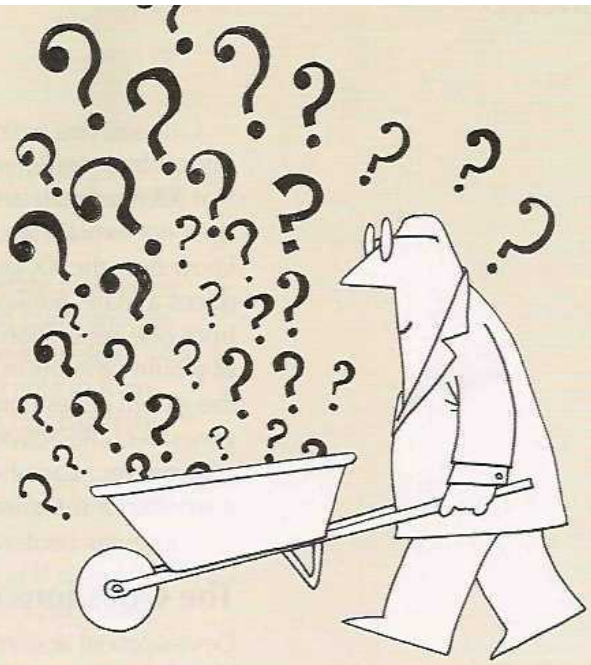
PART I



QUESTIONS

Introducing Developmental Biology

Between fertilization and birth, the developing organism is known as an embryo. The concept of an embryo is a staggering one, and forming an embryo is the hardest thing you will ever do. To become an embryo, you had to build yourself from a single cell. You had to respire before you had lungs, digest before you had a gut, build bones when you were pulpy, and form orderly arrays of neurons before you knew how to think. One of the critical differences between you and a machine is that a machine is never required to function until after it is built. Every animal has to function even as it builds itself.

For animals, fungi, and plants, the sole way of getting from egg to adult is by developing an embryo. The embryo mediates between genotype and phenotype, between the inherited genes and the adult organism. Whereas most fields of biology study adult structure and function, developmental biology finds the study of the transient stages leading up to the adult to be more interesting. Developmental biology studies the initiation and construction of organisms rather than their maintenance. It is a science of becoming, a science of process.

This development, this formation of an orderly body from relatively homogeneous material, provokes profound and fundamental questions that *Homo sapiens* have been asking since the dawn of self-awareness: How does the body form with its head always above its shoulders? Why is the heart on the left side of our body? Why do we have five fingers on each hand and not more or fewer? Why can't we regenerate limbs? How do the sexes develop their different anatomies? Why can only females have babies?

Our answers to these questions must respect the complexity of the inquiry and must form a coherent causal network from gene through functional organ. To say that XX mammals are usually females and that XY mammals are usually males does not explain sex determination to a developmental biologist, who wants to know *how* the XX genotype produces a female and *how* the XY genotype produces a male. Similarly, a geneticist might ask how globin genes are transmitted from one generation to the next, and a physiologist might ask about the function of globin proteins in the body. But the developmental biologist asks how it is that the globin genes come to be expressed only in red blood cells, and how these genes become active only at specific times in development. (We don't know the answers yet.) Each field of biology is defined by the questions it asks. *Welcome to a wonderful set of important questions!*

The Questions of Developmental Biology

Development accomplishes two major objectives. First, it generates cellular diversity and order within the individual organism; secondly, it ensures the continuity of life from one generation to the next. Put another way, there are two fundamental questions in developmental biology. How does the fertilized egg give rise to the adult body? And how does that adult body produce yet another body? These two huge questions can be subdivided into seven general categories of questions scrutinized by developmental biologists:

- **The question of differentiation.** A single cell, the fertilized egg, gives rise to hundreds of different cell types—muscle cells, epidermal cells, neurons, lens cells, lymphocytes, blood cells, fat cells, and so on. The generation of this cellular diversity is called *differentiation*. Since every cell of the body (with very few exceptions) contains the same set of genes, how can this identical set of genetic instructions produce different types of cells? How can a single cell, the fertilized egg, generate so many different cell types?*
- **The question of morphogenesis.** How can the cells in our body organize themselves into functional structures? Our differentiated cells are not randomly distributed. Rather, they become organized into intricate tissues and organs. During development, cells divide, migrate, and die; tissues fold and separate. Our fingers are always at the tips of our hands, never in the middle; our eyes are always in our heads, not in our toes or gut. This creation of ordered form is called *morphogenesis*, and it involves coordinating cell growth, cell migration, and cell death.
- **The question of growth.** If each cell in our face were to undergo just one more cell division, we would be considered horribly malformed. If each cell in our arms underwent just one more round of cell division, we could tie our shoelaces without bending over. How do our cells know when to stop dividing? Our arms are generally the same size on both sides of the body. How is cell division so tightly regulated?
- **The question of reproduction.** The sperm and egg are very specialized cells, and only they can transmit the instructions for making an organism

*There are more than 210 different cell types recognized in the adult human, but this number has little or no significance. There are many transient cell types that are formed during development but are not seen in the adult. Some of these embryonic cells are transitional stages or precursors of adult cell types. Other embryonic cell types perform particular functions in constructing an organ and then undergo programmed cell death after completing their tasks.

from one generation to the next. How are these germ cells set apart from the cells that are constructing the physical structures of the embryo, and what are the instructions in the nucleus and cytoplasm that allow them to form the next generation?

- **The question of regeneration.** Some organisms can regenerate their entire body. Some salamanders regenerate their eyes and legs, and many reptiles can regenerate their tails. Mammals are generally poor at regeneration, and yet there are some cells in our bodies—*stem cells*—that are able to form new structures even in adults. How do the stem cells retain this capacity, and can we harness it to cure debilitating diseases?
- **The question of evolution.** Evolution involves inherited changes in development. When we say that today's one-toed horse had a five-toed ancestor, we are saying that changes in the development of cartilage and muscles occurred over many generations in the embryos of the horse's ancestors. How do changes in development create new body forms? Which heritable changes are possible, given the constraints imposed by the necessity that the organism survive as it develops?
- **The question of environmental integration.** The development of many (perhaps all) organisms is influenced by cues from the environment that surrounds the embryo or larvae. The sex of many species of turtles, for instance, depends on the temperature the embryo experiences while in the egg. The formation of the reproductive system in some insects depends on bacteria that are transmitted inside the egg. Moreover, certain chemicals in the environment can disrupt normal development, causing malformations in the adult. How is the development of an organism integrated into the larger context of its habitat?

The study of development has become essential for understanding all other areas of biology. Indeed, the questions asked by developmental biologists have also become critical in molecular biology, physiology, cell biology, genetics, anatomy, cancer research, neurobiology, immunology, ecology, and evolutionary biology. In turn, the many advances of molecular biology, along with new techniques of cell imaging, have finally made these questions answerable. This makes developmental biologists extremely happy; for, as the Nobel Prize-winning developmental biologist Hans Spemann stated in 1927,

> We stand in the presence of riddles, but not without the hope of solving them. And riddles with the hope of solution—what more can a scientist desire?

So, like the man in the cartoon, I come bearing questions. They are questions bequeathed to us by earlier generations of biologists, philosophers, and parents. They are questions with their own history, questions discussed on an anatomical level by people such as Aristotle, William Harvey, St. Albertus Magnus, and Charles Darwin. More recently, these questions have been addressed on the cellular and molecular levels by men and women throughout the world, each of whom brings to the laboratory his or her own perspectives and training. For there is no one way to become a developmental biologist, and the field has benefitted by having researchers trained in cell biology, genetics, biochemistry, immunology, and even anthropology, engineering, physics, history, and art. You are now invited to become part of a community of question-askers for whom the embryo is a source of both wonder and the most interesting questions in the world.

The next three chapters outline some of the critical framework needed to answer these questions. Chapter 1 discusses *organismal* concepts, including life cycles, the

three germ layers that form the organs, and the migration of cells during development. Chapter 2 concentrates on the *genetic* approach to cell differentiation and outlines the principle of differential gene expression (which explains how different proteins can be made in different cells from the same set of inherited genes). Chapter 3 focuses on the *cellular* approach to morphogenesis, showing how communication between cells is critical for their formation into tissues and organs. Thus, you will be introduced to development at the organismal, genetic, and cellular, and much of the textbook thereafter will show how these levels are integrated to produce the remarkable panoply of animal development.

1 Developmental Anatomy

ACCORDING TO ARISTOTLE, the first embryologist known to history, science begins with wonder: "It is owing to wonder that people began to philosophize, and wonder remains the beginning of knowledge" (Aristotle, *Metaphysics*, ca. 350 BCE). The development of an animal from an egg has been a source of wonder throughout history. The simple procedure of cracking open a chick egg on each successive day of its 3-week incubation period provides a remarkable experience as a thin band of cells is seen to give rise to an entire bird. Aristotle performed this procedure and noted the formation of the major organs. Anyone can wonder at this remarkable—yet commonplace—phenomenon, but it is the scientist seeks to discover how development actually occurs. And rather than dissipating wonder, new understanding increases it.

Multicellular organisms do not spring forth fully formed. Rather, they arise by a relatively slow process of progressive change that we call **development**. In nearly all cases, the development of a multicellular organism begins with a single cell—the fertilized egg, or **zygote**, which divides mitotically to produce all the cells of the body. The study of animal development has traditionally been called **embryology**, after that phase of an organism that exists between fertilization and birth. But development does not stop at birth, or even at adulthood. Most organisms never stop developing. Each day we replace more than a gram of skin cells (the older cells being sloughed off as we move), and our bone marrow sustains the development of millions of new red blood cells every minute of our lives. In addition, some animals can regenerate severed parts, and many species undergo metamorphosis (such as the transformation of a tadpole into a frog, or a caterpillar into a butterfly). Therefore, in recent years it has become customary to speak of **developmental biology** as the discipline that studies embryonic and other developmental processes.

As the introduction to Part I notes, a scientific field is defined by the questions it seeks to answer. Most of the questions in developmental biology have been provided to it by its embryological heritage. We can identify three major approaches to studying embryology:

- Anatomical approaches
- Experimental approaches
- Genetic approaches

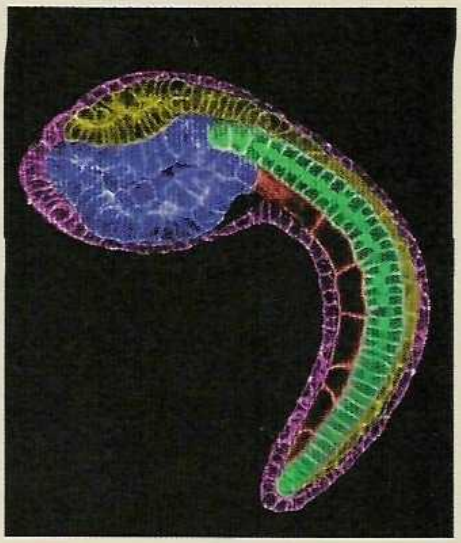
Each of these traditions has predominated during a different era. However, although it is true that anatomical approaches gave rise to experimental approaches, and that genetic approaches built on the foundations of the earlier two approaches, all three traditions persist to this day and continue to play a major role in developmental biology. The basis of all research in developmental biology is the changing anatomy of the organism. Today the anatomical approach to

It is a most beautiful thing to study the different changes of life, from the microscopic changes of conception to the more apparent ones of maturity and old age.

FRANKLIN MALL (CA. 1890)

The greatest progressive minds of embryology have not looked for hypotheses; they have looked at embryos.

