Toll protein. *Toll* gene product activity is found in all parts of a donor embryo when cytoplasm is extracted and tested by injection. Yet it induces ventral structures only in the appropriate location in normal development. An initial general distribution of *Toll* gene product must therefore in some way be converted into a concentration of active product by local events.

Toll is a transmembrane protein homologous to the vertebrate interleukin-1 (IL1) receptor. It is located in the plasma membrane of the egg cell, with its ligand-binding domain extending into the perivitelline space (2889). Binding of ligand is sufficient to activate the ventral-determining pathway. The reaction occurs on the ventral side of the perivitelline space. The spatzle ligand either cannot diffuse far from the site where it is generated, or perhaps it binds to Toll very rapidly, with the result that Toll is activated only on the ventral side of the embryo. Loss-of-function mutations in Toll are dorsalized, because the receptor cannot be activated. There are also dominant ($Toll^{D}$) mutations, which confer ventral properties on dorsal regions; these are gain-of-function mutations, which are ventralized because the receptor is constitutively active. Genetic analysis shows that *toll* acts via *tube* and *pelle*. Tube is probably an adaptor protein that recruits the kinase pelle to the activated receptor. The target for the pelle kinase is not proven, but its activation leads to phosphorylation of the product of *cactus*, which is the final regulator of the transcription factor coded by *dorsal*.

Last updated on 8-30-2002

Reviews

367. Morisato, D. and Anderson, K. V. (1995). *Signaling pathways that establish the dorsal-ventral pattern of the Drosophila embryo.* Annu. Rev. Genet. 29, 371-399.

References

896. Anderson, K. V., Bokla, L., and Nusslein-Volhard, C. (1985). *Establishment of dorsal-ventral polarity in the Drosophila embryo: the induction of polarity by the Toll gene product.* Cell 42, 791-798.

2883. Sen, J., Goltz, J. S., Stevens, L., and Stein, D. (1998). *Spatially restricted expression of pipe in the Drosophila egg chamber defines embryonic dorsal-ventral polarity.* Cell 95, 471-481.

2884. Morisato, D. and Anderson, K. V. (1994). *The spatzle gene encodes a component of the extracellular signaling pathway establishing the dorsal-ventral pattern of the Drosophila embryo.* Cell 76, 677-688.

2888. Stein, D., Roth, S., Vogelsang, E., Nusslein-Volhard, C., and Vogelsang, C. (1991). *The polarity of the dorsoventral axis in the Drosophila embryo is defined by an extracellular signal.* Cell 65, 725-735.

2889. Hashimoto, C., Hudson, K. L., and Anderson, K. V. (1988). *The Toll gene of Drosophila, required for dorsal-ventral embryonic polarity, appears to encode a transmembrane protein.* Cell 52, 269-279.

GRADIENTS, CASCADES, AND SIGNALING PATHWAYS

6.31.11 Dorsal protein forms a gradient of nuclear localization

Key Concepts

- The activation of dorsal is achieved by releasing it in the cytoplasm so that it can enter the nucleus.
- A gradient of dorsal with regard to nuclear localization is established along the ventral to dorsal axis.

Figure 31.18 shows the parallels between the *toll* signaling pathway in flies and the IL1 vertebrate pathway (where the biochemistry is well characterized). Activation of the receptor causes a complex to assemble that includes adaptor proteins (several in vertebrates), which bind a kinase. Activation of the vertebrate kinase (IRAK) in turn activates the kinase NIK, which phosphorylates I- κB. It is not clear whether the fly kinase (pelle) acts directly on cactus (the equivalent of I κ-B) or through an intermediate.

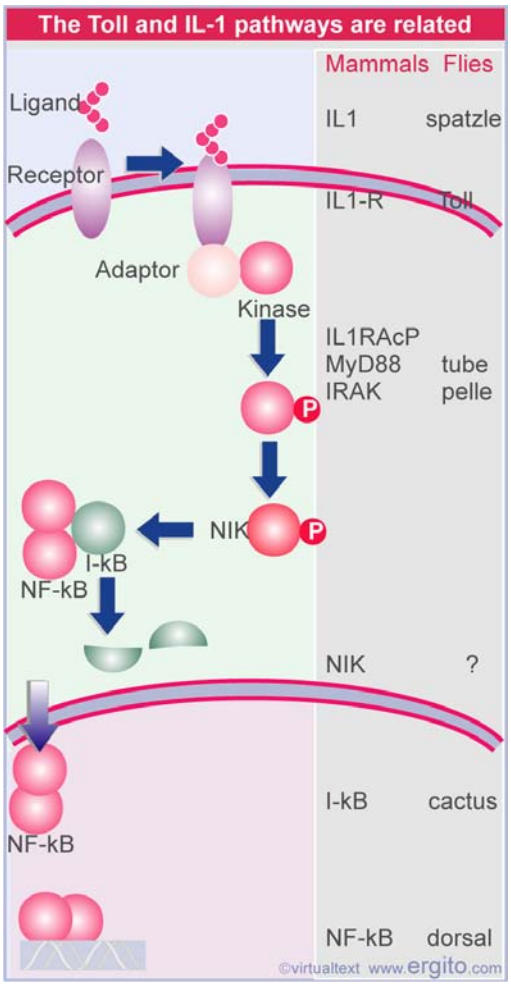

Figure 31.18 Activation of IL1 receptor triggers formation of a complex containing adaptor(s) and a kinase. The IRAK kinase activates NIK, which phosphorylates I- κ B. This triggers degradation of I- κ B, releasing NF- κ B, which translocates to the nucleus to activate transcription.

At all events, dorsal and cactus form an interacting pair of proteins that are related to the transcription factor NF- κB and its regulator I κB (3662; 3663; 3692). NF- κB consists of two subunits (related in sequence) which are bound by I κB in the cytoplasm. When I κB is phosphorylated, it releases NF- κB, which then moves into the nucleus, where it functions as a transcription factor of genes whose promoters have the κB sequence motif. (An example of the pathway is illustrated in **Figure 22.12**.) Cactus regulates dorsal in the same way that I κB regulates NF- κB (3693). A cactus-dorsal complex is inert in the cytoplasm, but when cactus is phosphorylated, it releases dorsal protein, which enters the nucleus. The pathway is therefore conserved from receptor to effector, since activation of interleukin-1 receptor has as a principal effect the activation of NF- κ B, and activation of Toll leads to activation of dorsal. A related pathway, triggered by a Toll-like receptor (TLR), is found in the system of innate immunity that is conserved from flies to mammals (see *Molecular Biology 5.25.21 Innate immunity utilizes conserved signaling pathways*).

As a result of the activation of Toll, a gradient of dorsal protein in the nucleus is established, from ventral to dorsal side of the embryo. On the ventral side, dorsal protein is released to the nucleus, but on the dorsal side of the embryo it remains in the cytoplasm. A steep gradient is established at the stage of syncytial blastoderm, and becomes sharper during the transition to cellular blastoderm. The proportion of dorsal protein that is in the nucleus correlates with the ventral phenotype that will be displayed by this region. An example of a gradient visualized by staining with antibody against dorsal protein is shown in **Figure 31.19**. The total amount of dorsal protein in the embryo does not change: the gradient is established solely by a redistribution of the protein between nucleus and cytoplasm (899; 3694; for review see 368).

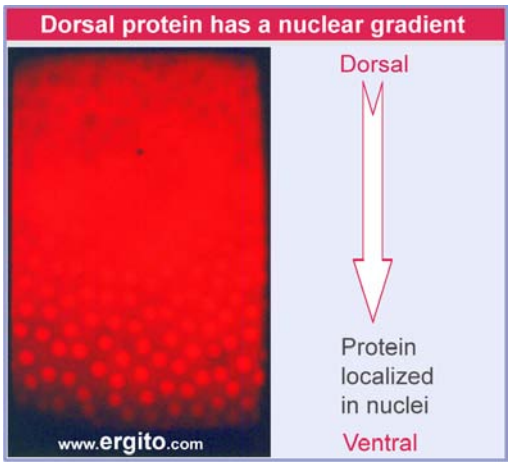

Figure 31.19 Dorsal protein forms a gradient of nuclear localization from ventral to dorsal side of the embryo. On the ventral side (lower) the protein identifies bright nuclei; on the dorsal side (upper) the nuclei lack protein and show as dark holes in the bright cytoplasm. Photograph kindly provided by Michael Levine.

Dorsal both activates and represses gene expression. It activates the genes *twist* and *snail*, which are required for the development of ventral structures. And it represses the genes *dpp* and *zen*, which are required for the development of dorsal structures; as a result, these genes are expressed only in the 40% most dorsal of the embryo (see *Molecular Biology 6.31.13 TGF β/BMPs are diffusible morphogens*).

One of the crucial aspects of dorsal-ventral development is the relationship between the different pathways. This is summarized in **Figure 31.20**. The ability of one system to repress the next is responsible for restricting the localized activities to the appropriate part of the embryo. The initial interaction between gurken and torpedo leads to the repression of spatzle activity on the dorsal side of the embryo. This restricts the activation of dorsal protein to the ventral side of the embryo. Nuclear localization of dorsal protein in turn represses the expression of *dpp*, so that it forms a gradient diffusing from the dorsal side. In this way, ventral structures are formed in the nuclear gradient of dorsal protein, and dorsal structures are formed in the gradient of dpp protein.

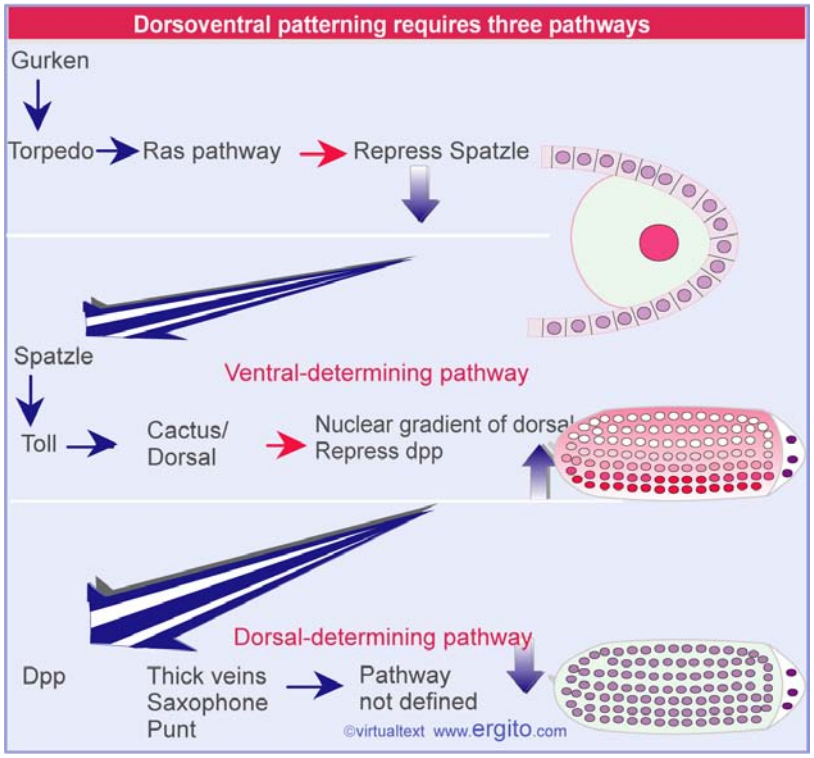

Figure 31.20 Dorsal-ventral patterning requires the successive actions of three localized systems.

The terminal system is initiated in a way that is similar to the dorsal-ventral system. A transmembrane receptor, coded by the *torso* gene, is produced by translation of a maternal RNA after fertilization. The receptor is localized throughout the embryo. It is activated at the poles by local production of an extracellular ligand. Torso protein has a kinase activity, which initiates a cascade that leads to local expression of the *tailless* and *huckebein* RNAs, which code for factors that regulate transcription.

Last updated on 8-30-2002

Reviews

368. Belvin, M. P. and Anderson, K. V. (1996). *A conserved signaling pathway: the Drosophila Toll-Dorsal pathway*. Annu. Rev. Cell Dev. Biol. 12, 393-416.

References

899. Roth, S., Stein, D., and Nusslein-Volhard, C. (1989). *A gradient of nuclear localization of the dorsal protein determines Dorsal-ventral pattern in the Drosophila embryo*. Cell 59, 1189-202.

3662. Kieran, M., Blank, V., Logeat, F., Vandekerckhove, J., Lottspeich, F., Le Bail, O., Urban, M. B., Kourilsky, P., Baeuerle, P. A., and Israel, A. (1990). *The DNA binding subunit of NF-kappa B is identical to factor KBF1 and homologous to the rel oncogene product.* Cell 62, 1007-1018.

3663. Ghosh, S., Gifford, A. M., Riviere, L. R., Tempst, P., Nolan, G. P., and Baltimore, D. (1990). *Cloning of the p50 DNA binding subunit of NF-kappa B: homology to rel and dorsal.* Cell 62, 1019-1029.

3692. Ip, Y. T., Kraut, R., Levine, M., and Rushlow, C. A. (1991). *The dorsal morphogen is a sequence-specific DNA-binding protein that interacts with a long-range repression element in Drosophila.* Cell 64, 439-446.

3693. Kidd, S. (1992). *Characterization of the Drosophila cactus locus and analysis of interactions between cactus and dorsal proteins.* Cell 71, 623-635.

3694. Rushlow, C. A., Han, K., Manley, J. L., and Levine, M. (1989). *The graded distribution of the dorsal morphogen is initiated by selective nuclear transport in Drosophila.* Cell 59, 1165-1177.

GRADIENTS, CASCADES, AND SIGNALING PATHWAYS

6.31.12 Patterning systems have common features

The pattern of regulators at each stage of development for each of the systems is summarized in **Figure 31.21**. Two types of mechanism are used to create the initial asymmetry. For the anterior-posterior axis, an RNA is localized at one end of the egg (bicoid for the anterior system, nanos for the posterior system); localization depends upon the interaction of sequences in the 3′ end of the RNA with maternal proteins. In the case of the dorsal-ventral and terminal systems, a receptor protein is specifically activated in a localized manner, as the result of the limited availability of its ligand. All of these interactions depend on RNAs and/or proteins expressed from maternal genes.