JANE OPPENHEIMER (1955)



development is continually expanded and enhanced by revolutions in microscopy, computer-aided graphical reconstructions of three-dimensional objects, and methods of applying mathematics to biology. Many of the beautiful photographs in this book reflect this increasingly important component of embryology.

The Cycle of Life

One of the major triumphs of descriptive embryology was the idea of a generalizable animal life cycle. Each animal, whether earthworm or eagle, termite or beagle, passes through similar stages of development. The stages of development between fertilization and hatching are collectively called **embryogenesis**.

Throughout the animal kingdom, an incredible variety of embryonic types exist, but most patterns of embryogenesis are variations on six fundamental processes: fertilization, cleavage, gastrulation, organogenesis, metamorphosis, and gametogenesis.

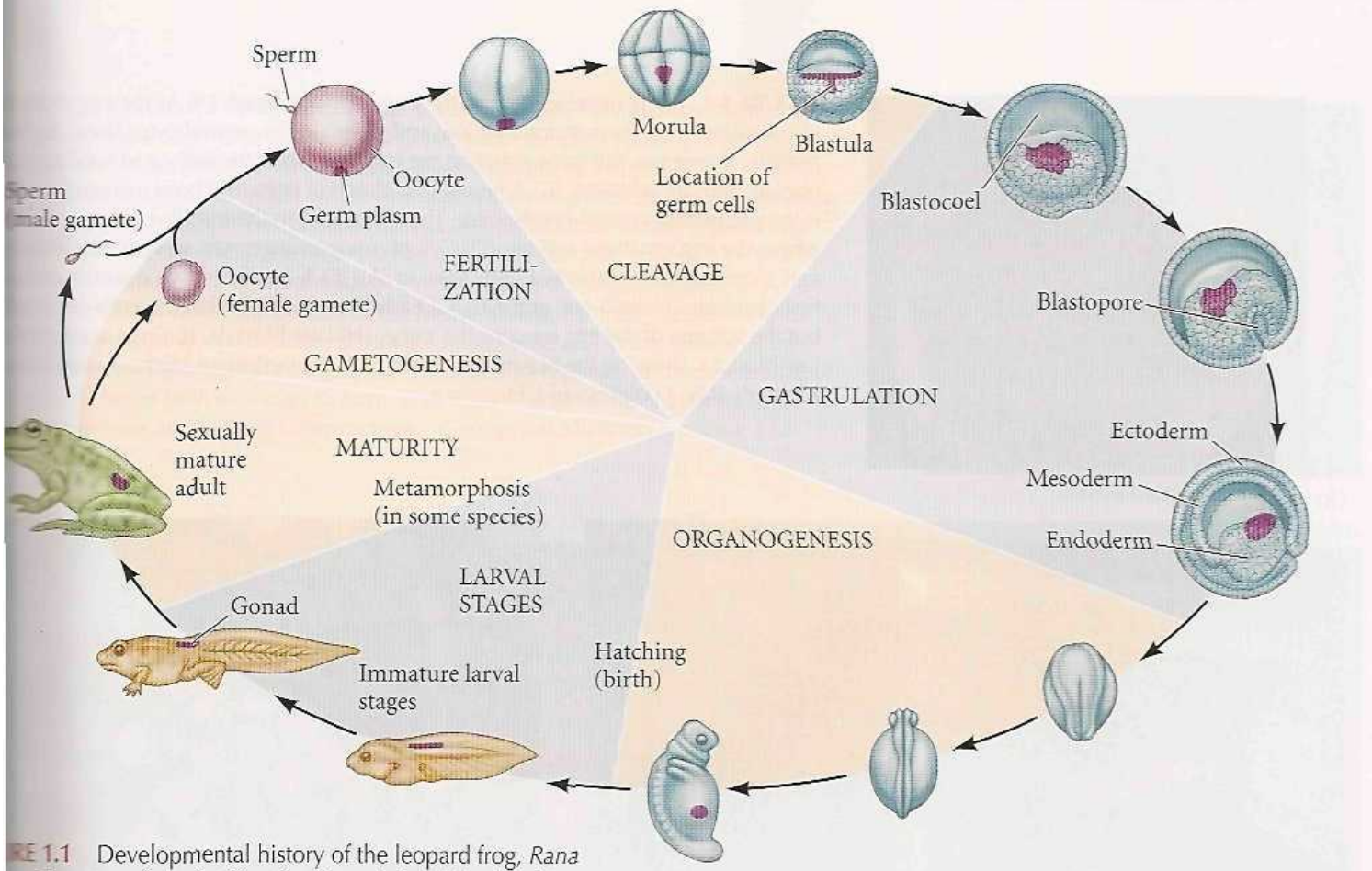
1. **Fertilization** involves the fusion of the mature sex cells, the sperm and egg, which are collectively called the **gametes**. The fusion of the gamete cells stimulates the egg to begin development and initiates a new individual. The subsequent fusion of the gamete nuclei (both of which have only half the normal number of chromosomes characteristic for the species) gives the embryo its **genome**, the collection of genes that helps instruct the embryo to develop in a manner very similar to that of it parents.
2. **Cleavage** is a series of extremely rapid mitotic divisions that immediately follow fertilization. During cleavage, the enormous volume of zygote cytoplasm is divided into numerous smaller cells called **blastomeres**. By the end of cleavage, the blastomeres have usually formed a sphere, known as a **blastula**.
3. After the rate of mitotic division slows down, the blastomeres undergo dramatic movements and change their positions relative to one another. This series of extensive cell rearrangements is called **gastrulation**, and the embryo is said to be in the **gastrula** stage. As a result of gastrulation, the embryo contains three germ layers that will interact to generate the organs of the body.
4. Once the germ layers are established, the cells interact with one another and rearrange themselves to produce tissues and organs. This process is called **organogenesis**. Many organs contain cells from more than one germ layer, and it is not unusual for the outside of an organ to be derived from one layer and the inside from another. For example, the outer layer of skin (epidermis) comes from the ectoderm, whereas the inner layer (the dermis) comes from the mesoderm. Also during organogenesis, certain cells undergo long migrations from their place of origin to their final location. These migrating cells include the precursors of blood cells, lymph cells, pigment cells, and sex cells.
5. In many species, the organism that hatches from the egg or is born into the world is not sexually mature. Rather, the organism needs to undergo **metamorphosis** to become a sexually mature adult. In most animals, the young organism is a called a **larva**, and it may look significantly different from the adult. In many species, the larval stage is the one that lasts the longest, and is used for feeding or dispersal. In such species, the adult is a brief stage whose sole purpose is to reproduce. In silkworm moths, for instance, the adults do not have mouthparts and cannot feed; the larvae must eat enough so that the adult has the stored energy to survive and mate. Indeed, most female moths mate as soon as they eclose from their pupa, and they fly only once—to lay their eggs. Then they die.
6. In many species, a group of cells is set aside to produce the next generation (rather than forming the current embryo). These cells are the precursors of the gametes. The gametes and their precursor cells are collectively called **germ cells**, and they are set aside for reproductive function. All the other cells of the body are called **somatic cells**. This separation of somatic cells (which give rise to the individual body) and germ cells (which contribute to the formation of a new generation) is often one of the first differentiations to occur during animal development. The germ cells eventually migrate to the gonads, where they differentiate into gametes. The development of gametes, called **gametogenesis**, is usually not completed until the organism has become physically mature. At maturity, the gametes may be released and participate in fertilization to begin a new embryo. The adult organism eventually undergoes senescence and dies, its nutrients often supporting the early embryogenesis of its offspring and its absence allowing less competition. Thus, the cycle of life is renewed.

A Frog's Life

All animal life cycles are modifications of the generalized one described above. Figure 1.1 shows the development of the leopard frog, *Rana pipiens*, and provides a good starting point for a more detailed discussion of a representative life cycle.

Gametogenesis and fertilization

The end of one life cycle and the beginning of the next are often intricately intertwined. Life cycles are often controlled by environmental factors (tadpoles wouldn't survive if they hatched in the fall, when their food is dying), so in most frogs, gametogenesis and fertilization are seasonal events. A combination of photoperiod (hours of daylight) and temperature informs the pituitary gland of the mature female frog that it is spring. The pituitary then secretes hormones that stimulate her ovary to make the hormone estrogen. Estrogen then instructs the liver to make and secrete yolk proteins, which are then transported through the



RE 1.1 Developmental history of the leopard frog, *Rana* s. The stages from fertilization through hatching (birth) are collectively as embryogenesis. The region set aside for pro- germ cells is shown in purple. Gametogenesis, which is eted in the sexually mature adult, begins at different times development, depending on the species. (The sizes of the ored wedges shown here are arbitrary and do not corre- to the proportion of the life cycle spent in each stage.)

d into the enlarging eggs in the ovary. The yolk is sported into the bottom portion of the egg, called the **tal hemisphere**, where it will serve as food for the loping embryo (Figure 1.2A). The upper half of the egg led the **animal hemisphere**.* Sperm formation also s on a seasonal basis. Male leopard frogs make sperm g the summer, and by the time they begin hiberna- n the fall they have produced all the sperm that will ailable for the following spring's breeding season. most species of frogs, fertilization is external. The frog grabs the female's back and fertilizes the eggs as male releases them (Figure 1.2B). Some species lay eggs in pond vegetation, and the egg jelly adheres to lants and anchors the eggs (Figure 1.2C). Other species their eggs into the center of the pond without any ort. So the first important thing to remember about ycles is that they are often intimately involved with nmental factors.

se of the terms *animal* and *vegetal* for the upper and lower spheres of the early frog embryo reflect the division rates of s. The upper cells divide rapidly and become actively mobile "animated"), while the yolk-filled cells of the lower half en as being immobile (hence like plants, or "vegetal").

Fertilization accomplishes several things. First, it allows the haploid nucleus of the egg (the **female pronucleus**) to merge with the haploid nucleus of the sperm (the **male pronucleus**) to form the diploid **zygote nucleus**. Second, fertilization causes the cytoplasm of the egg to move such that different parts of the cytoplasm find themselves in new locations (Figure 1.2D). This cytoplasmic migration will be important in determining the three embryonic axes of the frog: anterior-posterior (head-tail), dorsal-ventral (back-belly), and right-left. Third, fertilization activates those molecules necessary to begin cell cleavage and gastrulation (Rugh 1950).

Cleavage and gastrulation

During cleavage, the volume of the frog egg stays the same, but it is divided into tens of thousands of cells (Figure 1.2E–H). The cells in the animal hemisphere of the egg divide faster than those in the vegetal hemisphere, and the cells of the vegetal hemisphere become progressively larger the more vegetal the cytoplasm. Meanwhile, a fluid-filled cavity, the **blastocoel**, forms in the animal hemisphere (Figure 1.2I). This cavity will be important for allowing cell movements to occur during gastrulation.

Gastrulation in the frog begins at a point on the embryo surface roughly 180 degrees opposite the point of sperm entry with the formation of a dimple, called the **blastopore**. This dimple (which will mark the future dorsal side of the embryo) expands to become a ring, and cells migrating through the blastopore become the mesoderm (Figure 1.3A–C). The cells remaining on the outside become the ecto-

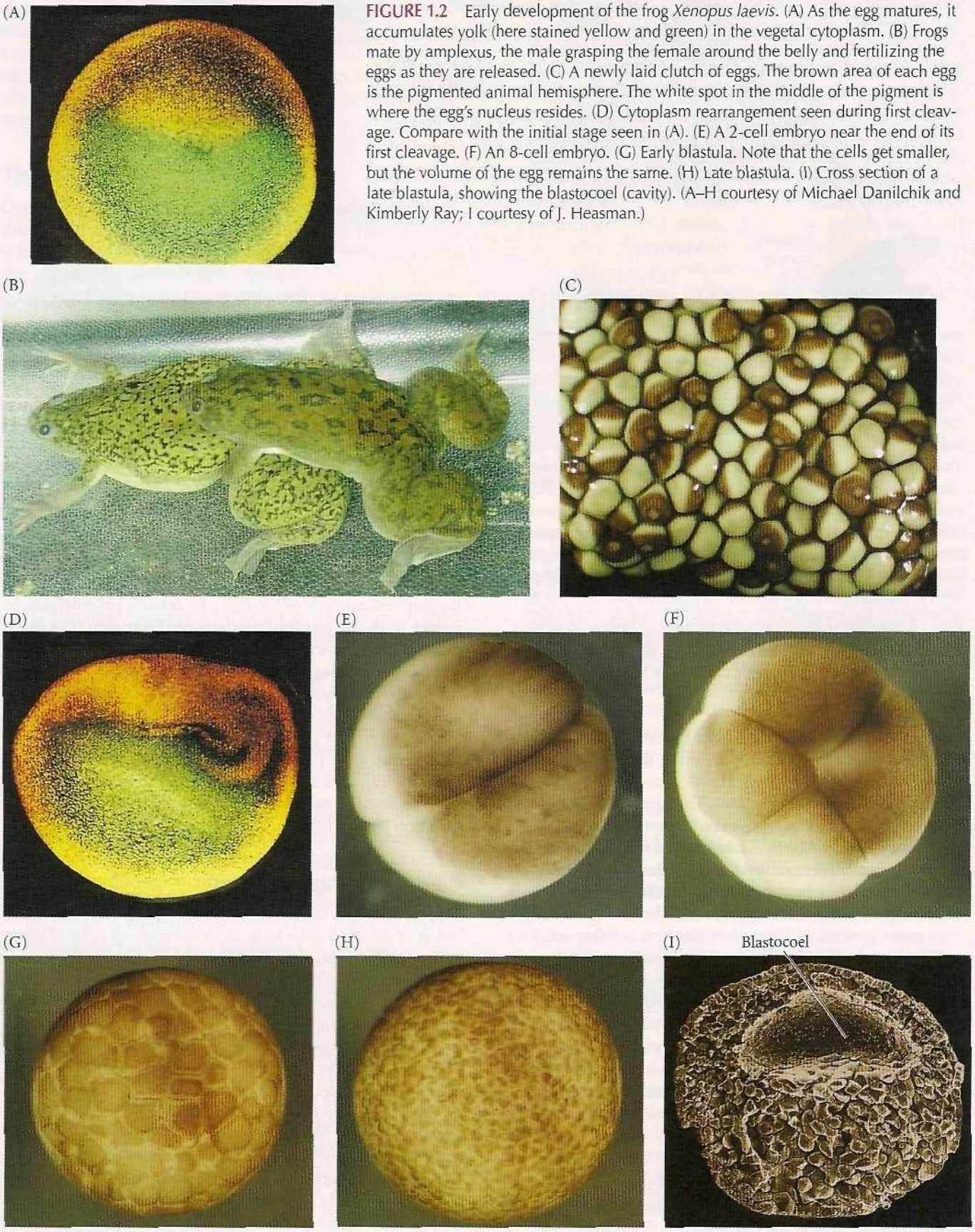
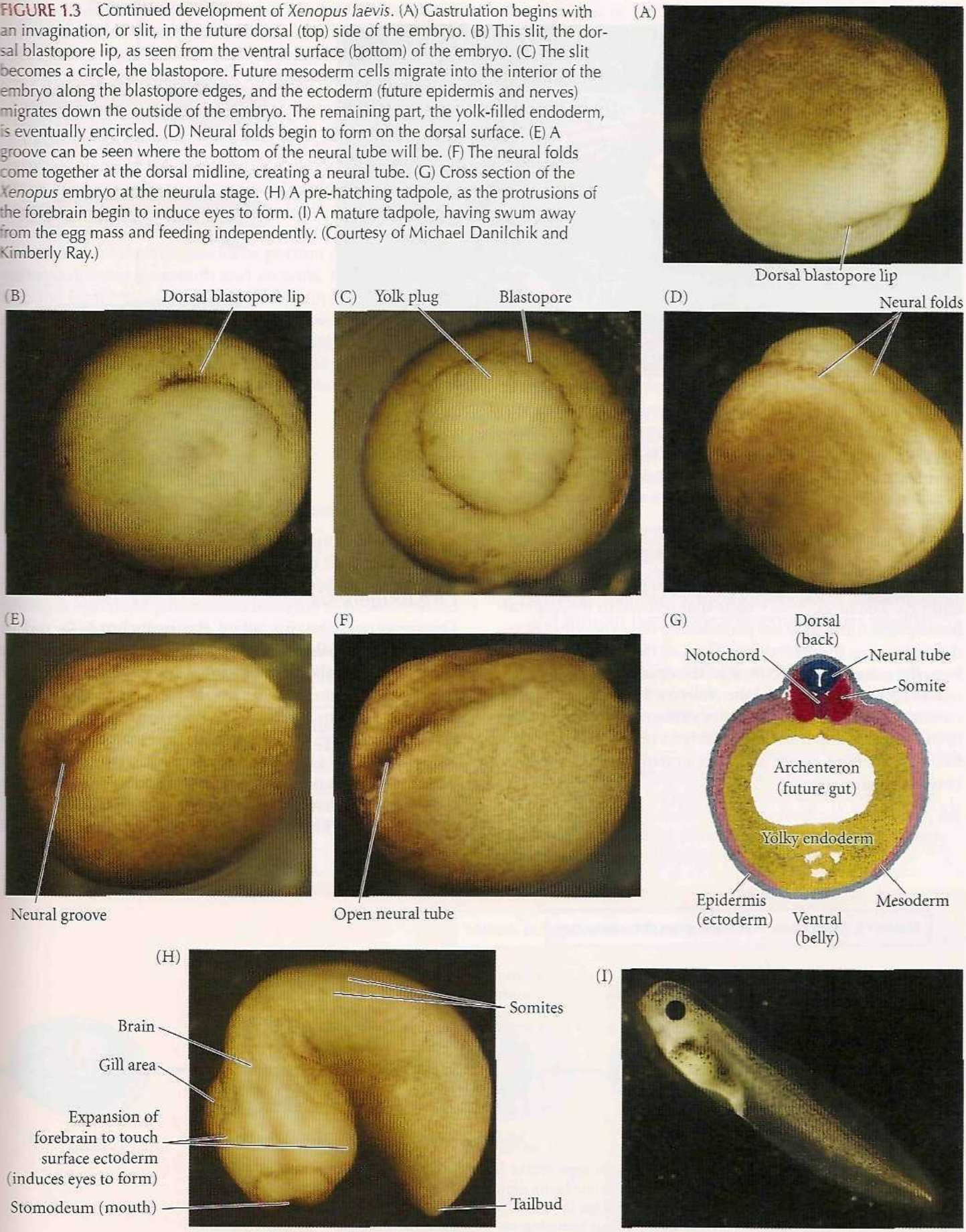


FIGURE 1.2 Early development of the frog *Xenopus laevis*. (A) As the egg matures, it accumulates yolk (here stained yellow and green) in the vegetal cytoplasm. (B) Frogs mate by amplexus, the male grasping the female around the belly and fertilizing the eggs as they are released. (C) A newly laid clutch of eggs. The brown area of each egg is the pigmented animal hemisphere. The white spot in the middle of the pigment is where the egg's nucleus resides. (D) Cytoplasm rearrangement seen during first cleavage. Compare with the initial stage seen in (A). (E) A 2-cell embryo near the end of its first cleavage. (F) An 8-cell embryo. (G) Early blastula. Note that the cells get smaller, but the volume of the egg remains the same. (H) Late blastula. (I) Cross section of a late blastula, showing the blastocoel (cavity). (A–H courtesy of Michael Danilchik and Kimberly Ray; I courtesy of J. Heasman.)

FIGURE 1.3 Continued development of *Xenopus laevis*. (A) Gastrulation begins with an invagination, or slit, in the future dorsal (top) side of the embryo. (B) This slit, the dorsal blastopore lip, as seen from the ventral surface (bottom) of the embryo. (C) The slit becomes a circle, the blastopore. Future mesoderm cells migrate into the interior of the embryo along the blastopore edges, and the ectoderm (future epidermis and nerves) migrates down the outside of the embryo. The remaining part, the yolk-filled endoderm, is eventually encircled. (D) Neural folds begin to form on the dorsal surface. (E) A groove can be seen where the bottom of the neural tube will be. (F) The neural folds come together at the dorsal midline, creating a neural tube. (G) Cross section of the *Xenopus* embryo at the neurula stage. (H) A pre-hatching tadpole, as the protrusions of the forebrain begin to induce eyes to form. (I) A mature tadpole, having swum away from the egg mass and feeding independently. (Courtesy of Michael Danilchik and Kimberly Ray.)