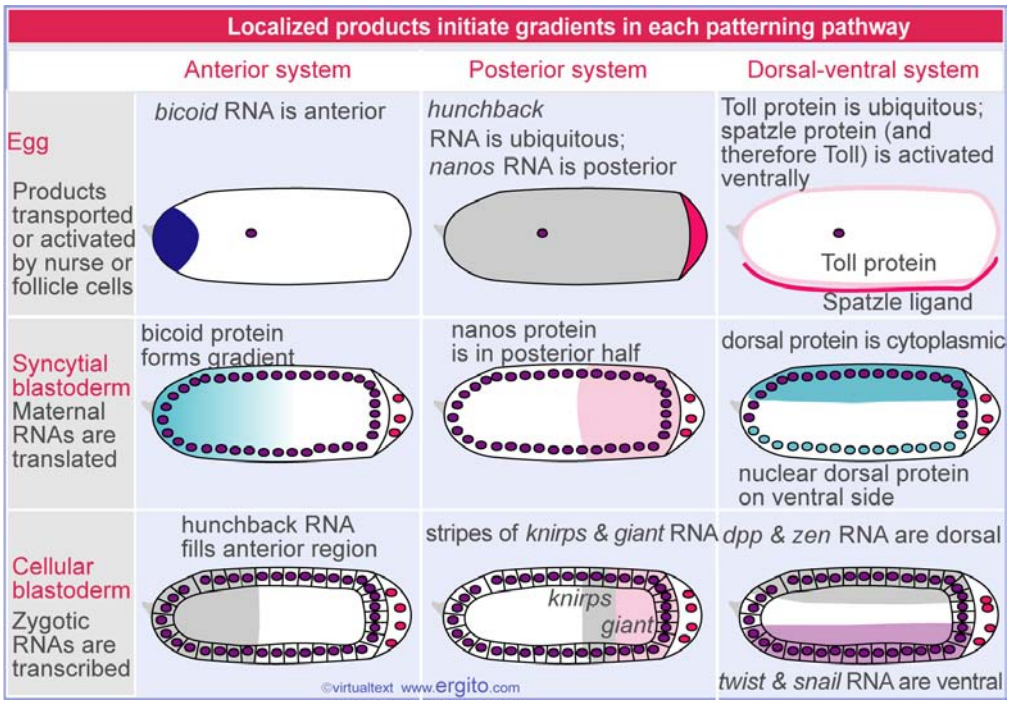

Figure 31.21 In each axis-determining system, localized products in the egg cause other maternal RNAs or proteins to be broadly localized at syncytial blastoderm, and zygotic RNAs are transcribed in bands at cellular blastoderm.

The local event leads to the production of a morphogen, which forms a gradient, either quantitatively (bicoid) or by nucleocytoplasmic distribution (dorsal), or is localized in a broad restricted region (nanos). The extent of the region in which the morphogen is active is ~50% across the egg for each of these systems. The morphogens are translated from maternal RNAs, and development is therefore still dependent on maternal genes up to this stage.

Establishing anterior-posterior and dorsal-ventral gradients is the first step in determining orientation and spatial organization of the embryo. Under the direction of maternal genes, gradients form across the common cytoplasm and influence the

behavior of the nuclei located in it. The next step is the development of discrete regions that will give rise to different body parts. This requires the expression of the zygotic genome, and the loci that now become active are called *zygotic genes.* Genes involved at this stage are identified by segmentation mutants.

The products of the segmentation genes form bands that distinguish individual regions on the anterior-posterior axis. When we consider the results of the anterior and posterior systems together, we see that there are several broad regions (the two regions generated by the anterior system are adjacent to the two regions defined by the posterior system; see **Figure 31.27**). On the dorsal-ventral axis, there are three rather broad bands that define the regions in which the mesoderm, neuroectoderm, and dorsal ectoderm form (proceeding from the ventral to the dorsal side).

Last updated on 8-30-2002

GRADIENTS, CASCADES, AND SIGNALING PATHWAYS

6.31.13 TGF β/BMPs are diffusible morphogens

Key Concepts

- The TGF β/BMP family provides ligands for receptors that activate Smads transcription factors.
- Synthesis of the Dpp member of this family is repressed on the ventral side of the fly embryo.
- It diffuses from a source on the ventral side and induces neural tissues.
- A similar pathway functions in vertebrates but is inverted with regard to the dorsal-ventral axis.

The principle of dorsal-ventral development in flies, amphibians, and mammals is the same. On one side of the animal, neural structures (including the CNS) develop. This is the ventral side in flies, and the dorsal side in vertebrates. On the other side, mesenchymal structures develop. This is the dorsal side in flies, and the ventral side in vertebrates. The important point here is that the same relative development is seen from one side of the animal to the other, but its absolute direction is reversed between flies and vertebrates. This must mean that the dorsal-ventral axis was inverted at some point during evolution, causing the CNS to be displaced from the ventral side to the dorsal side. This idea is supported by the fact that the same signaling pathway is initiated on the dorsal side of flies and on the ventral side of vertebrate embryos (for review see 369).

Mesenchymal (non-neural) structures are determined by diffusible factors in the TGF β/BMP family. These factors are small polypeptide ligands for receptors that activate the Smads transcription factors (see **Figure 28.46**). Formation of neural structures requires counteracting activities that also diffuse from a center; they prevent the TGF β/BMP ligands from activating the target receptors. (The names reflect the histories of their discoveries as transforming growth factor β and bone morphogenetic proteins; but in fact the most important role of these polypeptides is as morphogens in development.)

The involvement of this pathway in development was first described in *Drosophila*, where the product of *dpp* is a member of the TGF β growth factor family. The receptors typically are heterodimers that form transmembrane proteins with serine/threonine kinase activity. The heterodimer consists of a type I component and a type II component. In the *dpp* pathway, there are two type I members (coded by *thick veins* and *saxophone*) and a single type II member (coded by *punt*). Mutations in *tkv* and *punt* have the same phenotype as mutations in *dpp*, suggesting that the tkv/punt heterodimer is the principal receptor (3695; 3696).

The activated receptor phosphorylates the product *mad*. This is the founding member of the Smad family. The typical pattern of activation in mammalian cells is for the

regulated Smad to associate with a general partner to form a heterodimer that is imported into the nucleus, where it activates transcription (see *Molecular Biology 6.28.21 TGF β signals through Smads*).

Because the gene is repressed on the ventral side, Dpp protein is secreted from cells only across the dorsal side of the embryo, as depicted in **Figure 31.22**. So Dpp is in effect the morphogen that induces synthesis of dorsal structures (902). Several loci influence the production of Dpp, largely by post-translational mechanisms. The net result is to increase Dpp activity on the dorsal side, and to repress it on the ventral side, of the embryo. The concentration of Dpp directly affects the cell phenotype, the most dorsal phenotypes requiring the greatest concentration (2172).

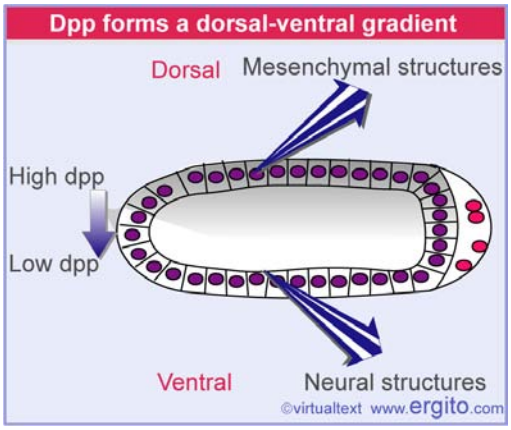

Figure 31.22 The morphogen Dpp forms a gradient originating on the dorsal side of the fly embryo. This prevents the formation of neural structures and induces mesenchymal structures.

The same pathway is involved in inducing the analogous structures in frog or mouse, but it is inverted with regard to the Dorsal-ventral axis. **Figure 31.23** shows that Bmp4 is secreted from one side of the egg. It is antagonized by a variety of factors. Neural tissues develop in the (dorsal) regions which Bmp4 is prevented from reaching.

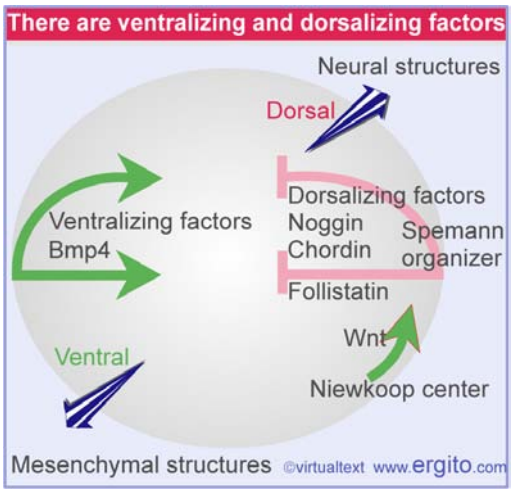

Figure 31.23 Two common pathways are used in early development of *Xenopus*. The Niewkoop center uses the Wnt pathway to induce the Spemann organizer. The organizer diffuses dorsalizing factors that counteract the effects of the ventralizing BMPs.

The crucial unifying feature is that neural tissues are induced when the activity of Dpp/Bmp4 is antagonized. Typically the Dpp/Bmp diffuses from a source, and different phenotypes may be produced by different concentrations of the morphogen. It is controversial whether the morphogen diffuses extracellularly or whether there may be a relay system that propagates it from cell to cell. Analogous pathways, triggered by different Bmps, are involved in the development of many organs.

Figure 31.24 compares the pathways in fly and frog. The basic principle is to control the availability of Dpp/Bmp. An antagonist binds to Dpp/Bmp and prevents it from binding to its receptor. The antagonists are large extracellular proteins. The antagonist is destroyed by a protease, releasing Dpp/Bmp. Neural tissue is formed in regions where Dpp/Bmp actions is prevented, whereas ectodermal tissue is formed in regions where Dpp/Bmp is activated.

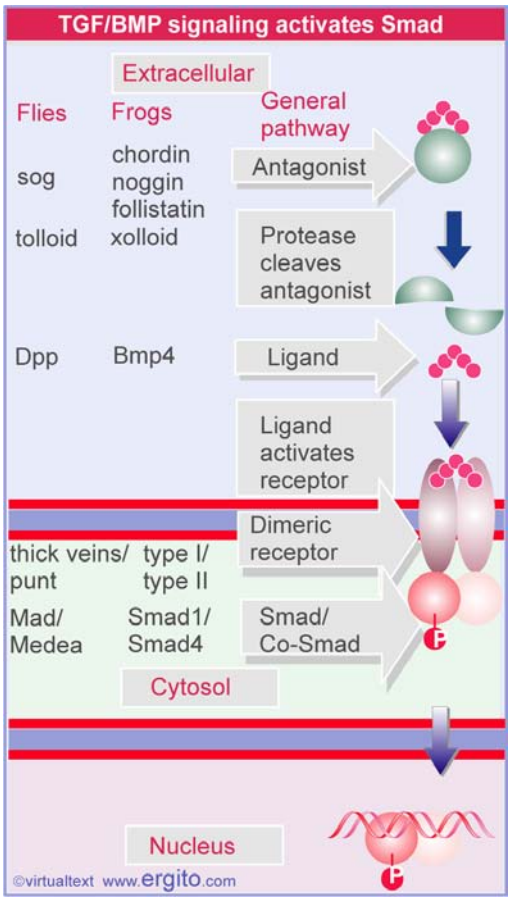

Figure 31.24 The TGF β /Bmp signaling pathway is conserved in evolution. The ligand may be sequestered by an antagonist, which is cleaved by a protease. Ligand binds to a dimeric receptor, causing the phosphorylation of a specific Smad, which together with a Co-Smad translocates to the nucleus to activate gene expression.

The fly pathway is well characterized for dorsal-ventral development. There are two types of mutants. Mutations in *sog* and *tsg* identify genes whose products antagonize Dpp, whereas mutations in *tolloid* suggest that it activates Dpp. Sog fulfills the role of antagonist illustrated in **Figure 31.24**, and is destroyed by the protease tolloid. Tsg is a sort of co-antagonist, which enhances the effect of Sog (1637; 1638; 1669).

The biochemical reactions actually have been better characterized for the corresponding frog proteins (Chordin is related to Sog). Frogs may have several such pathways, with a variety of Bmp ligands that interact in an overlapping manner with a family of receptors. The frog pathway shown in **Figure 31.24** is for the ventralizing effects of Bmp4, but the others are similar, although their specific effects on morphogenetic determination are of course different. There can be variation in specificity at each stage of the pathway. The antagonists, ligands, and receptors may be expressed in different places and times, providing specificity with regard to local concentrations of the morphogen, but there may also be partial redundancy. The genes for two proteins (Noggin and Chordin) both must be knocked out in mouse to produce a phenotype. Each receptor has specificity for certain Smads, so that

different target genes can be activated in different tissues.

Last updated on 4-4-2001

Reviews

369. De Robertis, E. M. and Sasi, Y. (1996). *A common plan for Dorsal-ventral patterning in Bilateria*. Nature 380, 37-40.

References

902. Ferguson, E. L. and Anderson, K. V. (1992). *decapentaplegic acts as a morphogen to organize dorsal-ventral pattern in the Drosophila embryo*. Cell 71, 451-461.

1637. Scott, I. C., Scott, I. C., Scott, I. C., Pappano, W. N., Maas, S. A., Cho, K. W., and Greenspan, D. S. (2001). *Homologues of Twisted gastrulation are extracellular cofactors in antagonism of BMP signalling.* Nature 410, 475-478.

1638. Ross, J. J., Shimmi, J. J., Vilmos, P., Petryk, A., Kim, H., Gaudenz, K., Hermanson, S., Ekker, S. C., O'Connor, M. B., and Marsh, J. L. (2001). *Twisted gastrulation is a conserved extracellular BMP antagonist.* Nature 410, 479-483.

1669. Chang, C., Holtzman, D. A., Chau, S., Chickering, T., Woolf, E. A., Holmgren, L. M., Bodorova, J., Gearing, D. P., Holmes, W. E., and Brivanlou, A. H. (2001). *Twisted gastrulation can function as a BMP antagonist.* Nature 410, 483-487.

2172. Nellen, D., Burke, R., Struhl, G., and Basler, K. (1996). *Direct and long-range action of a DPP morphogen gradient.* Cell 85, 357-368.

3695. Nellen, D., Affolter, M., and Basler, K. (1994). *Receptor serine/threonine kinases implicated in the control of Drosophila body pattern by decapentaplegic.* Cell 78, 225-237.

3696. Ruberte, E., Marty, T., Nellen, D., Affolter, M., and Basler, K. (1995). *An absolute requirement for both the type II and type I receptors, punt and thick veins, for dpp signaling in vitro.* Cell 80, 889-897.

GRADIENTS, CASCADES, AND SIGNALING PATHWAYS

6.31.14 Cell fate is determined by compartments that form by the blastoderm stage

Key Terms

Many organisms have a **segmented** body plan that divides the body into a number of repeating units, called segments, along the anterior-posterior axis.

In *Drosophila*, a **parasegment** is a unit composed of the rear of one segment and the front of the adjacent segment.

In *Drosophila* the **gap genes** are a set of genes that help set up the segmentation of the embryo. Gap genes encode transcription factors that are expressed in broad regions of the embryo. Gap genes activate transcription of the pair-rule genes.

In *Drosophila* the **pair-rule** genes are a set of genes that help set up the segmentation of the embryo. They are expressed in a striped pattern with one stripe in every other future segment.

In *Drosophila*, **segment polarity** genes are a set of genes that help set up the segmentation of the embryo. They are expressed in a striped pattern with one stripe in every future segment. Each stripe indicates the posterior margin of a segment.

Key Concepts

- A compartment of cells is defined at blastoderm and will give rise to a specific set of adult structures.
- Each segment consists of an anterior compartment and posterior compartment.
- Segmentation loci are divided into three groups of genes.