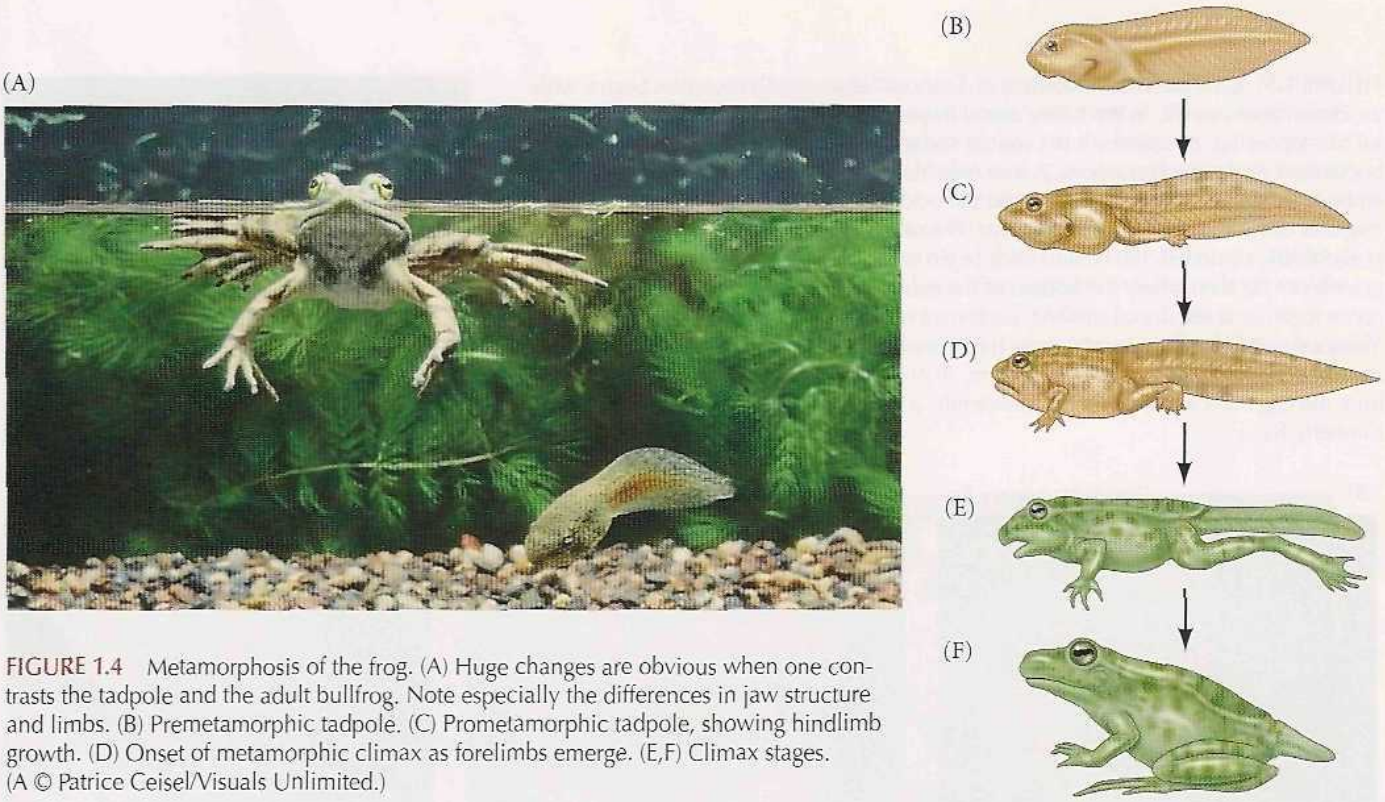


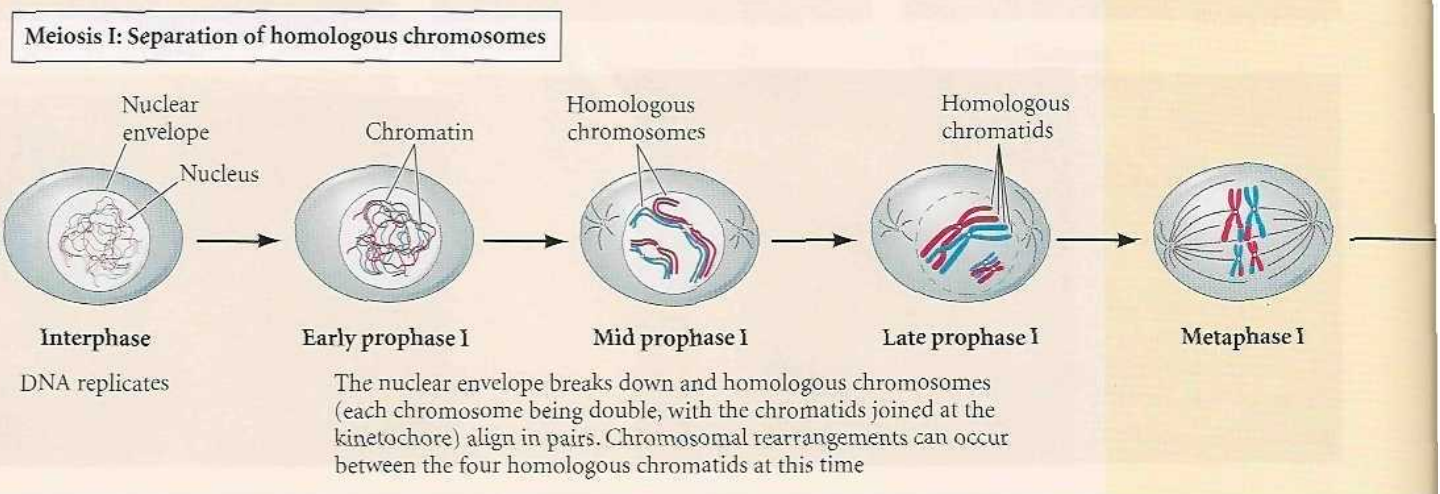
FIGURE 1.4 Metamorphosis of the frog. (A) Huge changes are obvious when one contrasts the tadpole and the adult bullfrog. Note especially the differences in jaw structure and limbs. (B) Premetamorphic tadpole. (C) Prometamorphic tadpole, showing hindlimb growth. (D) Onset of metamorphic climax as forelimbs emerge. (E,F) Climax stages. (A © Patrice Ceisel/Visuals Unlimited.)

derm, and this outer layer expands to enclose the entire embryo. The large, yolky cells that remain in the vegetal hemisphere (until they are encircled by the expanding ectoderm) become the endoderm. Thus, at the end of gastrulation, the ectoderm (precursor of the epidermis, brain, and nerves) is on the outside of the embryo, the endoderm (precursor of the gut and respiratory systems) is on the inside of the embryo, and the mesoderm (precursor of the connective tissue, blood, heart, skeleton, gonads, and kidneys) is between them.

Organogenesis

Organogenesis begins when the **notochord**—a rod of mesodermal cells in the most dorsal portion of the embryo*—signals the ectodermal cells above it that they

*The notochord consists of cells such as those mentioned on p. 2 of the Introduction—i.e., cells that are important for constructing the embryo but which, having performed their tasks, die. Although adult vertebrates do not have notochords, this embryonic organ is critical for establishing the fates of the ectodermal cells above it, as we shall see in Chapters 7–9.



are not going to become epidermis. Instead, these dorsal ectoderm cells form a tube and become the nervous system. At this stage, the embryo is called a **neurula**. The neural precursor cells elongate, stretch, and fold into the embryo, forming the **neural tube** (Figure 1.3D–F); the future epidermal cells of the back cover the neural tube.

Once the neural tube has formed, it and the notochord induce changes in their neighbors, and organogenesis continues. The mesodermal tissue adjacent to the neural tube and notochord becomes segmented into **somites** (Figure 1.3G,H), the precursors of the frog's back muscles, spinal vertebrae, and dermis (the inner portion of the skin). The embryo develops a mouth and an anus, and it elongates into the familiar tadpole structure (Figure 1.3I). The neurons make their connections to the muscles and to other neurons, the gills form, and the larva is ready to hatch from its egg jelly. The hatched tadpole will feed for itself as soon as the yolk supplied by its mother is exhausted.

See VADE MECUM
The amphibian life cycle

Metamorphosis and gametogenesis

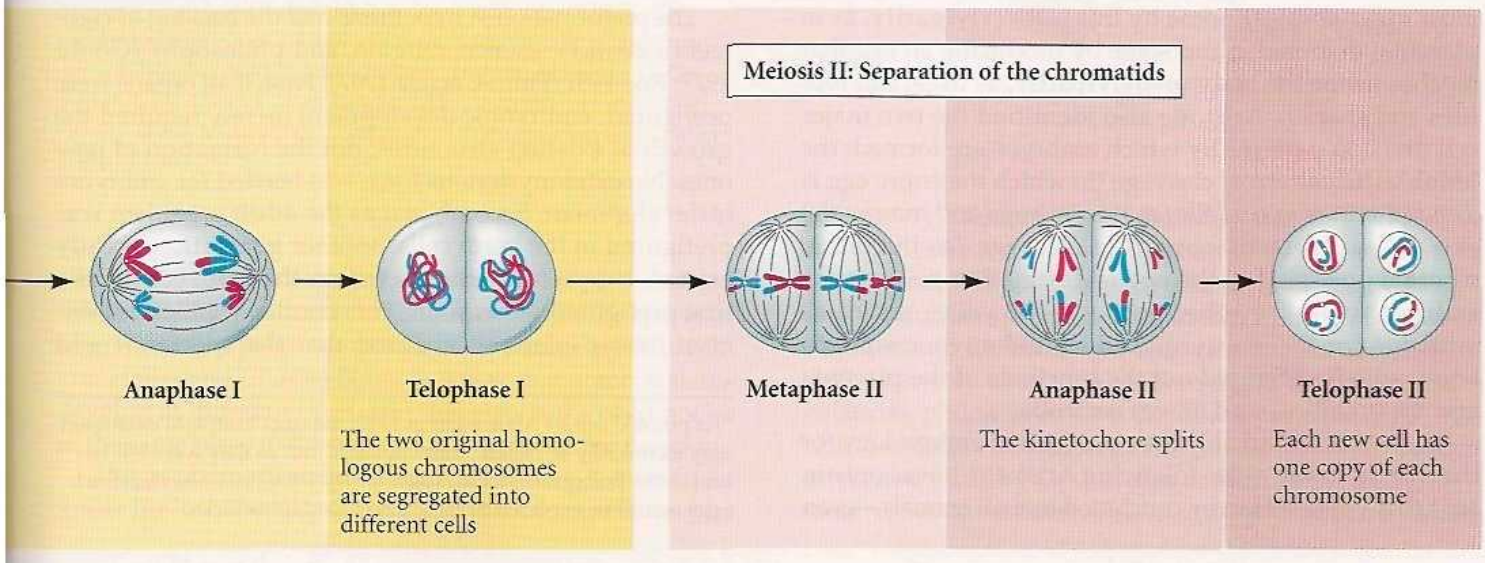
Metamorphosis of the fully aquatic tadpole larva into an adult frog that can live on land is one of the most striking transformations in all of biology. In amphibians, metamorphosis is initiated by hormones from the tadpole's thyroid gland. (The mechanisms by which thyroid hormones accomplish these changes will be discussed in Chapter 15.) In frogs, almost every organ is subject to modification, and the resulting changes in form are striking and very obvious (Figure 1.4). The hindlimbs and forelimbs the adult will use for locomotion differentiate as the tadpole's paddle tail recedes. The cartilaginous tadpole skull is replaced by the predominantly bony skull of the young frog. The horny teeth the tadpole uses to tear up pond plants disappear as the mouth and jaw take a new shape, and the fly-catching tongue muscle of the frog develops. Meanwhile, the tadpole's lengthy intestine—a characteristic of herbivores—shortens to suit the more carnivorous diet of the adult frog. The gills regress and the lungs enlarge. The speed of metamorphosis is carefully keyed to environmental pressures. In temperate regions, for instance, *Rana* metamorphosis must occur before ponds freeze in winter. An adult leopard frog can burrow into the mud and survive the winter; its tadpole cannot.

As metamorphosis ends, the development of the germ cells begins. Gametogenesis can take a long time. In *Rana pipiens*, it takes 3 years for the eggs to mature in the female's ovaries. (Sperm take less time; *Rana* males are often fertile soon after metamorphosis.) To become mature, the germ cells must be competent to complete **meiosis**.

Meiosis (Figure 1.5) is one of the most important evolutionary processes characteristic of eukaryotic organisms. It makes fertilization possible and is critical in recombining genes from the two parents. Genetics, development, and evolution throughout the animal kingdom are predicated on meiosis. We will discuss meiosis more thoroughly in Chapter 16, but the most important things to remember about meiosis are:

1. The chromosomes replicate prior to cell division, so that each gene is represented four times.
2. The replicated chromosomes (each called a chromatid) are held together by their kinetochores (centromeres), and the four homologous chromatids pair together.

FIGURE 1.5 Summary of meiosis. The DNA replicates during interphase. During first meiotic prophase, the nuclear envelope breaks down and the homologous chromosomes (each chromosome is double, with its two chromatids joined at the kinetochore) align together. Chromosome rearrangements ("crossing over") can occur at this stage. After the first metaphase, the kinetochore remains unsplit and the pairs of homologous chromosomes are sorted into different cells. During the second meiotic division, the kintochore splits and the sister chromatids are moved into separate cells, each with a haploid set of chromosomes.



3. The first meiotic division separates the chromatid pairs from one another.
4. The second meiotic division splits the kinetochore such that each chromatid becomes a chromosome.
5. The result is four germ cells, each with a haploid nucleus.

Having undergone meiosis, the mature sperm and egg nuclei can unite in fertilization, restoring the diploid chromosome number and initiating the events that lead to development and the continuation of the circle of life.

"How Are You?"

The fertilized egg has no heart. It has no eye. No limb is found in the zygote. So how did we become what we are? What part of the embryo forms the heart? How do the cells that form the eye's retina migrate the proper distance from the cells that form the lens? How do the tissues that form a bird's wing relate to the tissues that form fish fins or the human hand? What organs are affected by mutations in particular genes? These are the types of questions asked by developmental anatomists.