By the blastoderm stage, cells have begun to acquire information about the pathways they will follow and the structures they will therefore form. This information derives initially from the maternal regulators, and then it is further refined by the actions of zygotic genes. This makes it possible to draw a "fate map" of the blastoderm embryo to identify each region in terms of the adult segments that will develop from the descendants of the embryonic cells. The concept that is intrinsic in the fate map is that a region identified at blastoderm consists of a "compartment" of cells that will give rise specifically to a particular adult structure.

We can consider the development of *D. melanogaster* in terms of the two types of unit depicted in **Figure 31.25**: the segment and parasegment.

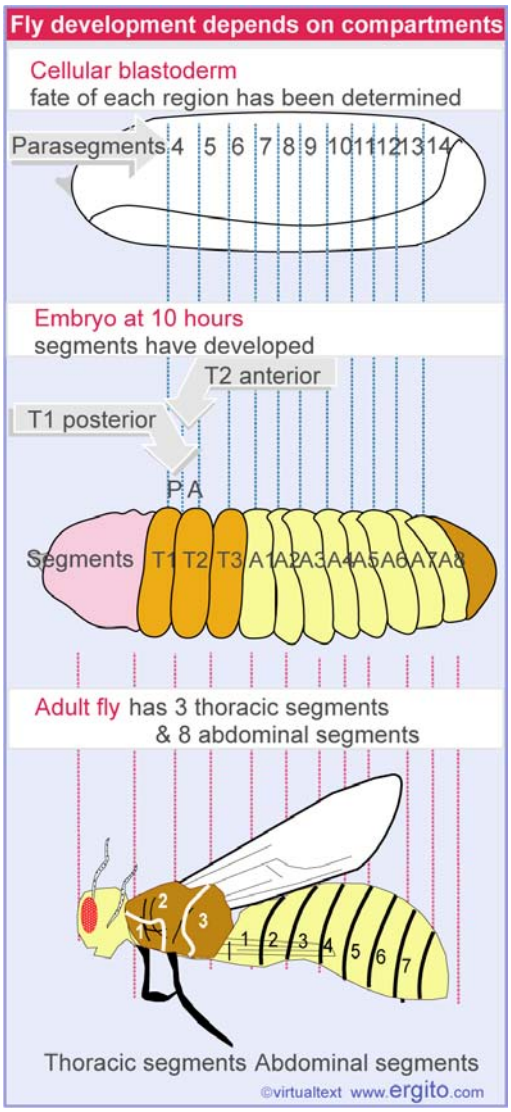

Figure 31.25 *Drosophila* development proceeds through formation of compartments that define parasegments and segments.

- The **segment** is a visible morphological structure. The adult fly consists of a series of clearly demarcated segments, and the larva has a series of corresponding segments separated by grooves. We are concerned primarily with the three thoracic (T) and eight abdominal (A) segments, about whose development most is known. The pattern of segmental units is determined by blastoderm, when the main mass of the embryo is divided into a series of alternating anterior (A) and posterior (P) compartments. So a segment consists of an A compartment succeeded by a P compartment; segment A3, for example, consists of compartments A3A and A3P.

- Another type of classification originates earlier, when divisions can first be seen at gastrulation. The embryo can be divided into **parasegments**, each consisting of a P compartment succeeded by an A compartment. Parasegment 8, for

example, consists of compartments A2P and A3A. In the 5-6 hour embryo, shallow grooves on the surface separate the adjacent parasegments. When segments form at around 9 hours, the grooves deepen and move, so that each segmental boundary represents the center of a parasegment. So the anterior part of the segment is derived from one parasegment, and the posterior part of the segment is derived from the next parasegment. In effect, the segmental units are initially evident as P-A pairs in parasegments, and then are recognized as A-P pairs in segments.

How are these compartments defined during embryogenesis? The general nature of segmentation mutants suggests that the functions of segmentation genes are to establish "rules" by which segments form; *a mutation changes a rule in such a way as to cause many or all segments to form improperly*. The drastic consequences of segment malformation make these mutants embryonic lethals – they die at various stages before metamorphosis into adults (880).

Probably ~30 loci are involved in segment formation. **Figure 31.26** shows that they can be classified according to the size of the unit that they affect:

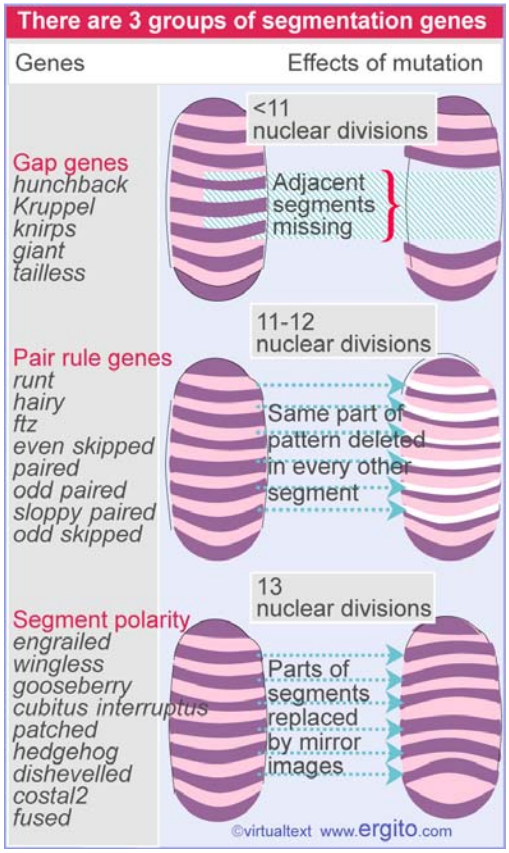

Figure 31.26 Segmentation genes affect the number of segments and fall into three groups.

- **Gap gene** mutants have a group of several adjacent segments deleted from the

final pattern. Four gap genes are involved in formation of the major body segments, and others are concerned with the head and tail structures.

- **Pair-rule** mutants have corresponding parts of the pattern deleted in every other segment. The afflicted segments may be even-numbered or odd-numbered. There are 8 pair-rule genes.

- **Segment polarity** mutants most often lose part of the P compartment of each segment, and it is replaced by a mirror image duplication of the A compartment. Some mutants cause loss of A compartments or middle segments. There are ~16 segment polarity genes.

These groups of genes are expressed at successive periods during development; and they define increasingly restricted regions of the egg, as can be seen from **Figure 31.27**. The maternal genes establish gradients from the anterior and posterior ends. The maternal gradients either activate or repress the gap genes, which are amongst the earliest to be transcribed following fertilization (following the 11th nuclear division); they divide the embryo into 4 broad regions. The gap genes regulate the pair-rule genes, which are transcribed slightly later; their target regions are restricted to *pairs* of segments. The pair-rule genes in turn regulate the segment polarity genes, which are expressed during the 13th nuclear division, and by now the target size is the *individual* segment.

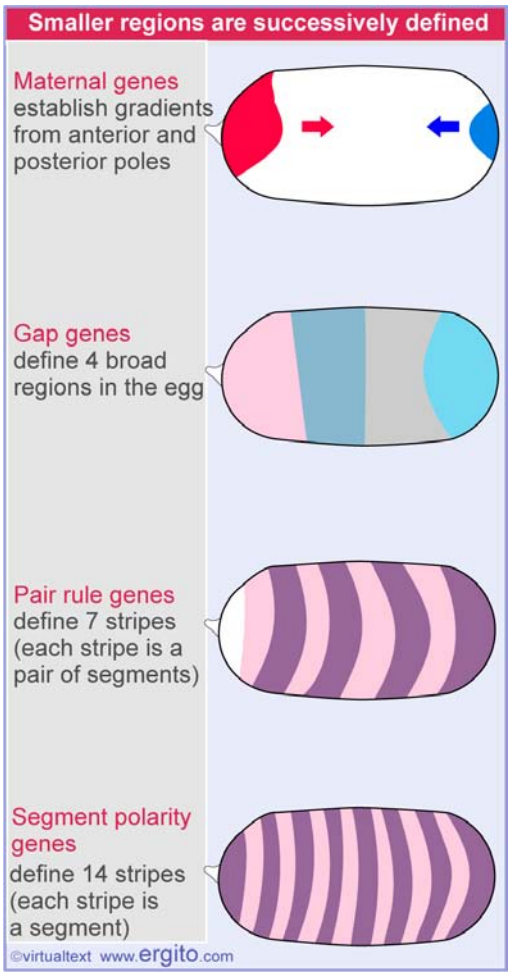

Figure 31.27 Maternal and segmentation genes act progressively on smaller regions of the embryo.

Many of the maternal genes, the gap genes, and the pair-rule genes are regulators of transcription. Their effects may be either to activate or to repress transcription; in some cases, a given protein may activate some target genes and repress other target genes, depending on its level or the context. The genes in any one class regulate one another as well as regulating the genes of the next class. When we reach the level of segment polarity genes, the nature of the regulatory event changes, and many of the gene products act on communication between cells to maintain borders between compartments, for example, to control the secretion of a protein from one cell to influence its neighbor (for review see 363; 358).

The principle that emerges from this analysis is that at each stage *a small number of maternal, gap, and pair-rule regulator proteins is used in combinatorial associations to specify the pattern of gene expression in a particular region of the embryo* (for review see 3705).

Reviews

358. Scott, M. P. and Carroll, S. B. (1987). *The segmentation and homeotic gene network in early Drosophila development*. Cell 51, 689-698.

363. Ingham, P. W. and Martinez-Arias, A. (1992). *Boundaries and fields in early embryos*. Cell 68, 221-235.

References

880. Nusslein-Vollhard, C. and Wieschaus, E. (1980). *Mutations affecting segment number and polarity in Drosophila*. Nature 287, 795-801.

3705. Lehmann, R. and Frohnhofer, H. G. (1989). *Segmental polarity and identity in the abdomen of Drosophila is controlled by the relative position of gap gene expression.* Development 107 Suppl, 21-29.

GRADIENTS, CASCADES, AND SIGNALING PATHWAYS

6.31.15 Gap genes are controlled by bicoid and by one another

Key Concepts

- Gap genes affect a group of segments, and are controlled by bicoid and by interactions among themselves.
- Pair-rule genes are expressed in either even- or odd-numbered segments, and are controlled by the gap genes.
- Segment polarity mutants are controlled by the pair-rule genes, and are expressed in segments where they affect anterior or posterior identification of the compartments.

The gap genes are controlled in two ways: they may respond directly to the bicoid morphogen; and they regulate one another. The four bands shown in **Figure 31.27** are created by the levels of the two proteins bicoid and hunchback (3704). The synthesis of hunchback mRNA is activated by bicoid (3702). Hunchback protein forms a gradient in the egg that in turn controls the expression of other genes.

The most anterior band in **Figure 31.27** consists of hunchback protein. The next band consists of Kruppel protein; transcription of the *Kruppel* gene is activated by hunchback protein. The next two bands consist of knirps and giant proteins. Transcription of these genes is repressed by hunchback. They are expressed in the posterior part of the embryo because nanos has prevented the expression of hunchback there.

Figure 31.28 examines the transition from the 4 band to the 7 striped stage in more detail. The detailed interactions among the gap proteins are determined by examining the pattern of the distribution of other gap proteins in a mutant lacking one particular gap protein. Hunchback plays an especially important role. It is expressed in a broad anterior region, with a gradient of decline in the middle of the embryo. High levels of hunchback repress *Kruppel;* this determines the anterior boundary of Kruppel expression, which rises just as hunchback falls off, in parasegment 3. But some level of hunchback is needed for *Kruppel* expression, so when the level of hunchback decreases further, *Kruppel* is turned off, around parasegment 5. In the same way, expression of *giant* responds to successive changes in the level of hunchback; and *knirps* expression requires the absence of hunchback.

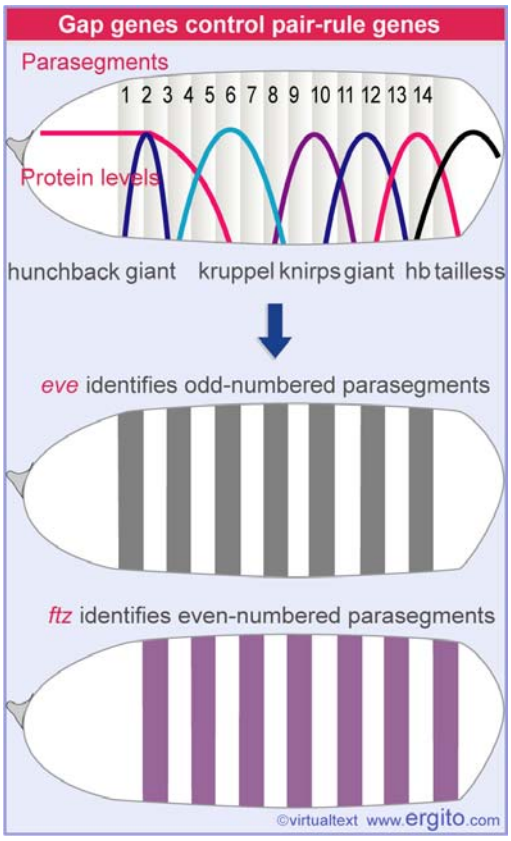

Figure 31.28 Expression of the gap genes defines adjacent regions of the embryo. The gap genes control the pair-rule genes, each of which is expressed in 7 stripes.

The control is refined further by interactions among the proteins. *The general principle is that one interaction may be required to express a protein in a particular region, and other interactions may be required to repress its expression at the boundaries.* The effects are worked out by examining pairwise interactions. For example, overexpression of *giant* causes the Kruppel band to become much narrower, suggesting that giant contributes to repressing the boundaries of Kruppel. The posterior margins of knirps and giant are determined by the operation of the terminal system. Altogether, these interactions mean that, as we proceed along the egg from anterior to posterior, any particular position can be defined by the levels of the various gap proteins.

References

3702. Tabata, T., Schwartz, C., Gustavson, E., Ali, Z., and Kornberg, T. B. (1995). *Creating a Drosophila wing de novo, the role of engrailed, and the compartment border hypothesis.* Development 121, 3359-3369.

3704. Simpson-Brose, M., Treisman, J., and Desplan, C. (1994). *Synergy between the hunchback and bicoid morphogens is required for anterior patterning in Drosophila.* Cell 78, 855-865.

GRADIENTS, CASCADES, AND SIGNALING PATHWAYS

6.31.16 Pair-rule genes are regulated by gap genes

Key Concepts

- Pair-rule genes are expressed in either even- or odd-numbered segments, and are controlled by the gap genes.

All the gap proteins are regulators of transcription, and in addition to regulating one another, they regulate the expression of the pair-rule genes that function at the next stage. Each pair-rule protein is found in a pattern of 7 "stripes" along the embryo, and **Figure 31.28** shows the approximate positions that these stripes will take as the result of expression of the gap genes (3707). (Of course, the parasegments have not developed yet, and are shown just to relate their positions to the protein distribution.) The 7 stripes of a pair-rule gene identify either all the odd-numbered parasegments (like *eve*) or all the even-numbered parasegments (like *ftz*). Two of the pair-rule genes, *hairy* and *eve*, are called primary pair-rule genes, because they are expressed first, and their pattern of expression influences the expression of the other pair-rule genes.

Recall that mutations in pair-rule genes delete half the segments. **Figure 31.29** compares the segmentation patterns of wild-type and *fushi tarazu* (*ftz*) larvae. The mutant has only half the number of segments, because every other segment is missing.

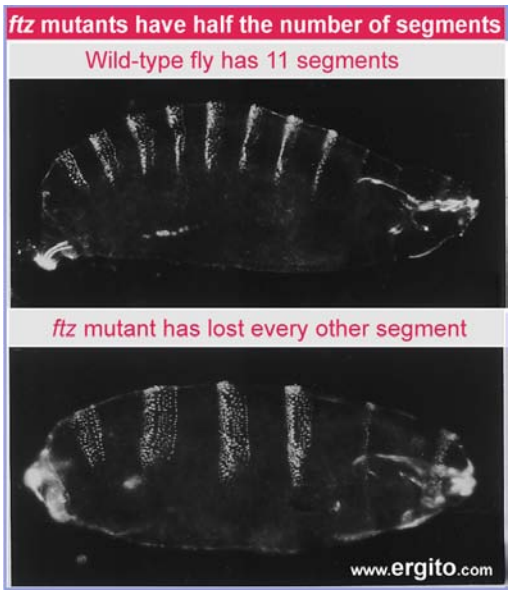

Figure 31.29 *ftz* mutants have half the number of segments present in wild-type. Photographs kindly provided by Walter Gehring.

The *ftz* mRNA is present from early blastoderm to gastrula stages of development.

Figure 31.30 shows the locations of the transcripts, visualized *in situ* at blastoderm in wild type. *The gene is expressed in 7 stripes, each 3-4 cells wide, running across the embryo.* As shown previously in **Figure 31.28**, the stripes correspond to even-numbered parasegments (4 = T1P/T2A, 6= T3P/A1A, 8 = A2P/A3A, etc.).

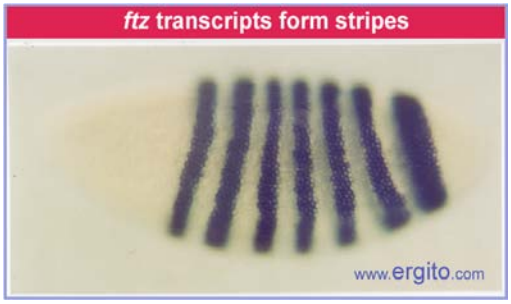

Figure 31.30 Transcripts of the *ftz* gene are localized in stripes corresponding to even numbered parasegments. The expressed regions correspond to the regions that are missing in the *ftz* mutant of the previous figure. Photograph kindly provided by Walter Gehring.

This pattern suggests a function for the *ftz* gene: *it must be expressed at blastoderm for the structures that will be descended from the even-numbered parasegments to develop.* Mutants in which *ftz* is defective lack these parasegments because the gene product is absent during the period when they must be formed. In other words, expression of *ftz* is required for survival of the cells in the regions in which it is expressed.

The expression of *ftz* is an example of the general rule that the stripes in which a pair-rule gene is expressed correspond to the regions that are missing from the embryo when the gene is mutated. *Compartments are therefore determined by the pattern of expression of segmentation genes.* The width of the stripe in which a gene is expressed corresponds to the size of the segmental unit that it affects. Different mechanisms are used to specify the expression patterns of different pair-rule genes; we have the most information about *ftz* and *eve*.

In the early embryo, *ftz* is uniformly expressed. If protein synthesis is blocked before the stripes develop, the embryo retains the initial pattern. So the development of stripes depends on the specific degradation of *ftz* RNA in the regions between the bands and at the anterior and posterior ends of the embryo. Once the stripes have developed, transcription of *ftz* ceases in the interbands and at the ends of the embryo. The specificity of transcription depends on regions upstream of the *ftz* promoter, and also on the function of several other segmentation genes. The transcription of *ftz* responds to other pair-rule genes (and perhaps gap genes) through elements that act on all stripes.

The expression pattern of *eve* is complementary to *ftz*, but has a different basis: it is controlled separately in each stripe. A detailed reconstruction using subregions of the *eve* promoter shows that the information for localization in each stripe is coded in a separate part of the promoter; the promoter can be divided into regions that respond to the local levels of gap gene products in particular parasegments. For example, the promoter region that is responsible for *eve* expression in parasegment 3 has binding

sites for the gap proteins bicoid, hunchback, giant, and Kruppel. **Figure 31.31** shows that this part of the promoter extends for 480 bp. It works in the following way. *eve* transcription is activated by hunchback and bicoid. The two boundaries are determined because the promoter is repressed by giant on the anterior side and by Kruppel on the posterior side (see also **Figure 31.28**). Other parts of the promoter respond to the protein levels in other parts of the embryo. So the different stripes of the primary pair-rule gene products are regulated by separate pathways, each of which is susceptible to activation by a particular combination of gap gene products and other regulators.

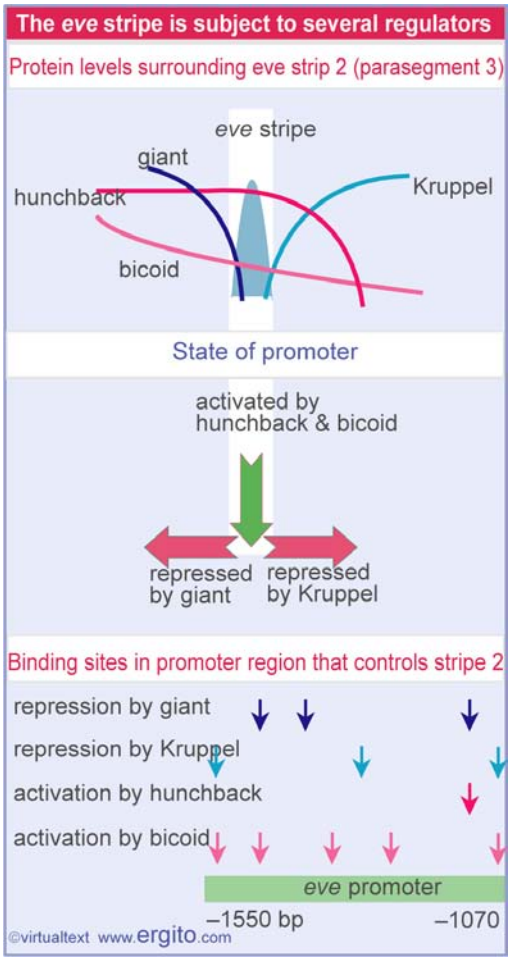

Figure 31.31 The *eve* stripe in parasegment 3 is activated by hunchback and bicoid. Repression by giant sets the anterior boundary; repression by Kruppel sets the posterior boundary. Multiple binding sites for these proteins in a 480 bp region of the promoter control expression of the gene.

This illustrates in miniature the principle that combinations of proteins control gene expression in local areas. The general principle is that generally distributed proteins (such as bicoid or hunchback) are needed for activation, whereas the borders are formed by selective repression (by giant and Kruppel in this particular example). We should emphasize that the hierarchy of gene control is not exclusively restricted to interactions between successive stages of control (maternal gap pair-rule). For

example, the involvement of bicoid protein in regulating *eve* transcription in parasegment 3 shows that a maternal gene may have a direct effect on a pair-rule gene.

The stripes of *eve* and *ftz* are fuzzy to begin with, and become sharper as development proceeds, corresponding to more finely defined units. **Figure 31.32** shows an example of an embryo simultaneously stained for expression of *ftz* and *eve*. Initially there is a series of alternating fuzzy stripes, but the stripes narrow from the posterior margin and sharpen on the anterior side as they intensify during development. This may depend on an autoregulatory loop, in which the expression of the gene is regulated by its own product.

Figure 31.32 Simultaneous staining for ftz (brown) and eve (grey) shows that they are first expressed as broad alternating stripes at the time of blastoderm (upper), but narrow during the next 1 hour of development (lower). Photographs kindly provided by Peter Lawrence.

References

3707. Lawrence, P. A. and Johnston, P. (1989). *Pattern formation in the Drosophila embryo: allocation of cells to parasegments by even-skipped and fushi tarazu.* Development 105, 761-767.

GRADIENTS, CASCADES, AND SIGNALING PATHWAYS

6.31.17 Segment polarity genes are controlled by pair-rule genes

Key Concepts

- Segment polarity genes are expressed in segments where they affect anterior or posterior identification of the compartments.

The pair-rule genes control the expression of the segment polarity genes, which are expressed in 14 stripes. Each stripe identifies a segment. The compartmental pattern in which segment polarity genes are expressed is exceedingly precise. Perhaps the ultimate demonstration of precision is provided by the pattern gene *engrailed*. The function coded by *engrailed* is needed in all segments and is concerned with the distinction between the A and P compartments. *engrailed* is expressed in every P compartment, but not in A compartments (3701; 3700; 3702). Mutants in this gene do not distinguish between anterior and posterior compartments of the segments.

Antibodies against the protein coded by *engrailed* react against the nucleus of cells expressing it. The regions in which *engrailed* is expressed form a pattern of stripes. When the stripes of engrailed protein first become apparent, *they are only one cell wide*.**Figure 31.33** shows the pattern at a stage when each segment has a stripe just 1 cell in width, with the stripe beginning to widen into several cells.

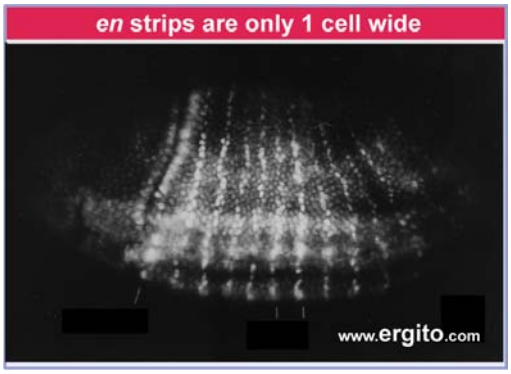

Figure 31.33 Engrailed protein is localized in nuclei and forms stripes as precisely delineated as 1 cell in width. Photograph kindly provided by Patrick O#Farrell.

Actually, the pattern of stripes becomes established over a 30 minute period, moving along the embryo from anterior to posterior. Initially one stripe is apparent; then every other segment has a stripe; and finally the complete pattern has a stripe 3-4 cells wide corresponding to the P compartment of every segment.

The expression of *engrailed* is of particular importance, because it defines the boundaries of the actual compartments from which adult structures will be derived.

The initial 1-cell-wide stripes of engrailed protein form at the anterior boundaries of both the *ftz* and *eve* stripes, and delineate what will become the anterior boundary of every P compartment. Why is *engrailed* initially transcribed exclusively in this anterior edge, within the broader stripes of *ftz* and *eve* expression? This question is a specific example of a more general question: how can a broad stripe be subdivided into more restricted, narrower stripes? We can consider two general types of model:

- A combinatorial model supposes that different genes are expressed in overlapping patterns of stripes. A pattern of stripes develops for each of the pair-rule genes. The different pair-rule gene stripes overlap, because they are out of phase with one another. As a result of these patterns, different cells in the cellular blastoderm express different combinations of pair-rule genes. Each compartment is defined by the particular combination of the genes that are expressed, and these combinations determine the responses of the cells at next stage of development. In other words, the segmentation genes are controlled by the pair-rule genes in the same general manner that the pair-rule genes are controlled by the gap genes.

- A boundary model supposes that a compartment is defined by the striped pattern of expression, but that interactions involving cell-cell communication at the boundaries cause subdivisions to arise within the compartment. In the case of *engrailed*, we would suppose that some unique event is triggered by the juxtaposition of cells possessing ftz (or eve) with cells that do not, and this is necessary to trigger *engrailed* expression.

Each of the 14 segments is subdivided further into anterior and posterior compartments by the activities of the segment polarity genes. The actions of the segment polarity genes are the same in every segment. For example, *engrailed* distinguishes the A and P compartments.

engrailed is a transcription factor, but other segment polarity genes have different types of functions. The products of the segment polarity genes include secreted proteins, transmembrane proteins, kinases, cytoskeletal proteins, as well as transcription factors. Cell-cell interactions become important at this stage for defining and maintaining the nature of the compartments (see *Molecular Biology 6.31.18 Wingless and engrailed expression alternate in adjacent cells*).

References

3700. Kornberg, T, Sidén, I, O'Farrell, P, and Simon, M (1985). *The engrailed locus of Drosophila: in situ localization of transcripts reveals compartment-specific expression.* Cell 40, 45-53.

3701. Kornberg, T. (1981). *Engrailed: a gene controlling compartment and segment formation in Drosophila.* Proc. Natl. Acad. Sci. USA 78, 1095-1099.

3702. Tabata, T., Schwartz, C., Gustavson, E., Ali, Z., and Kornberg, T. B. (1995). *Creating a Drosophila wing de novo, the role of engrailed, and the compartment border hypothesis.* Development 121, 3359-3369.

GRADIENTS, CASCADES, AND SIGNALING PATHWAYS

6.31.18 Wingless and engrailed expression alternate in adjacent cells

Key Concepts

- Wingless and engrailed have a mutually reinforcing interaction between neighboring cells.
- Wingless is secreted at the posterior boundary of a cell.
- It activates the Fz (or dFz2) receptor in the adjacent cell, which triggers the translocation of Armadillo to the nucleus.
- Armadillo causes engrailed to be expressed, and engrailed causes secretion of hedgehog at the anterior boundary, where it acts on the neighboring cell to maintain Wingless expression.

The circuit that defines the boundaries between anterior and posterior compartments is based on mutual interactions between segment polarity genes. *wingless* codes for a protein that is secreted and taken up by the adjacent row of cells. It is initially expressed in a row of cells immediately adjacent to the anterior side of the cells expressing *engrailed;* so *wingless* comes to identify the posterior boundary of the preceding parasegment. The initial expression of *engrailed* in response to ftz and eve is shortly replaced by an autoregulatory loop in which secretion of wingless protein from the adjacent cells is needed for expression of *engrailed;* and expression of *engrailed* is needed for expression of *wingless*. This keeps the boundary sharp.

The wingless signaling pathway is one of the most interesting, and has close parallels in all animal development. Like other signaling pathways utilized in development, it is initiated by an extracellular ligand, and results in the expression of a transcription factor, although the interactions between components of the pathway are somewhat unusual.

In fly embryonic development at the stage of segmental definition, the cells that define the boundaries of the A and P compartments express wingless (Wg) and engrailed (En) in a reciprocal relationship. **Figure 31.34** shows that wingless protein is secreted from a cell at a boundary, and acts upon the cell on its posterior side. The *wingless* signaling pathway causes the *engrailed* gene to be expressed. Engrailed causes the production of hedgehog (Hh) protein, which in turn is secreted. Hedgehog acts on the cell on its anterior side to maintain *wingless* expression. Wg is also required for patterning of adult eyes, legs, and wings (hence its name) (898).