Several strands weave together to form the anatomical approaches to development. The first strand is **comparative embryology**, the study of how anatomy changes during the development of different organisms. The second strand, based on the first, is **evolutionary embryology**, the study of how changes in development may cause evolutionary change and of how an organism's ancestry may constrain the types of changes that are possible. The third strand of the anatomical approach to developmental biology is **teratology**, the study of birth defects.

Comparative embryology

The first known study of comparative developmental anatomy was undertaken by Aristotle in the fourth century BCE. In *The Generation of Animals* (ca. 350 BCE), he noted some of the variations on the life cycle themes: some animals are born from eggs (**oviparity**, as in birds, frogs, and most invertebrates); some by live birth (**viviparity**, as in placental mammals); and some by producing an egg that hatches inside the body (**ovoviviparity**, as in certain reptiles and sharks). Aristotle also identified the two major cell division patterns by which embryos are formed: the **holoblastic** pattern of cleavage (in which the entire egg is divided into smaller cells, as it is in frogs and mammals) and the **meroblastic** pattern of cleavage (as in chicks, wherein only part of the egg is destined to become the embryo, while the other portion—the yolk—serves as nutrition for the embryo). And should anyone want to know who first figured out the functions of the placenta and the umbilical cord, it was Aristotle.

There was remarkably little progress in embryology for the two thousand years following Aristotle. It was only in 1651 that William Harvey concluded that all animals—even mammals—originate from eggs. *Ex ovo omnia* ("All from the egg") was the motto on the frontispiece of Harvey's *On the Generation of Living Creatures*, and this precluded the spontaneous generation of animals from mud or excrement. This statement was not made lightly, for Harvey knew that it went against the views of Aristotle, whom Harvey still venerated. (Aristotle had thought that menstrual fluid formed the material of the embryo, while the semen gave it form and animation.) Harvey also was the first to see the **blastoderm** of the chick embryo (the small region of the egg containing the yolk-free cytoplasm that gives rise to the embryo), and he was the first to notice that "islands" of blood tissue form before the heart does. Harvey also suggested that the amniotic fluid might function as a "shock absorber" for the embryo.

As might be expected, embryology remained little but speculation until the invention of the microscope allowed detailed observations. In 1672, Marcello Malpighi published the first microscopic account of chick development. Here, for the first time, the neural groove (precursor of the neural tube), the muscle-forming somites, and the first circulation of the arteries and veins—to and from the yolk—were identified (Figure 1.6).

Epigenesis and preformation

With Malpighi begins one of the great debates in embryology: the controversy over whether the organs of the embryo are formed de novo ("from scratch") at each generation, or whether the organs are already present, in miniature form, within the egg (or sperm). The first view, called **epigenesis**, was supported by Aristotle and Harvey. The second view, called **preformation**, was reinvigorated with Malpighi's support. Malpighi showed that the unincubated* chick egg already had a great deal of structure, and this observation provided him with reasons to question epigenesis. According to the preformationist view, all the organs of the adult were prefigured in miniature within the sperm or (more usually) the egg. Organisms were not seen to be "constructed" but rather "unrolled."

The preformationist hypothesis had the backing of eighteenth-century science, religion, and philosophy (Gould 1977; Roe 1981; Pinto-Correia 1997). First, if all organs were prefigured, embryonic development merely required the growth of existing structures, not the formation of new ones. No extra mysterious force was needed for embryonic development. Second, just as the adult organism was prefigured in the germ cells, another generation already existed in a prefigured state within the germ cells of the first prefigured generation. This corollary, called *emboîtment* (encapsulation), ensured that the species would

*As pointed out by Maître-Jan in 1722, the eggs Malpighi examined may technically be called "unincubated," but as they were left sitting in the Bolognese sun in August, they were not unheated. Such eggs would be expected to have developed into chicks.

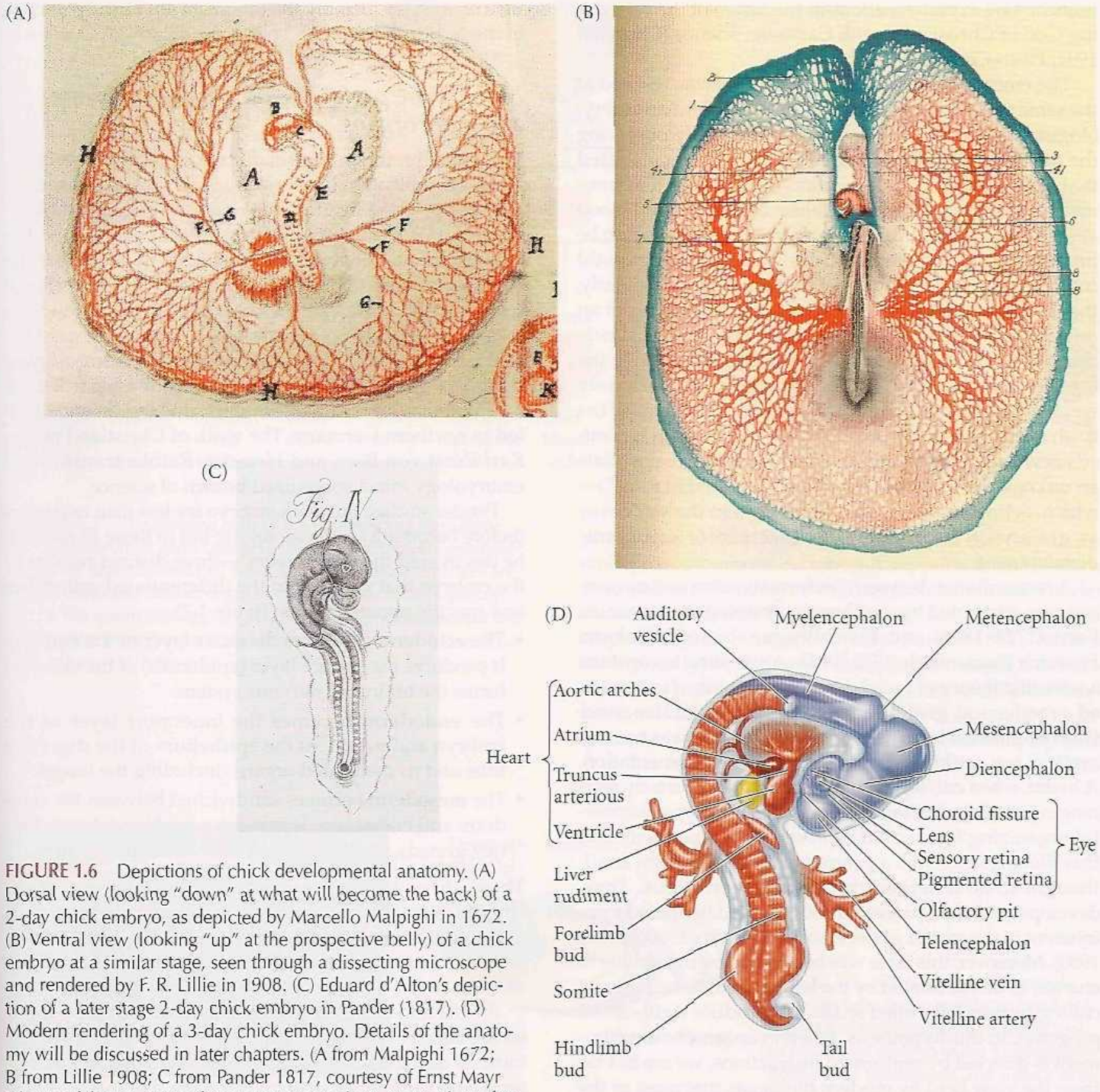


FIGURE 1.6 Depictions of chick developmental anatomy. (A) Dorsal view (looking "down" at what will become the back) of a 2-day chick embryo, as depicted by Marcello Malpighi in 1672. (B) Ventral view (looking "up" at the prospective belly) of a chick embryo at a similar stage, seen through a dissecting microscope and rendered by F. R. Lillie in 1908. (C) Eduard d'Alton's depiction of a later stage 2-day chick embryo in Pander (1817). (D) Modern rendering of a 3-day chick embryo. Details of the anatomy will be discussed in later chapters. (A from Malpighi 1672; B from Lillie 1908; C from Pander 1817, courtesy of Ernst Mayr Library of the Museum of Comparative Zoology, Harvard; D after Carlson 1981.)

remain constant. Although certain microscopists claimed to see fully formed human miniatures within the sperm or egg, the major proponents of this hypothesis—Albrecht von Haller and Charles Bonnet—knew that organ systems develop at different rates, and that structures need not be in the same place in the embryo as they are in the newborn.

The preformationists had no cell theory to provide a lower limit to the size of their preformed organisms (the cell theory arose in the mid-1800s), nor did they view mankind's tenure on Earth as potentially infinite. Rather, said Bonnet (1764), "Nature works as small as it wishes," and the human species existed in that finite time between Creation and Resurrection. This view was in accord with the best science of its time, conforming to the French mathematician-philosopher René Descartes' principle of the infinite divisibility of a mechanical nature initiated, but not interfered with, by God. It also conformed to Enlightenment views of the Deity. The scientist-priest Nicolas Male-

branche saw in preformationism the fusion of the rule-giving God of Christianity with Cartesian science (Churchill 1991; Pinto-Correia 1997).*

The embryological case for epigenesis was revived at the same time by Kaspar Friedrich Wolff, a German embryologist working in St. Petersburg. By carefully observing the development of chick embryos, Wolff demonstrated that the embryonic parts develop from tissues that have no counterpart in the adult organism. The heart and blood vessels (which, according to preformationism, had to be present from the beginning to ensure embryonic growth) could be seen to develop anew in each embryo. Similarly, the intestinal tube was seen to arise by the folding of an originally flat tissue. This latter observation was explicitly detailed by Wolff, who proclaimed in 1767 that "When the formation of the intestine in this manner has been duly weighed, almost no doubt can remain, I believe, of the truth of epigenesis." To explain how an organism is created anew each generation, however, Wolff had to postulate an unknown force—the *vis essentialis* ("essential force")—which, acting according to natural laws in the same way as gravity or magnetism, would organize embryonic development.

A reconciliation between preformationism and epigenesis was attempted by the German philosopher Immanuel Kant (1724–1804) and his colleague, biologist Johann Friedrich Blumenbach (1752–1840). Attempting to construct a scientific theory of racial descent, Blumenbach postulated a mechanical, goal-directed force he called *Bildungstrieb* ("developmental force"). Such a force, he said, was not theoretical, but could be shown to exist by experimentation. A hydra, when cut, regenerates its amputated parts by rearranging existing elements (see Chapter 15). Some purposeful organizing force could be observed in operation, and this *Bildungstrieb* was a property of the organism itself, thought to be inherited through the germ cells. Thus, development could proceed through a predetermined force inherent in the matter of the embryo (Cassirer 1950; Lenoir 1980). Moreover, this force was believed to be susceptible to change, as demonstrated by the left-handed variant of snail coiling (where left-coiled snails can produce right-coiled progeny). In this hypothesis, wherein epigenetic development is directed by preformed instructions, we are not far from the view held by modern biologists that most of the instructions for forming the organism are already present in the fertilized egg.*

*Preformation was a conservative theory, emphasizing the lack of change between generations. Its principal failure was its inability to account for the variations revealed by the limited genetic evidence of the time. It was known, for instance, that matings between white and black parents produced children of intermediate skin color, an impossibility if inheritance and development were solely through either the sperm or the egg. In more controlled experiments, the German botanist Joseph Kölreuter (1766) produced hybrid tobacco plants having the characteristics of both species. Moreover, by mating the hybrid to either the male or female parent, Kölreuter was able to "revert" the hybrid back to one or the other parental type after several generations. Thus, inheritance seemed to arise from a mixture of parental components.