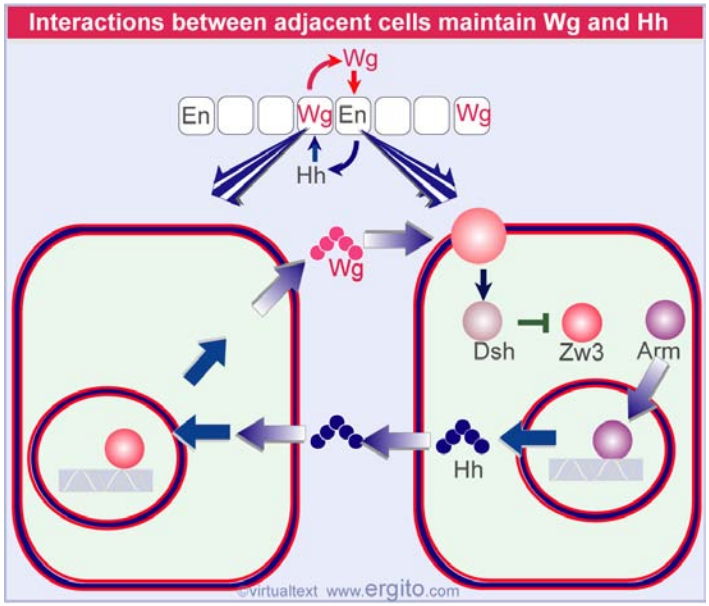

Figure 31.34 Reciprocal interactions maintain Wg and Hh signaling between adjacent cells. Wg activates a receptor, which activates a pathway leading to translocation of Arm to the nucleus. This activates engrailed, which leads to expression of Hedgehog protein, which is secreted to act on the neighboring cell, where it maintains Wg expression.

The identification of the receptor for Wg on the posterior cell has actually been very difficult. Wg interacts with frizzled *in vitro*, but mutational analysis suggests that the related protein, DFz2 (*Drosophila frizzled-2*) is the receptor. It is possible that these may play redundant roles. Another protein that is required for reception/signaling is the product of *arrow*, which is a single-pass membrane protein and is classified as a coreceptor (2894). The frizzled family members are 7-membrane pass proteins, with the appearance of classical receptors (although the major pathway does not appear to involve G proteins) (903; 2896).

Last updated on 9-3-2002

References

898. Rijsewijk, F. M. (1987). *The Drosophila homolog of the mouse mammary oncogene int-1 is identical to the segment polarity gene wingless*. Cell 50, 649-657.

903. Bhanot, P. et al. (1995). *A new member of the frizzled family from Drosophila functions as a wingless receptor*. Nature 382, 225-230.

2894. Wehrli, M., Dougan, S. T., Caldwell, K., O'Keefe, L., Schwartz, S., Vaizel-Ohayon, D., Schejter, E., Tomlinson, A., and DiNardo, S. (2000). *arrow encodes an LDL-receptor-related protein essential for Wingless signalling.* Nature 407, 527-530.

2896. Muller, H., Samanta, R., and Wieschaus, E. (1999). *Wingless signaling in the Drosophila embryo: zygotic requirements and the role of the frizzled genes.* Development 126, 577-586.

GRADIENTS, CASCADES, AND SIGNALING PATHWAYS

6.31.19 The wingless/wnt pathway signals to the nucleus

Key Concepts

- Wingless (in Drosophila) and Wnt (in vertebrates) activate a receptor that blocks the action of a cytosolic Ser/Thr kinase.
- The kinase phosphorylates Armadillo/ β-catenin that is localized in cytosolic complexes.
- In the absence of phosphorylation, Armadillo/ β-catenin is stabilized, and translocates to the nucleus where it activates transcription.
- A separate pool of Armadillo/ β-catenin is present in complexes at the cell surface, but is not a target for the pathway.
- The function of the cancer-causing gene APC is to destabilize β-catenin, and colon cancers caused by mutation in APC have elevated levels of β-catenin.

The interaction between Wingless and its receptor activates a signaling pathway. The effector of the pathway is a protein called Arm (Armadillo) in *Drosophila* and β-catenin in vertebrate cells. A cytoplasmic pool of the effector protein is constitutively degraded, and the role of the signaling pathway is to block the degradation. When this happens, Arm/ β-catenin is transported to the nucleus, where it activates transcription of a set of target genes.

Mutants in other genes that have segment polarity defects similar to *wg* mutants identify the other components of the pathway. They signal positively to execute the pathway. Mutations that have the opposite phenotype, and that block the pathway, identify proteins that are required to degrade Arm/ β-catenin. **Figure 31.35** shows the results of ordering the genes genetically, and defining the biochemical interactions between their products.

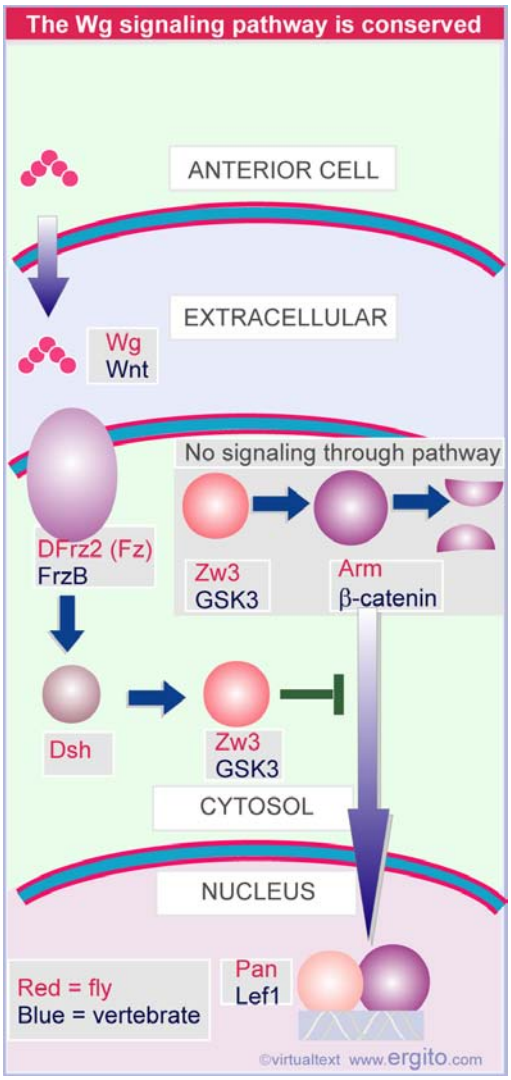

Figure 31.35 Wg secretion is assisted by porc. Wg activates the Dfz2 receptor, which inhibits Zw3 kinase. Active Zw3 causes turnover of Arm. Inhibition of Zw3 stabilizes Arm, allowing it to translocate to the nucleus. In the nucleus, Arm partners Pan, and activates target genes (including engrailed). A similar pathway is found in vertebrate cells (components named in blue).

The fly and vertebrate transduction pathways have homologous components. The components in the order of their function in the pathway are:

- Dsh (coded by *Dishevelled*) is a phosphoprotein that responds to the interaction of wingless/Wnt with the frizzled receptor.
- Dsh signals to a Ser/Thr kinase, called Zw3 in Drosophila, but called GSK3 in vertebrates (named for its historical identification as glycogen synthase kinase, but in fact a homologue of Zw3). Zw3/GSK3 is constitutively active, unless and until it is inactivated by Dsh.

Zw3/GSK phosphorylates a serine in the N-terminus of Arm/ β-catenin. This creates a binding site for the small polypeptide ubiquitin, which causes the protein to be degraded by the proteasome (see *Molecular Biology 2.8.31 Ubiquitination targets proteins for degradation*).

Zw3/GSK is the catalytic component of a protein aggregate called *the* β-*catenin destruction complex*. A scaffolding protein called Axin enables Zw3/GSK to bind to the Arm/ β-catenin target. Another component, called APC, dissociates the phosphorylated Arm/ β-catenin from Zw3/GSK after the reaction is complete 4518.

This signaling pathway is also implicated in colon cancer. Mutations in *APC* (adenomatous polyposis coli) are common in colon cancer. As a component of the β-catenin destruction complex, APC is required for destabilizing Arm/ β-catenin. The mutant proteins found in colon cancer allow levels of β-catenin to increase. Mutations in β-catenin that increase its stability have the same effect. Inappropriate activation of the Wnt signaling pathway therefore contributes to colon cancer.

When its degradation is inhibited and Arm/ β-catenin accumulates, it translocates to the nucleus. There it binds to a partner (called Pan in Drosophila, and called Tcf/LEF1 in vertebrates, depending on the system). The complex activates transcription at promoters that are bound by the Pan/Tcf subunit. When Tcf1 binds to DNA, β-catenin can activate transcription at the target promoters (904; 905).

So wingless/Wnt signaling controls the availability of Arm/ β-catenin by causing its degradation to be inhibited. The most surprising feature of this pathway is the nature of the Arm/ β-catenin protein. It has two unconnected activities.

- It is a component of a complex that links the cytoskeleton at adhesion complexes. β-catenin binds to cadherin. Mutations of *armadillo* that disrupt the cadherin-binding site show a defect in cell adhesion.
- A separate domain of Arm/ β-catenin has a transactivation function when the protein translocates to the nucleus (2167).

How does Arm/ β-catenin participate in two so very different activities? It is in fact bound by a large number of potential partners. Most of them recognize a series of repeats in the central sequence of the protein, with the result that most of these complexes are mutually exclusive. The various complexes are localized in different places in the cell. When Arm/ β-catenin binds to cadherins or certain other proteins of the plasma membrane, it forms a complex that participates in cell-cell adhesion. This complex is not a target for the wingless/Wnt pathway, which acts on Arm/ β-catenin that is free in the cytosol (for review see 2897).

Last updated on January 6, 2004

Reviews

2897. Zhurinsky, J., Shtutman, M., and Ben-Ze'ev, A. (2000). *Plakoglobin and beta-catenin: protein interactions, regulation and biological roles.* J. Cell Sci. 113 , 3127-3139.

References

904. Molenaar, M. et al. (1996). *XTcf-3 transcription factor mediates β -catenin-induced axis formation in Xenopus embryos*. Cell 86, 391-399.

905. Brunner, E. (1997). *pangolin encodes a Lef-1 homologue that acts downstream of Armadillo to transduce the Wingless signal in Drosophila*. Nature 385, 829-833.

2167. Graham, T. A., Weaver, C., Mao, F., Kimelman, D., and Xu, W. (2000). *Crystal structure of a beta-catenin/Tcf complex.* Cell 103, 885-896.

4518. Xing, Y., Clements, W. K., Kimelman, D., and Xu, W. (2003). *Crystal structure of a beta-catenin/axin complex suggests a mechanism for the beta-catenin destruction complex.* Genes Dev. 17, 2753-2764.

GRADIENTS, CASCADES, AND SIGNALING PATHWAYS

6.31.20 Complex loci are extremely large and involved in regulation

Key Terms

Homeotic genes are defined by mutations that convert one body part into another; for example, an insect leg may replace an antenna.

A **complex locus** (of *D. melanogaster*) has genetic properties inconsistent with the function of a gene representing a single protein. Complex loci are usually very large (>100 kb) at the molecular level.

Key Concepts

- Complex loci were identified by interallelic interactions that did not fit the usual complementation behavior.
- They are extremely large and may include multiple protein-coding units as well as *cis*-acting regulatory sites.
- The order of mutations from upstream to downstream corresponds to the order of the body parts that are affected from anterior to posterior.
- *ANT-C* includes several genes that affect head segments, whereas *BX-C* has only three genes and many regulatory sites.
- A segment is distinguished from the preceding (anterior-side) segment by expression of an additional protein-coding unit.
- Loss of a protein-coding unit causes a homeotic transformation in which a segment has the identity of the segment on its anterior-side.

Segment polarity genes control the anterior-posterior pattern within each segment. **Homeotic genes** impose the program that determines the unique differentiation of each segment. Most homeotic genes are expressed in a spatially restricted manner that corresponds to parasegments.

Homeotic genes interact in complicated interlocking patterns. Many homeotic genes code for transcription factors that act upon other homeotic genes as well as upon other target loci. As a result, a mutation in one homeotic gene influences the expression of other homeotic genes. The consequence is that *the final appearance of a mutant depends not only on the loss of one homeotic gene function, but also on how other homeotic genes change their spatial patterns in response to the loss.*

Homeotic genes act during embryogenesis. Their expression depends on the prior expression of the segmentation genes; we might regard the homeotic genes as integrating the pattern of signals established by the segmentation genes. Homeotic mutants "transform" part of a segment or an entire segment into another type of segment; they may cause one segment of the abdomen to develop as another, legs to

develop in place of antennae, or wings to develop in place of eyes. Note that homeotic genes do not *create* patterns *de novo;* they modify cell fates that are determined by genes such as the segment polarity genes, by switching the set of genes that functions in a particular place. Indeed, the segment polarity genes are active at about the same time as the peak of expression of the homeotic genes.

The genetic properties of some homeotic mutations are unusual and led to the identification of complex loci (for review see 357; 2881; 2882). A conventional gene – even an interrupted one – is identified at the level of the genetic map by a cluster of noncomplementing mutations. In the case of a large gene, the mutations might map into individual clusters corresponding to the exons. A hallmark of a **complex locus** is that, in addition to rather well-spaced groups of mutations, extending over a relatively large map distance, there are complex patterns of complementation, in which some pairwise combinations complement but others do not. The individual mutations may have different and complex morphological effects on the phenotype. These relationships are caused by the existence of an array of regulatory elements. Many of the bizarre results that are obtained in complementation assays turn out to result from mutations in promoters or enhancers that affect expression in one cell type but not another. We now recognize that complex loci do not have any novel features of genetic organization, apart from the fact that they have many regulatory elements that control expression in different parts of the embryo.

Two of the complex loci are involved in regulating development of the adult insect body. The *ANT-C* and *BX-C* complex loci together provide a continuum of functions that specify the identities of all of the segmented units of the fly. Each of these complexes contains several homeotic genes. The two separate complexes may have evolved from a split in a single ancestral complex, as suggested by the evolution of the corresponding genes in other species. In the beetle *Tribolium,* the *ANT-C* and *BX-C* complexes are found together at a single chromosomal location. The individual genes may have been derived from duplications and mutations of an original ancestral gene. And in mammals, there are arrays of related genes whose individual members are related sequentially to the genes of the *ANT-C* and *BX-C* complexes (see *Molecular Biology 6.31.22 The homeobox is a common coding motif in homeotic genes*).

The homeotic genes clustered at the *ANT-C* and *BX-C* complexes show a relationship between genetic order and the position in which they are expressed in the body of the fly. *Proceeding from left to right, each homeotic gene in the complex acts upon a more posterior region of the fly. The basic principle is that formation of a compartment requires the gene product(s) expressed in the previous compartment, plus a new function coded by the next gene along the cluster.* So loss-of-function mutations usually cause one compartment to have the phenotype of the corresponding compartment on its anterior side. The individual genes code for a set of transcription factors that have related DNA-binding domains (see next section).

The identities of the most anterior parts of the fly (parasegments 1-4) are specified by *ANT-C,* which contains several homeotic genes, including *labial (lab), proboscipedia (pb), Deformed (Dfd), Sex combs reduced (Scr),* and *Antennapedia (Antp).* The homeotic genes lie in a cluster over a region of ~350 kb, but several other genes are interspersed; most of these genes are regulators that function at different stages of development (356).