Naming the parts: The primary germ layers and early organs

The end of preformationism did not come until the 1820s, when a combination of new staining techniques, improved microscopes, and institutional reforms in German universities created a revolution in descriptive embryology. The new techniques enabled microscopists to document the epigenesis of anatomical structures, and the institutional reforms provided audiences for these reports and students to carry on the work of their teachers. Among the most talented of this new group of microscopically inclined investigators were three friends, born within a year of each other, all of whom came from the Baltic region and studied in northern Germany. The work of Christian Pander, Karl Ernst von Baer, and Heinrich Rathke transformed embryology into a specialized branch of science.

Pander studied the chick embryo for less than two years (before becoming a paleontologist), but in those 15 months he discovered the **germ layers**†—three distinct regions of the embryo that give rise to the differentiated cells types and specific organ systems (Figure 1.7).

- The **ectoderm** generates the outer layer of the embryo. It produces the surface layer (epidermis) of the skin and forms the brain and nervous system.
- The **endoderm** becomes the innermost layer of the embryo and produces the epithelium of the digestive tube and its associated organs (including the lungs).
- The **mesoderm** becomes sandwiched between the ectoderm and endoderm. It generates the blood, heart, kidney, gonads, bones, muscles, and connective tissues.

These three layers are found in the embryos of all **triploblastic** ("three-layer") animals. Some phyla, such as the poriferans (sponges) and ctenophores (comb jellies), lack a true mesoderm and are considered **diploblastic** animals.

Pander and Rathke also made observations that weighted the balance in favor of epigenesis. Rathke followed the intricate development of the vertebrate skull, excretory systems, and respiratory systems, showing that these became increasingly complex. He also showed that their complexity took on different trajectories in different classes of vertebrates. For instance, Rathke was the first to identify the

*But, as we shall see, not *all* the instructions there. Later in this book, we will see that temperature, diet, predators, symbionts, crowding, and other environmental agents normally regulate gene expression in the embryo and can cause particular phenotypes to occur.

†From the same root as *germination*, the Latin *germen* means "sprout" or "bud." The names of the three germ layers are from the Greek: ectoderm from *ektos* ("outside") plus *derma* ("skin"); mesoderm from *mesos* ("middle"); and endoderm from *endon* ("within").

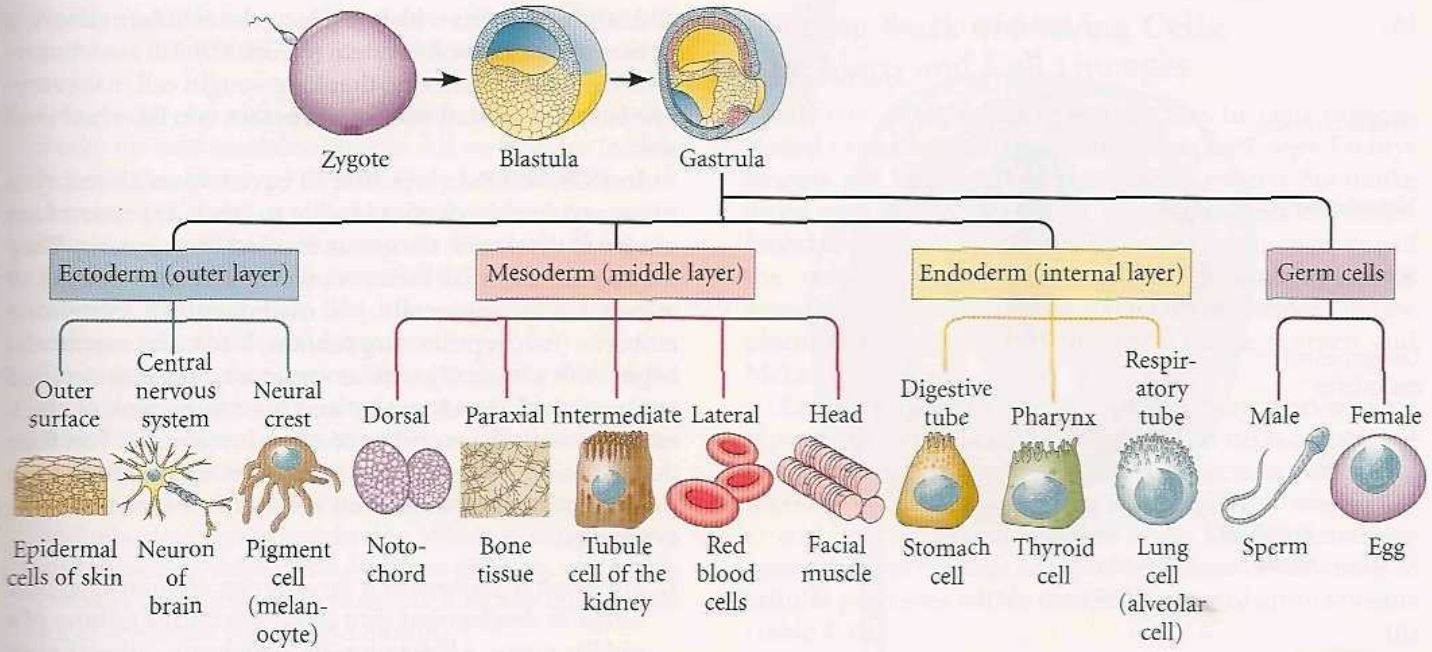


FIGURE 1.7 The dividing cells of the fertilized egg form three distinct embryonic germ layers. Each of the germ layers gives rise to myriad differentiated cell types (only a few representatives are shown here) and distinct organ systems. The germ cells (precursors of the sperm and egg) are set aside early in development and do not arise from any particular germ layer.

pharyngeal arches (Figure 1.8). He showed that these same embryonic structures became gill supports in fish and the jaws and ears (among other things) in mammals. Pander demonstrated that the germ layers did not form their respective organs autonomously (Pander 1817). Rather, each germ layer "is not yet independent enough to indicate what it truly is; it still needs the help of its sister travelers, and therefore, although already designated for different ends, all three influence each other collectively until each has reached an appropriate level." Pander had dis-

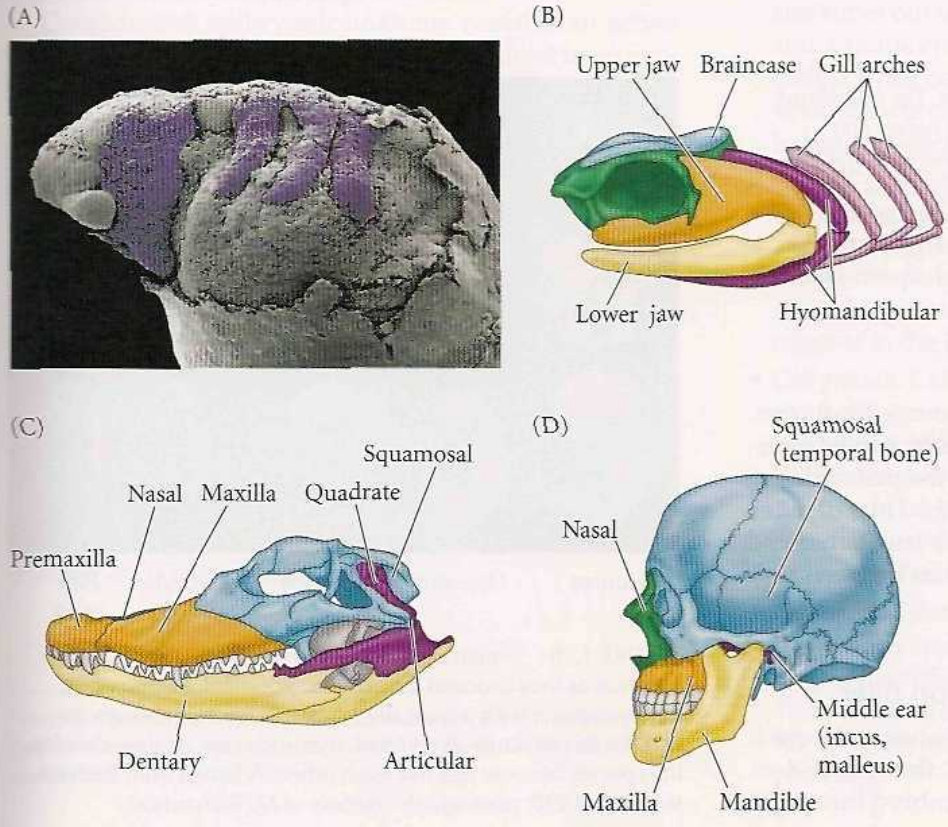


FIGURE 1.8 Evolution of pharyngeal arch structures in the vertebrate head. (A) Pharyngeal arches (also called branchial arches) in the embryo of the salamander *Ambystoma mexicanum*. The surface ectoderm has been removed to permit visualization of the arches (highlighted in color) as they form. (B) In adult fish, pharyngeal arch cells form the hyomandibular jaws and gill arches. (C) In amphibians, birds, and reptiles (a crocodile is shown here), these same cells form the quadrate bone of the upper jaw and the articular bone of the lower jaw. (D) In mammals, the quadrate has become internalized and forms the incus of the middle ear. The articular bone retains its contact with the quadrate, becoming the malleus of the middle ear. (A courtesy of P. Falck and L. Olsson; B–D after Zangerl and Williams 1975.)

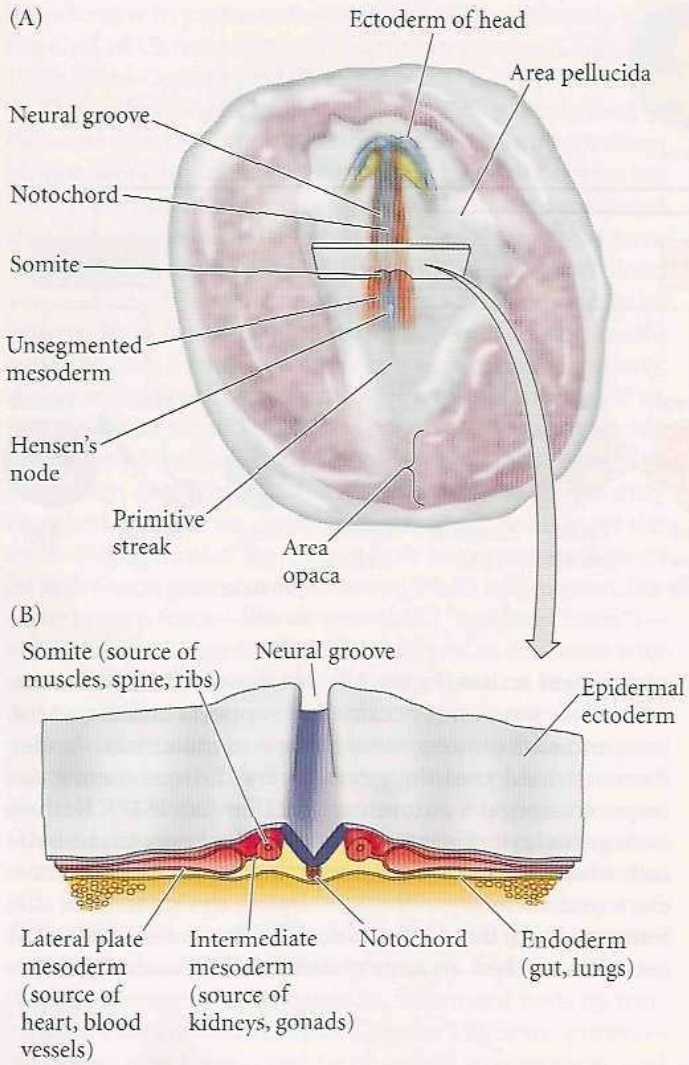


FIGURE 1.9 Notochord in chick development. The notochord separates vertebrate embryos into right and left halves and instructs the ectoderm above it to become the nervous system. (A) Dorsal view of the 24-hour chick embryo. (B) Cross section through the trunk region shows the notochord and developing neural tube. By comparing this illustration and Figure 1.6, you can see the remarkable changes between days 1, 2, and 3 of chick egg incubation. (A after Patten 1951.)

covered the tissue interactions that we now call *induction*. No tissue is able to construct organs by itself; it must interact with other tissues. (We will discuss the principles of induction more thoroughly in Chapter 3.) Thus, Pander showed that preformation could not be true, since the organs come into being through interactions between simpler structures.