Figure 31.36 correlates the organization of *ANT-C* with its effects upon body parts. Adjacent genes are expressed in successively more posterior parts of the embryo, ranging from the leftmost gene *labial* (the most anterior acting, which affects the head) to the rightmost gene *Antp* (the most posterior acting, which affects segments T2–T3).

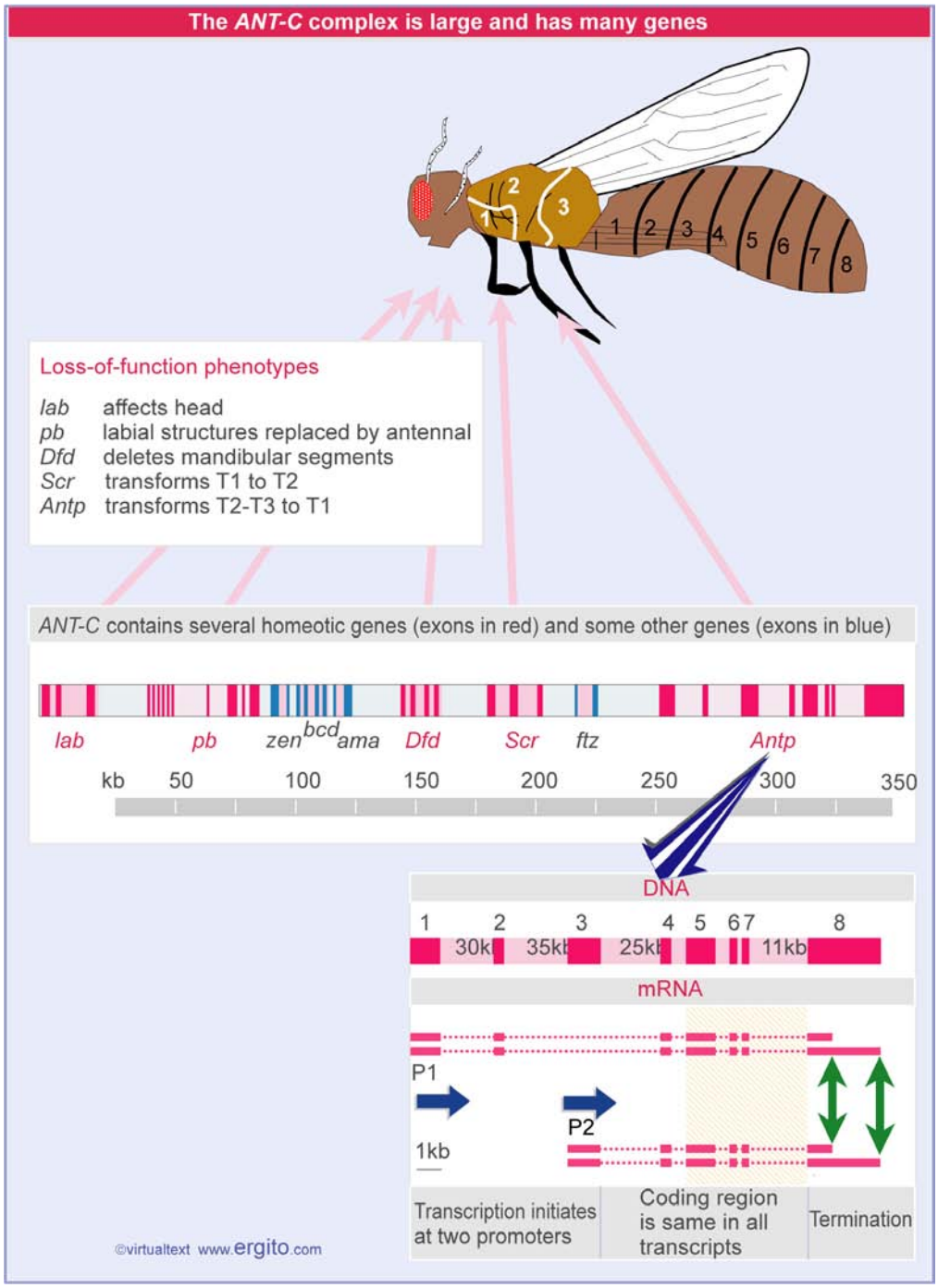

Figure 31.36 The homeotic genes of the *ANT-C* complex confer identity on the most anterior segments of the fly. The genes vary in size, and are interspersed with other genes. The *antp* gene is very large and has alternative forms of expression.

The *Antp* gene gave its name to the complex, and among the mutations in it are alleles that change antennae into second legs, or second and third legs into first legs. *Antp* usually functions in the thorax; it is needed both to promote formation of segments T2-T3 and to suppress formation of head structures. Loss of function therefore causes T2-T3 to resemble the more anterior structure of T1; gain of function, for example, by overexpression in the head, causes the anterior region to develop structures of the thorax. (The molecular action of *Antp* is to prevent the action of genes *hth* and *exd* that promote formation of antennal structures. Hth causes exd to be imported into the nucleus, where it switches on the genes that make the antenna.)

Figure 31.36 summarizes the organization of the gene. It has 8 exons, separated by very large introns, and altogether spanning ~103 kb. The single open reading frame begins only in exon 5, and apparently gives rise to a protein of 43 kD. The discrepancy between the length of the locus and the size of the protein means that only 1% of its DNA codes for protein (882).

Transcription starts at either of two promoters, located ~70 kb apart! One promoter is located upstream of exon 1, the other upstream of exon 3. Use of the first promoter is associated with omission of exon 3. The transcripts generated from either promoter end either within or after exon 8. All the transcripts appear to code for the same protein. Each promoter has its own tissue-specific expression pattern (3708). We do not know if there is any significance to the difference in the structure of the two types of transcript.

The other genes of the *ANT-C* complex are expressed in the head and first thoracic segment. In the most anterior compartments, *lab, pb, Dfd,* have unique patterns of expression, so that deletions of segmental regions can result from loss-of-function mutations. An exception to the left-right/anterior-posterior order of action is that loss of *Scr* allows the overlapping *Antp* to function, that is, the direction of transformation is opposite from usual.

Reviews

356. Regulski, M., Harding, K., Kostriken, R., Karch, F., Levine, M., and McGinnis, W. (1985). *Homeo box genes of the Antennapedia and bithorax complexes of Drosophila*. Cell 43, 71-80.

357. Scott, M. P. (1987). *Complex loci of Drosophila*. Annu. Rev. Biochem. 56, 195-227.

2881. Montgomery, G. (2002). *E. B. Lewis and the bithorax complex. Part I.* Genetics 160, 1265-1272.

References

882. Scott, M. P. et al. (1983). *The molecular organization of the Antennapedia locus of Drosophila*. Cell 35, 763-766.

2882. Duncan, I. and Montgomery, G. (2002). *E. B. Lewis and the Bithorax Complex. Part ii. from cis-trans test to the genetic control of development.* Genetics 161, 1-10.

3708. Jorgensen, E. M. and Garber, R. L. (1987). *Function and misfunction of the two promoters of the Drosophila Antennapedia gene.* Genes Dev. 1, 544-555.

GRADIENTS, CASCADES, AND SIGNALING PATHWAYS

6.31.21 The *bithorax* complex has *trans*-acting genes and *cis*-acting regulators

Key Terms

The **bithorax complex** is a group of homeotic genes which are responsible for the diversification of the different segments of the fly.

Key Concepts

- *bithorax* controls body structures from T2 to A8.
- The ultrabithorax domain has the *Ubx* transcription unit.
- The infraabdominal domain has the *AbdA* and *AbdB* transcription units.
- The order of units on the genetic map coincides with the order of body parts.
- Expression of additional units specifies more posterior body parts.

The classic complex homeotic locus is *BX-C*, the **bithorax complex**, characterized by several groups of homeotic mutations that affect development of the thorax, causing major morphological changes in the abdomen. When the whole complex is deleted, the insect dies late in embryonic development. Within the complex, however, are mutations that are viable, but which change the phenotype of certain segments. An extreme case of homeotic transformation is shown in **Figure 31.37**, in which a triple mutation converts T3A (which carries the halteres [truncated wings]) into the tissue type of the T2 (which carries the wings). This creates a fly with four wings instead of the usual two (881).

Figure 31.37 A four-winged fly is produced by a triple mutation in *abx*, *bx*, and *pbx* at the *BX-C* complex. Photograph kindly provided by Ed Lewis.

The genetic map of *BX-C* is correlated with the body structures that it controls in the fly in **Figure 31.38**. The body structures extend from T2 to A8. The *BX-C* complex is therefore concerned with the development of the major part of the body of the fly.

Like *ANT-C*, a crucial feature of this complex is also that mutations affecting particular segments lie in the same order on the genetic map as the corresponding segments in the body of the fly. Proceeding from left to right along the genetic map, mutations affect segments in the fly that become successively more posterior (895; for review see 3709).

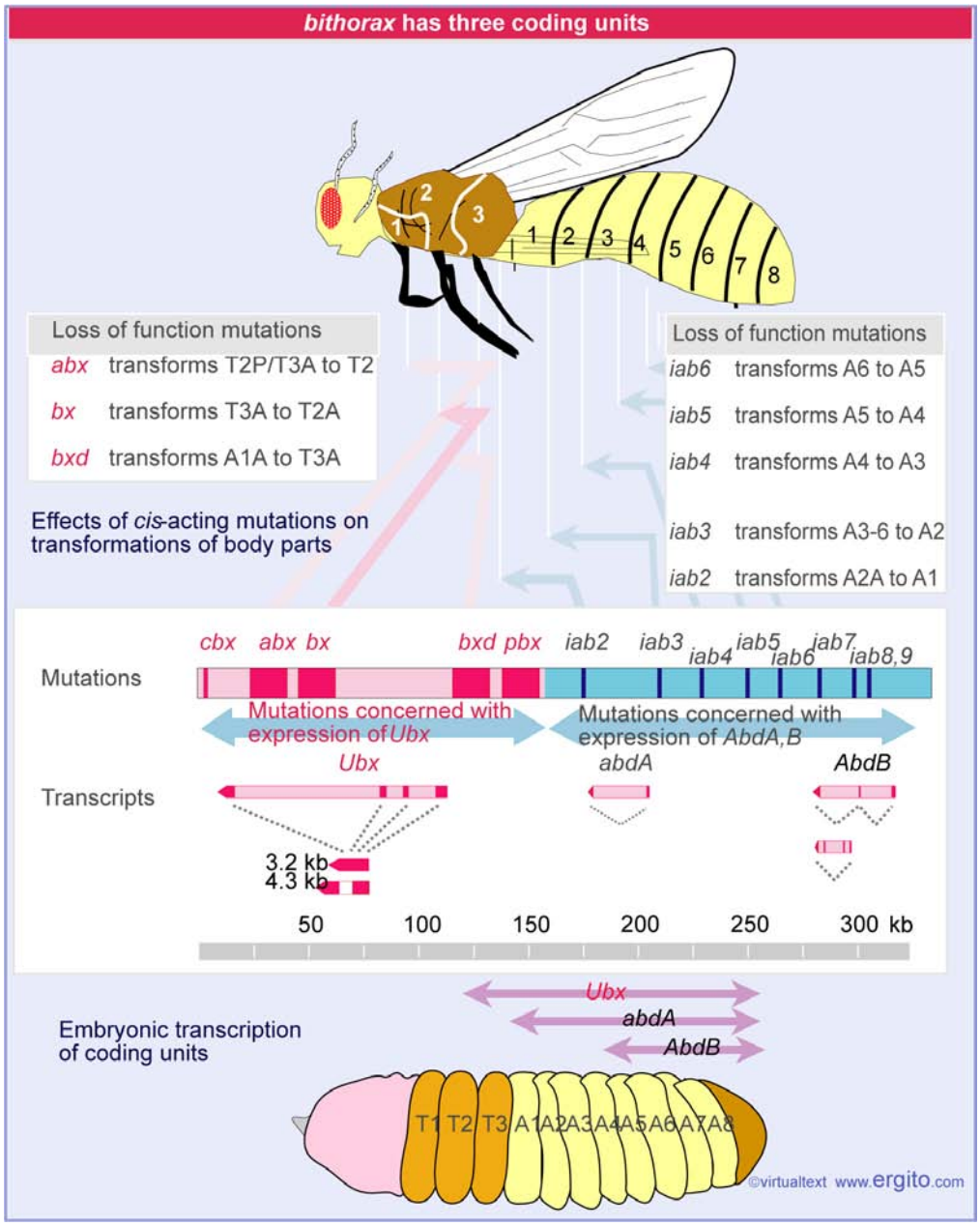

Figure 31.38 The bithorax (*BX-C*) locus has 3 coding units. A series of regulatory mutations affects successive segments of the fly. The sites of the regulatory mutations show the regions within which deletions, insertions, and translocations confer a given phenotype.

A difference between *ANT-C* and *BX-C* is that *ANT-C* functions largely or exclusively via its protein-coding loci, but *BX-C* displays a complex pattern of *cis*-acting interactions in addition to the effects of mutations in protein-coding

regions. The *BX-C* occupies 315 kb, of which only 1.4% codes for protein. The individual mutations fall into two classes:

- Three transcription units (*Ubx*, *abdA*, *AbdB*) produce mRNAs that code for proteins. The transcription units are large (>75 kb for *Ubx*, and >20 kb each for *abdA* and *AbdB*). Each contains several large introns. (The *bxd* and *iab4* regions produce RNAs that do not code for proteins; again, the transcription units are large, and the RNA products are spliced. Their functions are unknown.)
- There are *cis*-acting mutations at intervals throughout the entire cluster. They control expression of the transcription units. *Cis*-acting mutations of any particular type may occur in a large region. The locations shown on the map are only approximate, and the boundaries within which mutations of each type may occur are not well defined.

As a historical note, the complex was originally defined in terms of two "domains." Mutations in the *Ultrabithorax* domain were characterized first; they have the thoracic segments T2P-T3P and the abdominal compartment A1A as their targets (this corresponds to parasegments 5-6). These mutations lie either in the *Ubx* transcription unit or in the *cis*-acting sites that control it. The mutations within the ultrabithorax domain are named for their phenotypes. The *bx* and *bxd* types are identified by a series of mutations, in each case dispersed over ~10 kb. The *abx* and *pbx* mutations are caused by deletions, which vary from 1-10 kb.

Mutations in the *Infraabdominal* domain were found later; they have the abdominal segments A1P-A8P (parasegments 7-14) as targets. These mutations lie either in the *AbdA,B* transcription units, or in the *cis*-acting sites that control them. Within the infraabdominal domain, *cis*-acting mutations are named systematically as *iab2-9.* These mutations affect individual compartments, or sometimes adjacent sets of compartments, as shown at the top of the figure (885).

Proceeding from left to right along the cluster, transcripts are found in increasingly posterior parts of the embryo, as shown at the bottom of the figure. The patterns overlap. *Ubx* has an anterior boundary of expression in compartment T2P (parasegment 5), *abdA* is expressed from compartment A1P (parasegment 7), and *AbdB* is expressed in compartments posterior to compartment A4P (parasegment 10).

Transcription of *Ubx* has been studied in the most detail. The *Ubx* transcription unit is ~75 kb, and has alternative splicing patterns that give rise to several short RNAs. A transient 4.7 kb RNA appears first, and then is replaced by RNAs of 3.2 and 4.3 kb. A feature common to both the latter two RNAs is their inclusion of sequences from both ends of the primary transcript. Of course, there may be other RNAs that have not yet been identified.

We do not yet have a good idea of whether there are significant differences in the coding functions of these RNAs. The first and last exons are quite lengthy, but the interior exons are rather small. Small exons from within the long transcription unit may enter mRNA products by means of alternative splicing patterns. So far, however, we do not know of any functional differences in the Ubx proteins produced by the various modes of expression.

Ubx proteins are found in the compartments that correspond to the sites of transcription, that is in T2P-A1 and at lower levels in A2-A8. So the Ubx unit codes for a set of related proteins that are concentrated in the compartments affected by mutations in the *Ultrabithorax* domain. Ubx proteins are located in the nucleus, and they fall into the general type of transcriptional regulators whose DNA-binding region consists of a homeodomain.

We can understand the general function of the *BX-C* complex by considering the effects of loss-of-function mutations. If the entire complex is deleted, the larva cannot develop the individual types of segments. In terms of parasegments (which are probably the affected units), all the parasegments differentiate in the same way as parasegment 4; the embryo has 10 repetitions of the repeating structure T1P/T2A all along its length, in place of the usual compartments between parasegments 5 and 14. In effect, the absence of *BX-C* functions allows *Antp* to be expressed throughout the abdomen, so that all the segments take on the characteristic of a segment determined by *Antp*; *BX-C* functions are needed to add more posterior-type information.

Each of the transcription units affects successive segments, according to its pattern of expression. So if *Ubx* alone is present, the larva has parasegment 4 (T1P/T2A), parasegment 5 (T2P/T3A), and then 8 copies of parasegment 6 (T3P/A1A). This suggests that the expression of *Ubx* is needed for the compartments anterior to A1A. *Ubx* is also expressed in the more posterior segments, but in the wild type, *abdA* and *AbdB* are also present. If they are removed, the expression of *Ubx* alone in all the posterior segments has the same effect that it usually has in parasegment 6 (T3P/A1A).

The addition of *abdA* to *Ubx* adds the wild-type pattern to parasegments 7, 8, and 9. In other words, *Ubx* plus *abdA* can specify up to compartments A3P/A4A, and in the absence of *AbdB*, this continues to be the default pattern for all the more posterior compartments. The addition of *AbdB* is needed to specify parasegments 10-14.

The general model for the function of the *ANT-C* and *BX-C* complexes is to suppose that additional functions are added to define successive segments proceeding in the posterior direction. It functions by reliance on a *combinatorial pattern* in which the addition of successive gene products confers new specificities. This explains the rule that a loss-of-function mutation in one of the genes of the *ANT-C/BX-C* complexes generally allows the gene on the more anterior side of the mutated gene to determine phenotype, that is, loss-of-function results in homeotic transformation of posterior regions into more anterior phenotypes (884).

Expression of *Ubx* in a more anterior segment than usual should have the opposite effect to a loss-of-function; the segment develops a more anterior phenotype. When this is tested by arranging for *Ubx* to be expressed in the head, the anterior segments are converted to the phenotype of parasegment 6. So lack of expression of *Ubx* causes a homeotic transformation in which posterior segments acquire more anterior phenotypes; and overexpression of *Ubx* causes a homeotic transformation in which anterior segments acquire more posterior phenotypes.

This type of relationship is true generally for the cluster as a whole, and explains the properties of *cis*-acting mutations as well as those in the transcription units. These regulatory mutations cause loss of the protein in part of its domain of expression or

cause additional expression in new domains. So they may have either loss-of-function or gain-of-function phenotypes (or sometimes both). The most common is loss-of-function in an individual compartment. For example, *bx* specifically controls expression of *Ubx* in compartment T3A; a *bx* mutation loses expression of *Ubx* in that compartment, which is therefore transformed to the more anterior type of T2A. This example is typical of the general rule for individual *cis*-acting mutations in the complex; each converts a target compartment *so that it develops as though it were located at the corresponding position in the previous segment.* The order of the *cis*-acting sites of mutation on the chromosomes reflects the order of the compartments in which they function. So the expression of *Ubx* in parasegments 4, 5, 6 is controlled sequentially by *abx* (affects parasegment 5), *bx* (affects T3A), etc.

The presence of only 3 genes within the *BX-C* complex poses two major questions. First, how do the combinations of 3 proteins specify the identity of 10 parasegments? One possibility is that there are quantitative differences in the various regions, allowing for the same sort of varying responses in target genes that we described previously for the combinatorial functioning of the segmentation genes. Second, how do the proteins function in different tissue types? The pattern of expression described above refers generally to the epidermis; the development of other tissues is controlled in a way that is parallel, but not identical. For example, although *Ubx* is expressed in all posterior segments up to A8 in the epidermis, in mesoderm, it is repressed posterior of segment A7. The posterior boundary reflects repression by *abdA*, since in *abdA* mutants, *Ubx* expression extends posterior in the mesoderm.

Why are loci involved in regulating development of the adult insect from the embryonic larva different from genes coding for the everyday proteins of the organism? Is their enormous length necessary to generate the alternative products? Could it be connected with some timing mechanism, determined by how long it takes to transcribe the unit? At a typical rate of transcription, it would take ~100 minutes to transcribe *Antp*, which is a significant proportion of the 22 hour duration of *D. melanogaster* embryogenesis.

Proceeding from anterior to posterior along the embryo, we encounter the changing patterns of expression of the genes of the *ANT-C* and *BC-C* loci. What controls their transcription? As in the case of the segmentation loci, the homeotic loci are controlled *partially by the genes that were expressed at the previous stage of development, and partially by interactions among themselves.* For example, the expression of *Ubx* is changed by mutations in *bicoid, hunchback,* or *Kruppel.* The anterior boundary of expression respects the parasegment border defined by *ftz* and *eve*. The general principle is that all of these regulatory genes function by controlling transcription, either by activating it or by repressing it, and that the gene products may exert specific effects by both qualitative and quantitative combinations.

Reviews

3709. Lewis, E. B. (1985). *Regulation of the genes of the bithorax complex in Drosophila.Drosophila*. Cold Spring Harb Symp Quant Biol 50, 155-164.

References

881. Lewis, E. B. (1978). *A gene complex controlling segmentation in Drosophila.* Nature 276, 565-570.

884. Beachy, P. A., Helfand, S. L., and Hogness, D. S. (1985). *Segmental distribution of bithorax complex proteins during Drosophila development*. Nature 313, 545-551.

885. Karch, F., Weiffenbach, B., Peifer, M., Bender, W., Duncan, I., Celniker, S., Crosby, M., and Lewis, E. B. (1985). *The abdominal region of the bithorax complex*. Cell 43, 81-96.

895. Martin, C. H. et al. (1995). *Complete sequence of the bithorax complex of Drosophila*. Proc. Natl. Acad. Sci. USA 92, 8398-8402.

GRADIENTS, CASCADES, AND SIGNALING PATHWAYS

6.31.22 The homeobox is a common coding motif in homeotic genes

Key Terms

The **homeobox** describes the conserved sequence that is part of the coding region of *D. melanogaster* homeotic genes; it is also found in amphibian and mammalian genes expressed in early embryonic development.

Paralogs are highly similar proteins that are coded by the same genome.

Key Concepts

- The homeobox codes for a 60 amino acid protein domain that is a DNA-binding motif.
- Many Drosophila homeotic and segmentation genes code for transcription factors that use a homeodomain to bind DNA.
- The homeodomain does not fully specify the target DNA site, which is influenced by protein-protein combinatorial interactions.
- Homeodomains are found in important regulators of development in a wide range of organisms.
- A vertebrate Hox cluster contains several genes that have homeoboxes, related to the genes in the ANT-C and BX-C fly loci.
- The Hox genes play roles in vertebrate development that are analogous to the roles of the fly genes.

The three groups of genes that control *D. melanogaster* development – maternal genes, segmentation genes, and homeotic genes – regulate one another and (presumably) target genes that code for structural proteins. Interactions between the regulator genes have been defined by analyses that show defects in expression of one gene in mutants of another. However, we have identified rather few of the structural targets on which these groups of genes act to cause differentiation of individual body parts.

Consistent with the idea that the segmentation genes code for proteins that regulate transcription, the genes of 3 gap loci (*hb*, *Kr*, *kni*) contain zinc finger motifs. As first identified in the transcription factors TFIIIA and Sp1 (see **Figure 22.13**), these motifs are responsible for making contacts with DNA. The products of other loci in the gap class also have DNA-binding motifs; *giant* encodes a protein with a basic zipper motif, and *tailless* encodes a protein that resembles the steroid receptors. This suggests that the general function of gap genes is to function as transcriptional regulators.

Conserved motifs are found in many of the homeotic and segmentation genes. The

most common of the conserved motifs is the **homeobox**, a 180 bp region located near the 3′ end in several segmentation and homeotic genes. There are ~40 genes in *Drosophila* that contain a homeobox, and almost all are known to be involved in developmental regulation. (The homeobox was first identified by its predominance in the homeotic genes, from which it took its name.) The protein sequence coded by the homeobox is called the homeodomain; it is a DNA-binding motif in transcription factors (see **Figure 22.24** in *Molecular Biology 5.22.14 Homeodomains bind related targets in DNA*) (883).

The fly homeodomains fall into several groups. A major group in *Drosophila* consists of the homeotic genes in the *BX-C/ANT-C* complexes; they are called the *Antennapedia* group. Their homeodomains are 70-80% conserved, and usually occur at the C-terminal end of the protein (see **Figure 22.22**). A distinct homeodomain sequence is found in the related genes *engrailed* and *invected;* it has only 45% sequence conservation with the *Antennapedia* group (see **Figure 22.23**). Other types of homeodomain sequences are represented in 2-4 genes each (for review see 360).

Many of the *Drosophila* genes that contain homeoboxes are organized into clusters. Three of the homeotic genes in the *BX-C* cluster have homeoboxes, the *ANT-C* complex contains a group of 5 homeotic genes with homeoboxes, and 4 other genes at *ANT-C* also contain homeoboxes. The homeotic genes at *BX-C* and *ANT-C* are sometimes described under the general heading of *HOM-C* genes (for review see 356).

What is the basic function of the *HOM-C* genes in determining identity on the anterior-posterior axis? We assume that homeodomains with different amino acid sequences recognize different target sequences in DNA. Experiments in which regions have been swapped between different proteins suggest that a major part of the specificity of these proteins rests with the homeodomain. However, the ability to bind to a particular DNA target site may not account entirely for their properties. For example, some of these proteins can either activate or repress transcription in response to the context, that is, their actions depend on the set of other proteins that are bound, not just on recognition of the DNA-binding site (for review see 365).

The similarities between the homeodomains of the more closely related members of the group suggest that they could recognize overlapping patterns of target sites. This would open the way for combinatorial effects that could be based on quantitative as well as qualitative differences, that is, there could be competition between proteins with related homeodomains for the same sites. In some cases, different homeoproteins recognize the same target sites on DNA, which poses a puzzle with regard to defining their specificity of action; we assume that there are subtle differences in DNA-binding yet to be discovered, or there are other interactions, such as protein-protein interactions, that play a role.

The homeobox motif is extensively represented in evolution. A striking extension of the significance of homeoboxes is provided by the discovery that a DNA probe representing the homeobox hybridizes with the genomes of many eukaryotes. Genes containing homeoboxes have been characterized in detail in frog, mouse, and human DNA. The frog and mammalian genes are expressed in early embryogenesis, which strengthens the parallel with the fly genes, and suggests the possibility that genes containing homeoboxes are involved in regulation of embryogenesis in a variety of species.

Genes in mammals (and possibly all animals) that are related to the *HOM-C* group have a striking property: like those of the *BX-C/ANT-C* complexes, they are organized in clusters. The individual mammalian genes are called *Hox* genes. A cluster of *Hox* genes may extend 20-100 kb and contain up to 10 genes. Four *Hox* clusters of genes containing homeoboxes have been characterized in the mouse and human genomes. Their organization is compared with the two large fly clusters in **Figure 31.39** (889).

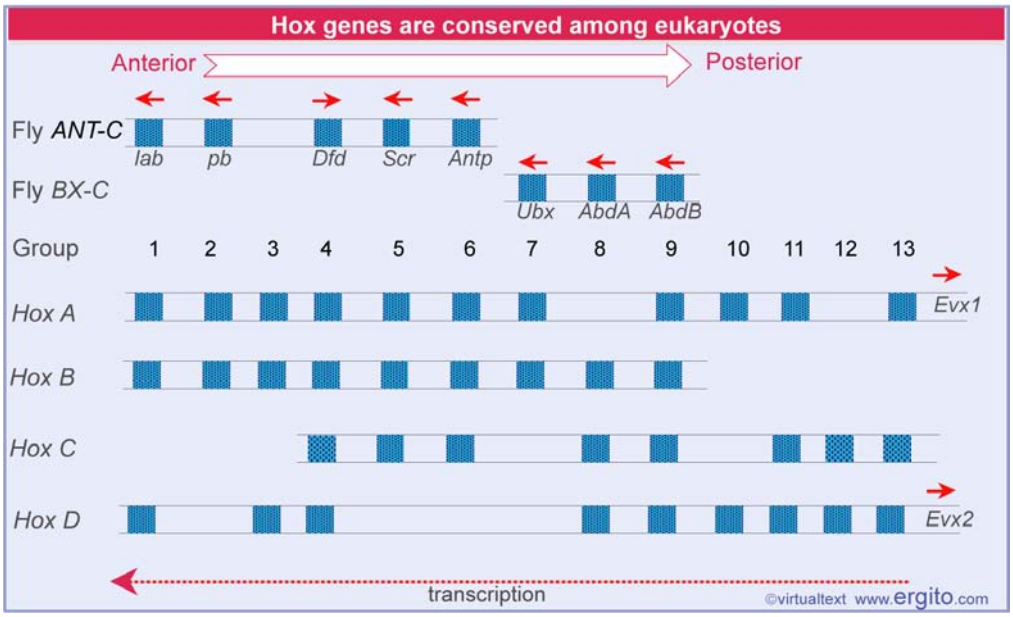

Figure 31.39 Mouse and human genomes each contain 4 clusters of genes that have homeoboxes. The order of genes reflects the regions in which they are expressed on the anterior-posterior axis. The Hox genes are aligned with the fly genes according to homology, which is strong for groups 1, 2, 4, and 9. The genes are named according to the group and the cluster, e.g., HoxA1 is the most anterior gene in the HoxA group. All Hox genes are present in both man and mice except for some mouse genes missing from cluster C.