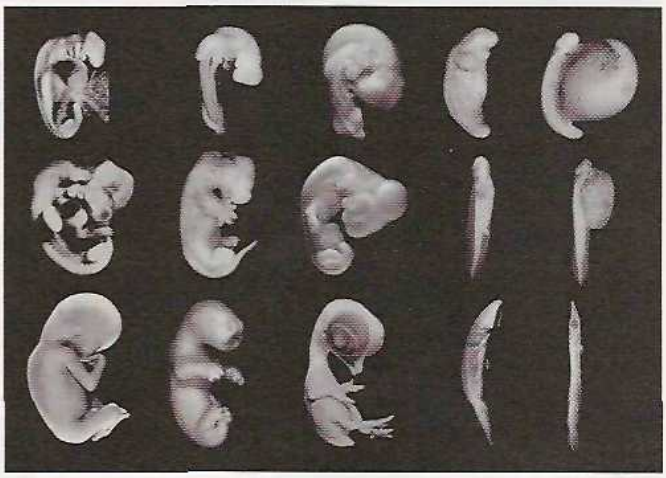
The four principles of Karl Ernst von Baer

Karl Ernst von Baer extended Pander's studies of the chick embryo. He discovered the notochord, the rod of dorsalmost mesoderm that separates the embryo into right and left halves and which instructs the ectoderm above it to become the nervous system (Figure 1.9). He also discovered the mammalian egg, that long-sought cell that everyone believed existed but no one before von Baer had ever seen.*

In 1828, von Baer reported, "I have two small embryos preserved in alcohol, that I forgot to label. At present I am unable to determine the genus to which they belong. They may be lizards, small birds, or even mammals." Figure 1.10 allows us to appreciate his quandary. All vertebrate embryos (fish, reptiles, amphibians, birds, and mammals) begin with a basically similar structure. From his detailed study of chick development and his comparison of chick embryos with the embryos of other vertebrates, von Baer derived four generalizations. Now often referred to as "von Baer's laws," they are stated here with some vertebrate examples.

1. *The general features of a large group of animals appear earlier in development than do the specialized features of a smaller group.* All developing vertebrates appear very similar right after gastrulation. It is only later in development that the special features of class, order, and

*von Baer could hardly believe that he had at last found what so many others—Harvey, de Graaf, von Haller, Prevost, Dumas, and even Purkinje—had searched for and failed to find. "I recoiled as if struck by lightening ... I had to try to relax a while before I could work up enough courage to look again, as I was afraid I had been deluded by a phantom. Is it not strange that a sight which is expected, and indeed hoped for, should be frightening when it eventually materializes?"



Human Opossum Chicken Salamander (axolotl) Fish (gar)

FIGURE 1.10 Similarities and differences among vertebrate embryos as they proceed through development. Each species' embryos begin with a basically similar structure, although they acquire this structure at different ages and sizes. As they develop, the species become less like each other. (Adapted from Richardson et al. 1998; photograph courtesy of M. Richardson.)

finally species emerge. All vertebrate embryos have gill arches, a notochord, a spinal cord, and primitive kidneys.

2. *Less general characters develop from the more general, until finally the most specialized appear.* All vertebrates initially have the same type of skin. Only later does the skin develop fish scales, reptilian scales, bird feathers, or the hair, claws, and nails of mammals. Similarly, the early development of limbs is essentially the same in all vertebrates. Only later do the differences between legs, wings, and arms become apparent.
3. *The embryo of a given species, instead of passing through the adult stages of lower animals, departs more and more from them.** The visceral clefts of embryonic birds and mammals do not resemble the gill slits of adult fish in detail. Rather, they resemble the visceral clefts of *embryonic* fish and other *embryonic* vertebrates. Whereas fish preserve and elaborate these clefts into true gill slits, mammals convert them into structures such as the eustachian tubes (between the ear and mouth).
4. *Therefore, the early embryo of a higher animal is never like a lower animal, but only like its early embryo.* Human embryos never pass through a stage equivalent to an adult fish or bird. Rather, human embryos initially share characteristics in common with fish and avian embryos. Later, the mammalian and other embryos diverge, none of them passing through the stages of the others.

von Baer also recognized that there is a common pattern to all vertebrate development: each of the three germ layers generally gives rise to the same organs, whether the organism itself is a fish, a frog, or a chick.

Comparative embryonic anatomy remains an active field of research today, although it is now done in an evolutionary context. What embryonic interactions, for instance, cover the squirrel's tail with fur but provide scales on the rat's tail? The author's own research concerns how turtles get their shells—a skeletal feature generally composed of 59 bones that no other vertebrate possesses. What is the relationship of these 59 bones to the bones found in alligators and prehistoric marine reptiles? What changes in the "typical" development of the vertebrate skeleton allowed these unique bones to form? Jack Horner and Hans Larsson are looking at the similarities between the developmental anatomy of chick and dinosaur embryos and have found that the embryonic chick, unlike the dinosaur, regresses its tail. They are conducting experiments to block this regression, and actually hope to obtain a chick that more closely resembles its dinosaur ancestors (Horner and Gorman 2009).

*von Baer formulated these generalizations prior to Darwin's theory of evolution. "Lower animals" would be those having simpler anatomies.

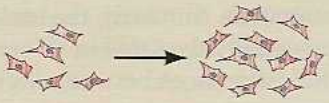
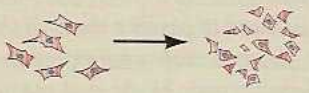
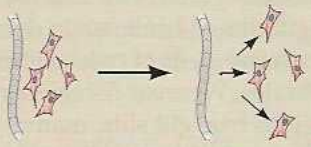
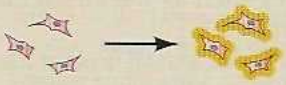
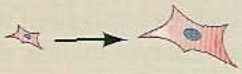
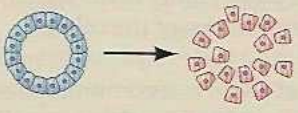
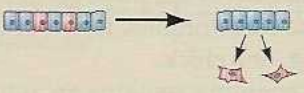
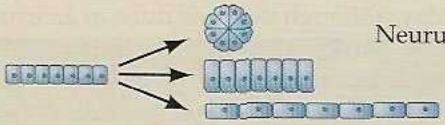
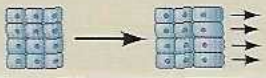
Keeping Track of Moving Cells: Fate Maps and Cell Lineages

By the late 1800s, the cell had been conclusively demonstrated to be the basis for anatomy and physiology. Embryologists, too, began to base their field on the cell. But unlike those who studied the adult, developmental anatomist found that *cells do not stay still in the embryo.* Indeed, one of the most important conclusions of developmental anatomists is that embryonic cells do not remain in one place, nor do they keep the same shape (Larsen and McLaughlin 1987).

Early embryologists recognized that there are two major types of cells in the embryo: **epithelial cells**, which are tightly connected to one another in sheets or tubes; and **mesenchymal cells**, which are unconnected to one another and operate as independent units. Morphogenesis is brought about through a limited repertoire of variations in cellular processes within these two types of arrangements (Table 1.1):

- *Direction and number of cell divisions.* Think of the faces of two dog breeds—say, a German shepherd and a poodle. The faces are made from the same cell types, but the number and orientation of the cell divisions are different. Think also of the legs of a German shepherd compared with those of a dachshund. The skeleton-forming cells of the dachshund have undergone fewer cell divisions than those of taller dogs (see Figure 1.21).
- *Cell shape changes.* Cell shape change is a critical part of not only of development but also of cancer. In development, change in the shapes of epithelial cells often creates tubes out of sheets (as when the neural tube forms); and a shape change from epithelial to mesenchymal is critical when cells migrate away from the epithelium (as when muscle cells are formed). This same type of epithelial-to-mesenchymal change allows cancer cells to migrate and spread from the primary tumor to new sites.
- *Cell movement.* Cell migration is critical to get cells to their appropriate places. The germ cells have to migrate into the developing gonad, and the primordial heart cells meet in the middle of the vertebrate neck and then migrate to the left part of the chest.
- *Cell growth.* Cells can change in size. This is most apparent in the germ cells: the sperm eliminates most of its cytoplasm and becomes smaller, whereas the developing egg conserves and adds cytoplasm, becoming comparatively huge. Many cells undergo an "asymmetric" cell division that produces one big cell and one small cell, each of which may have a completely different fate.
- *Cell death.* Death is a critical part of life. The cells that in the womb constitute the webbing between our toes and fingers die before we are born. So do the cells of our tails. The orifices of our mouth, anus, and reproductive glands all form through cells dying at particular times and places.

TABLE 1.1 Summary of major morphogenic processes regulated by mesenchymal and epithelial cells

Process	Action	Morphology	Example
MESENCHYMAL CELLS			
Condensation	Mesenchyme becomes epithelium		Cartilage mesenchyme
Cell division	Mitosis produces more cells (hyperplasia)		Limb mesenchyme
Cell death	Cells die		Interdigital mesenchyme
Migration	Cells move at particular times and places		Heart mesenchyme
Matrix secretion and degradation	Synthesis or removal of extracellular layer		Cartilage mesenchyme
Growth	Cells get larger (hypertrophy)		Fat cells
EPITHELIAL CELLS			
Dispersal	Epithelium becomes mesenchyme (entire structure)		Müllerian duct degeneration
Delamination	Epithelium becomes mesenchyme (part of structure)		Chick hypoblast
Shape change or growth	Cells remain attached as morphology is altered		Neurulation
Cell migration (intercalation)	Rows of epithelia merge to form fewer rows		Vertebrate gastrulation
Cell division	Mitosis within row or column		Vertebrate gastrulation
Matrix secretion and degradation	Synthesis or removal of extracellular matrix		Vertebrate organ formation
Migration	Formation of free edges		Chick ectoderm

- *Changes in the composition of the cell membrane or secreted products.* Cell membranes and secreted cell products influence the behavior of neighboring cells. For instance, extracellular matrices secreted by one set of cells will allow the migration of their neighboring cells. Extracellular matrices made by other cell types will *prohibit* the migration of the same set of cells. In this way, "paths and guiderails" are established for migrating cells.

Fate maps

Given such a dynamic situation, one of the most important programs of descriptive embryology became the tracing of **cell lineages**: following individual cells to see what those cells become. In many organisms, resolution of individual cells is not possible, but one can label *groups* of embryonic cells to see what that area becomes in the adult organism. By bringing such studies together, one can construct a **fate map**. These diagrams "map" larval or adult structures onto the region of the embryo from which they arose. Fate maps constitute an important foundation for experimental embryology, providing researchers with information on which portions of the embryo normally become which larval or adult structures. Figure 1.11 shows fate maps of some vertebrate embryos at the early gastrula stage.