By comparing the sequences of the homeoboxes (and sometimes other short regions), the mammalian genes can be placed into groups that correspond with the fly genes. This is shown by vertical alignment in the figure. For example, *HoxA4* and *HoxB4* are best related to *Dfd*. When these relationships are defined for the cluster as a whole, it appears that within each cluster we can recognize a series of genes that are related to the genes in the *ANT-C* and *BX-C* clusters. Groups 1-9 in the mammalian loci are defined as corresponding to the genes of the *ANT-C* and *BX-C* loci organized end to end in anterior-posterior orientation. Groups 10-13 appear to have arisen by tandem duplications and divergence of group 9 (the *AbdB* homologue). The corresponding loci in each cluster are sometimes called **paralogs** (for example, *HoxA4* and *HoxB4* are paralogous).

This situation could have arisen if the fly and mammalian loci diverged at a point when there was only a single complex, containing all of the genes that define anterior-posterior polarity. The organism *Amphioxus*, which corresponds to a line of evolution parallel to the vertebrates, has a single *Hox* cluster containing one member of each paralogous group; this appears to be a direct representative of the original cluster. During evolution, the *Drosophila* genes broke into two separate clusters,

while the entire group of mammalian genes became duplicated, some individual members being lost from each complex after the duplication (891).

The parallel between the mouse and fly genes extends to their pattern of spatial expression. The genes within a *Hox* cluster are expressed in the embryo in a manner that matches their organization in the genome. Progressing from the left toward the right end of the cluster drawn in **Figure 31.39**, genes are expressed in the embryo in locations progressively more restricted to the posterior end. The patterns of expression for fly and mouse are compared schematically in **Figure 31.40**. The domain of expression extends strongly to the posterior boundary shown in the figure, and then tails off into more posterior segments (for review see 361; 364).

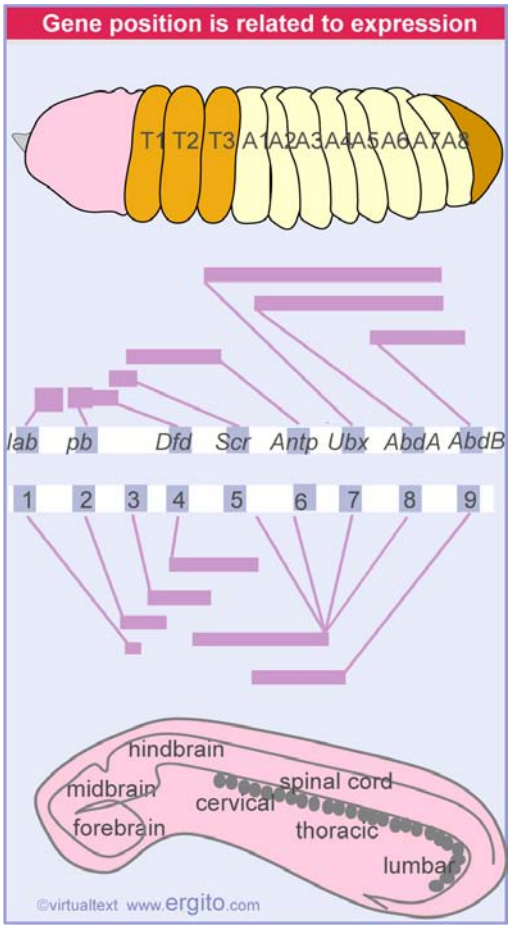

Figure 31.40 A comparison of *ANT-C/BX-C* and HoxB expression patterns shows that the individual gene products share a progressive localization of expression towards the more posterior of the animal proceeding along the gene cluster from left to right. Expression patterns show the regions of transcription in the fly epidermis at 10 hours, and in the central nervous system of the mouse embryo at 12 days.

These results raise the extraordinary possibility that the clusters of genes not only share a common evolution, but also have maintained a common general function in which genome organization is related to spatial expression in fly and mouse, and

there is some correspondence between the homologous genes. The idea of such a relationship is strengthened by the observation that ectopic expression of mouse *HoxD4* or *HoxB6* in *Drosophila* cause homeotic transformations virtually identical to those caused by homeotic expression of *Dfd* or *Antp*, respectively! Since the homology between these mouse and fly proteins rests almost exclusively with their homeodomains, this reinforces the view that these domains determine specificity.

There are some differences in the apparent behavior of the vertebrate Hox clusters and the *ANT-C/BX-C* fly clusters:

- The *Hox* genes are small, and there is a greater number of protein-coding units. The mouse *HoxB* cluster is ~120 kb and contains 9 genes. The connection between genomic position and embryonic expression is analogous to that in *Drosophila*, but describes only the genes themselves; we have no information about *cis*-acting sites. Of course, this may be a consequence of the much greater difficulty in generating mutations in vertebrates. However, our present information identifies the control of *Hox* genes only by promoters and enhancers in the region upstream of the startpoint. It remains to be seen whether there is any counterpart to the very extensive and complex regulatory regions of the *Drosophila* homeotic genes.

- In *Drosophila*, each gene is unique; but in vertebrates, the duplication of the clusters enables multiple genes (paralogs) to have the same or very similar patterns of expression. If the paralogs have redundant or partially redundant functions, so that the absence of one product may be at least partially substituted by the corresponding protein of another cluster, the effects of mutations will be minimized.

Disruptions of *Hox* genes in mice often generate recessive lethals. In the examples of *HoxA1* and *HoxA3* various structures of the head and thorax are absent. Not all of the structures that usually express the mutant genes are missing, suggesting that there is indeed some functional redundancy, that is, other *Hox* genes of group 1 or group 3 can substitute in some but not all other tissues for the absence of the *HoxA* gene.

Homeotic transformations are less common with mutants in mice than in *Drosophila*, but sometimes occur. Loss of *HoxC8*, for example, causes some skeletal segments to show more anterior phenotypes. It remains to be seen whether this is a general rule.

Ectopic expression of *Hox* genes has been used successfully to demonstrate that gain-of-function can transform the identity of a segment towards the identity usually conferred by the gene. The most common type of effect in *Drosophila* is to transform a segment into a phenotype that is usually more posterior; in effect, the expression of the homeotic gene has added additional information that confers a more posterior identity on the segment. Similar effects are observed in some cases in the mouse. However, the pattern is not completely consistent (892).

Taken together, these results make it clear that the *Hox* genes resemble their counterparts in *Drosophila* in determining patterning along the anterior-posterior axis. It may be the case that there is a combinatorial code of *Hox* gene expression, or there may be differences in degree of functional redundancy between paralogs, but we cannot yet provide a systematic model for their role in determining pattern.

The most striking feature of organization of the *Hox* loci still defies explanation: why has the organization of the cluster, in which genomic position correlates with embryonic expression, been maintained in evolution? The obvious explanation is that there is some overall control of gene expression to ensure that it proceeds through the cluster, with the result that a gene could be properly expressed *only* when it is within the cluster. But this does not appear to be true, at least for those individual cases in which genes have been removed from the cluster. Analysis of promoter regions suggests that a *Hox* gene may be controlled by a series of promoter or enhancer elements that together ensure its overall pattern of expression. Usually these elements are in the region upstream of the startpoint. For example, *HoxB4* expression can be reconstructed as the sum of the properties of a series of such elements, tested by introducing appropriate constructs to make transgenic mice. But then why should there have been evolutionary pressure to retain genes in an ordered cluster? One possibility is that an enhancer for one gene might be embedded within another gene, in such a way that, even if an individual gene could function when translocated elsewhere, its removal would impede the expression of other gene(s). An indication that there may be something special about the organization of the region is given by the fact that it has an unusually high density of conserved noncoding sequences and an unusually low density of insertions such as transposons (1442). This suggests the existence of large scale regulatory elements that we have not yet identified.

Last updated on 2-16-2001

Reviews

356. Regulski, M., Harding, K., Kostriken, R., Karch, F., Levine, M., and McGinnis, W. (1985). *Homeo box genes of the Antennapedia and bithorax complexes of Drosophila*. Cell 43, 71-80.

360. Scott, M. P (1989). *The structure and function of the homeodomain*. Biochim. Biophys. Acta 989, 25-48.

361. Hunt, P. and Krumlauf, R. (1992). *Hox codes and positional specification in vertebrate embryonic axes*. Annu. Rev. Cell Biol. 8, 227-256.

364. Krumlauf, R. (1994). *Hox genes in vertebrate development*. Cell 78, 191-201.

365. McGinnis, W. and Krumlauf, R. (1992). *Homeobox genes and axial patterning*. Cell 68, 283-302.

References

883. McGinnis, W. et al. (1984). *A homologous protein-coding sequence in Drosophila homeotic genes and its conservation in other metazoans*. Cell 37, 403-408.

889. Graham, A., Papalopulu, N., and Krumlauf, R. (1989). *The murine and Drosophila homeobox gene complexes have common features of organization and expression*. Cell 57, 367-378.

891. Garcia-Fernandez, J. and Holland, P. W. H. (1994). *Archetypal organization of the amphioxus Hox gene cluster*. Nature 370, 563-566.

892. Malicki, J., Schughart, K., and McGinnis, W. (1990). *Mouse hox-22 specifies thoracic segmental identity in Drosophila embryos and larvae*. Cell 63, 961-967.

1442. Sachidanandam, R. et al. (2001). *A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. The International SNP Map Working Group.* Nature 409, 928-933.

GRADIENTS, CASCADES, AND SIGNALING PATHWAYS

6.31.23 Summary

The development of segments in *Drosophila* occurs by the actions of segmentation genes that delineate successively smaller regions of the embryo. Asymmetry in the distribution of maternal gene products is established by interactions between the oocyte and surrounding cells. This leads to the expression of the gap genes, in 4 broad regions of the embryo. The gap genes in turn control the pair-rule genes, each of which is distributed in 7 stripes; and the pair-rule genes define the pattern of expression of the segment polarity genes, which delineate individual compartments. At each stage of expression, the relevant genes are controlled both by the products of genes that were expressed at the previous stage, and by interactions among themselves. The segmentation genes act upon the homeotic genes, which determine the identities of the individual compartments.

Each of the 4 maternal systems consists of a cascade which generates a locally distributed or locally active morphogen. The morphogen either is a transcription factor or causes the activation of a transcription factor. The transcription factor is the last component in each pathway.

The major anterior-posterior axis is determined by two systems: the anterior system establishes a gradient of bicoid from the anterior pole; and the posterior system produces nanos protein in the posterior half of the egg. These systems function to define a gradient of hunchback protein from the anterior end, with broad bands of knirps and giant in the posterior half. The terminal system acts to produce localized events at both termini. The dorsal-ventral system produces a gradient of nuclear localization of dorsal protein on the ventral side, which represses expression of *dpp* and *zen*; this leads to the ventral activation of *twist* and *snail*, and the dorsal-side activation of *dpp* and *zen*.

Each system is initiated by localization of a morphogen in the egg as a result of its interaction with the surrounding cells. For the anterior and posterior systems, this takes the form of localizing an RNA; *bicoid* mRNA is transported into and localized at the anterior end, and *nanos* mRNA is transported to the posterior end. For the dorsal-ventral system, the Toll receptor is located ubiquitously on the oocyte membrane, but the spatzle ligand is activated ventrally and therefore triggers the pathway on the ventral side. The pathway resembles the mammalian IL-1 signal transduction pathway and culminates in the phosphorylation of cactus, which regulates the dorsal transcription factor. On the dorsal side of the embryo, the morphogen Dpp is released; it is a member of the TGF β family that diffuses to interact with its receptor. A ligand-receptor interaction involving related members of the TGF β/receptor families is also employed in a comparable role in vertebrate development.

The early embryo consists of a syncytium, in which nuclei are exposed to common cytoplasm. It is this feature that allows all 4 maternal systems to control the function of a nucleus according to the coordinates of its position on the anterior-posterior and dorsal-ventral axes. At cellular blastoderm, zygotic RNAs are transcribed, and the developing embryo becomes dependent upon its own genes. Cells form at the

blastoderm stage, after which successive interactions involve a cascade of transcriptional regulators.

Three gap genes are zinc-finger proteins, and one is a basic zipper protein. Their concentrations control expression of the pair-rule genes, which are also transcription factors. In particular, the expression of *eve* and *ftz* controls the boundaries of compartments, functioning in every other segment. The segment polarity genes represent the first step in the developmental cascade that involves functions other than transcription factors. Interactions between the segmentation gene products define unique combinations of gene expression for each segment.

The segment polarity genes include proteins involved in cell-cell interactions as well as transcription factors. The basic circuitry that determines the anterior and posterior polarities of compartments is maintained by an autoregulatory interaction between the cells at the boundary. An anterior compartment secretes wingless protein, which acts upon the cell on the posterior side. This causes engrailed to be expressed in the posterior cell, which in turn causes secretion of hedgehog on the anterior side. Hedgehog causes the anterior cell to express wingless.

Homeotic genes impose the program that determines the unique differentiation of each segment. The complex loci *ANT-C* and *BX-C* each contain a cluster of functions, whose spatial expression on the anterior-posterior axis reflects genetic position in the cluster. Each cluster contains one exceedingly large transcription unit as well as other, shorter units. Many of the transcription units (including the largest genes, *Ubx* and *Antp*) have patterns of alternative splicing, but no significance has been attributed to this yet. Proceeding from left to right in each cluster, genes are expressed in more posterior tissues. The genes are expressed in overlapping patterns in such a way that addition of a function confers new features of posterior identity; thus loss of a function results in a homeotic transformation from posterior to more anterior phenotype. The genes are controlled in a complex manner by a series of regulatory sites that extend over large regions; mutations in these sites are *cis*-acting, and may cause either loss-of-function or gain-of-function. The *cis*-acting mutations tend to act on successive segments of the fly, by controlling expression of the homeotic proteins.

The genes of the *ANT-C* and *BX-C* loci, and many segmentation genes (including the maternal gene *bicoid* and most of the pair-rule genes) contain a conserved motif, the homeobox. Homeoboxes are also found in genes of other eukaryotes, including worms, frogs, and mammals. In each case, these genes are expressed during early embryogenesis. In mammals, the *Hox* genes (which specify homeodomains in the *Antennapedia* class) are organized in clusters. There are 4 *Hox* clusters in both man and mouse. These clusters can be aligned with the *ANT-C/BX-C* clusters in such a way as to recognize homologies between genes at corresponding positions. Proceeding towards the right in a *Hox* cluster, a gene is expressed more towards the posterior of the embryo. The *Hox* genes have roles in conferring identity on segments of the brain and skeleton (and other tissues). The analogous clusters represent regulators of embryogenesis in mammals and flies. *Hox* clusters may be a characteristic of all animals.

Drosophila genes containing homeoboxes form an intricate regulatory network, in which one gene may activate or repress another. The relationship between the sequence of the homeodomain, the DNA target it recognizes, and the regulatory

consequences, remains to be fully elucidated. Specificity in target choice appears to reside largely in the homeodomain; we have yet to explain the abilities of a particular homeoprotein to activate or to repress gene transcription at its various targets. The general principle is that segmentation and homeotic genes act in a transcriptional cascade, in which a series of hierarchical interactions between the regulatory proteins is succeeded by the activation of structural genes coding for body parts.