Fate maps can be generated in several ways, and the technology has changed greatly over the past few years. Construction of these maps is an ongoing research program, and some of the results can be controversial. Recent examples of controversial fate maps include the map for the region of the frog embryo that specifies heart and blood cell precursors (Lane and Sheets 2006), and that for the region of the embryonic turtle that becomes the bones of the plastron (Cebra-Thomas et al. 2007). As we will see later in this book, researchers are currently constructing, refining, and arguing about the fate maps of mammalian embryos.

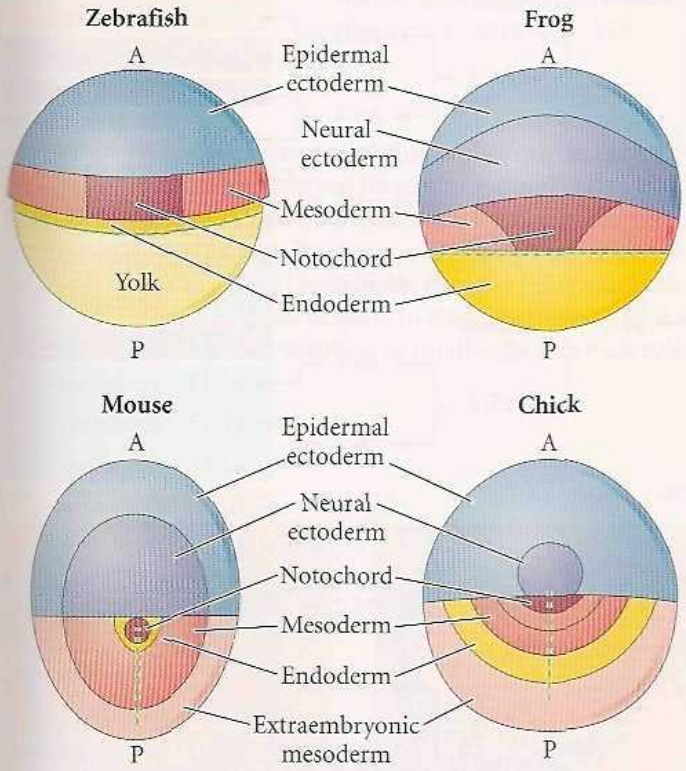


FIGURE 1.11 Fate maps of vertebrates at the early gastrula stage. All are dorsal surface views (looking "down" on the embryo at what will become its back). Despite the different appearances of the adult animals, fate maps of these four vertebrates show numerous similarities among the embryos. The cells that will form the notochord occupy a central dorsal position, while the precursors of the neural system lie immediately anterior to it. The neural ectoderm is surrounded by less dorsal ectoderm, which will form the epidermis of the skin. A indicates the anterior end of the embryo, P the posterior end. The dashed green lines indicate the site of ingression—the path cells will follow as they migrate from the exterior to the interior of the embryo.

Direct observation of living embryos

Some embryos have relatively few cells, and the cytoplasm in each of the early blastomeres has a different pigment. In such fortunate cases, it is actually possible to look through the microscope and trace the descendants of a particular cell into the organs they generate. E. G. Conklin patiently followed the fates of each early cell of *Styela partita*, a tunicate (sea squirt) that resides in waters off the coast of Massachusetts (Conklin 1905). The muscle-forming cells of the embryo always had a yellow color, derived from a region of cytoplasm found in the B4.1 blastomere (Figure 1.12). Conklin's fate map was confirmed by cell-removal experiments. Removal of the B4.1 cell (which according to Conklin's map should produce all the tail musculature) in fact resulted in a larva with no tail muscles (Reverberi and Minganti 1946).

See WEBSITE 1.1 Conklin's art and science

See VADE MECUM The compound microscope

Dye marking

Most embryos are not so accommodating as to have cells of different colors. In the early years of the twentieth century, Vogt (1929) traced the fates of different areas of amphibian eggs by applying **vital dyes** to the region of interest. Vital dyes stain cells but do not kill them. Vogt mixed such dyes with agar and spread the agar on a microscope slide to dry. The ends of the dyed agar were very thin. Vogt cut chips from these ends and placed them on a frog embryo. After the dye stained the cells, he removed the agar chips and could follow the stained cells' movements within the embryo (Figure 1.13).

One problem with vital dyes is that as they become more diluted with each cell division, they become difficult to detect. One way around this is to use **fluorescent dyes** that are so intense that once injected into individual cells, they can still be detected in the progeny of these cells many divisions later. Fluorescein-conjugated dextran, for example, can be injected into a single cell of an early embryo, and the descendants of that cell can be seen by examining the embryo under ultraviolet light (Figure 1.14).

See VADE MECUM
Histotechniques

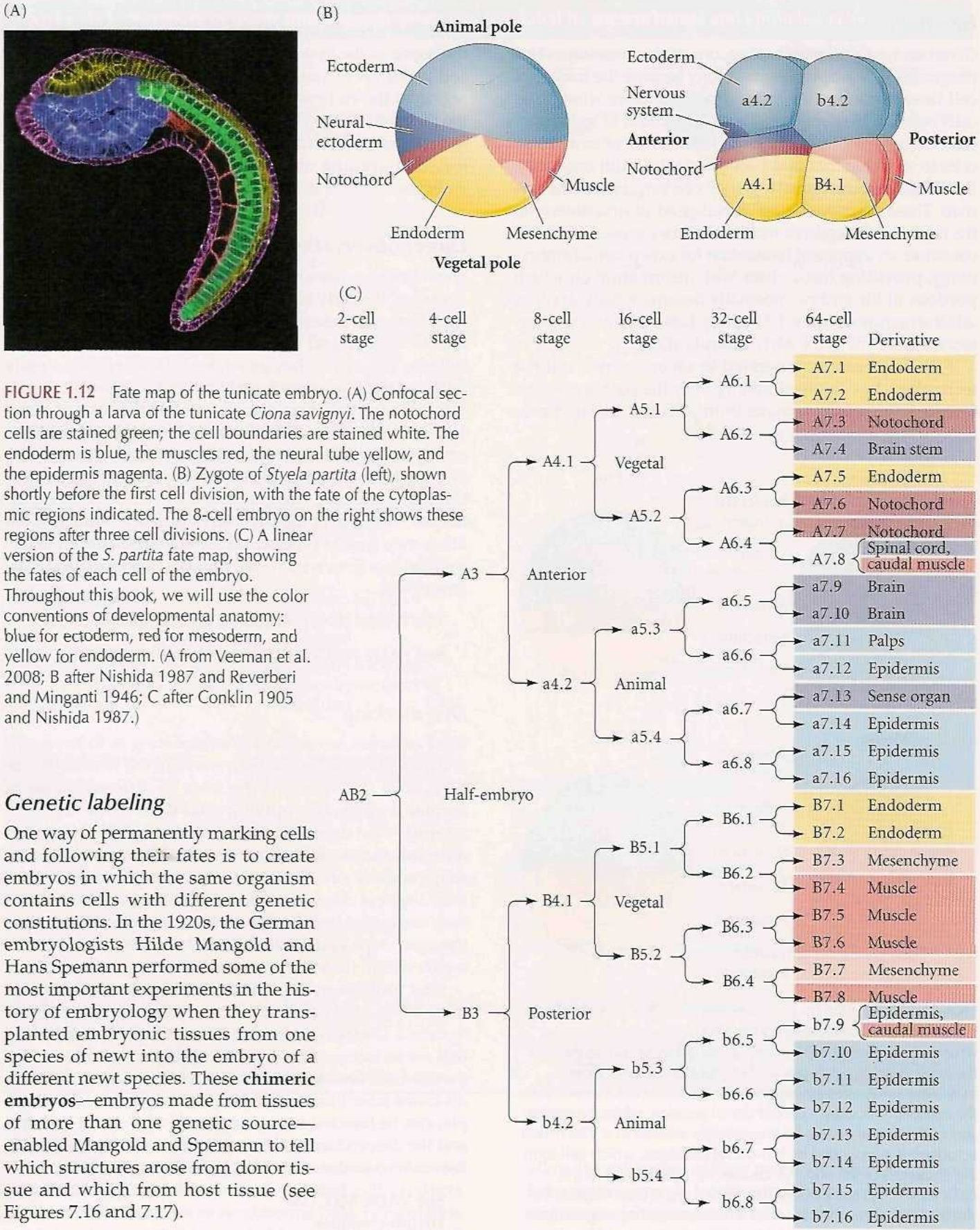


FIGURE 1.12 Fate map of the tunicate embryo. (A) Confocal section through a larva of the tunicate *Ciona savignyi*. The notochord cells are stained green; the cell boundaries are stained white. The endoderm is blue, the muscles red, the neural tube yellow, and the epidermis magenta. (B) Zygote of *Styela partita* (left), shown shortly before the first cell division, with the fate of the cytoplasmic regions indicated. The 8-cell embryo on the right shows these regions after three cell divisions. (C) A linear version of the *S. partita* fate map, showing the fates of each cell of the embryo. Throughout this book, we will use the color conventions of developmental anatomy: blue for ectoderm, red for mesoderm, and yellow for endoderm. (A from Veeman et al. 2008; B after Nishida 1987 and Reverberi and Minganti 1946; C after Conklin 1905 and Nishida 1987.)

Genetic labeling

One way of permanently marking cells and following their fates is to create embryos in which the same organism contains cells with different genetic constitutions. In the 1920s, the German embryologists Hilde Mangold and Hans Spemann performed some of the most important experiments in the history of embryology when they transplanted embryonic tissues from one species of newt into the embryo of a different newt species. These **chimeric embryos**—embryos made from tissues of more than one genetic source—enabled Mangold and Spemann to tell which structures arose from donor tissue and which from host tissue (see Figures 7.16 and 7.17).

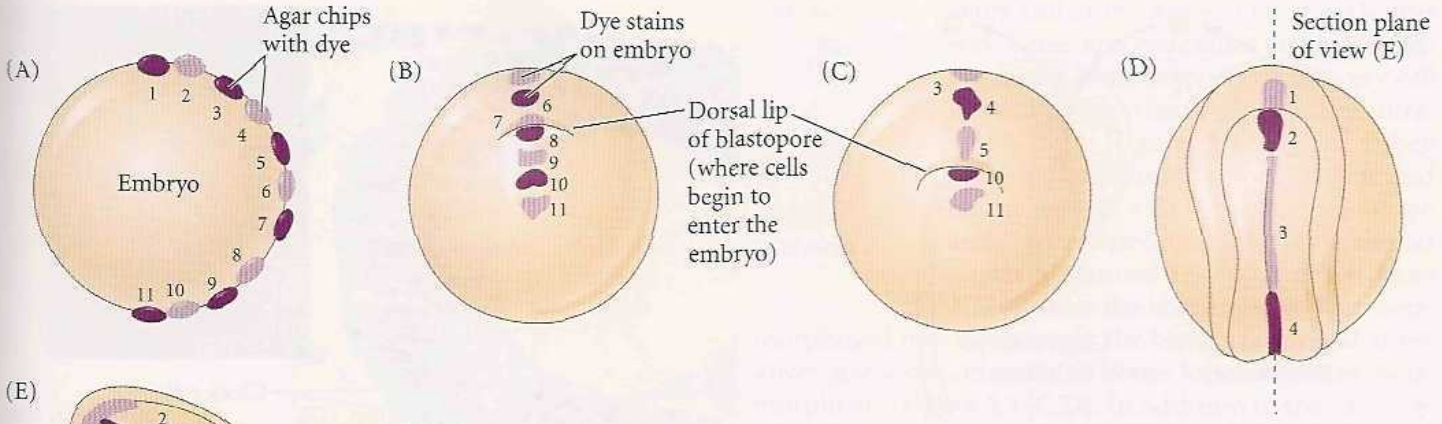


FIGURE 1.13 Vital dye staining of amphibian embryos. (A) Vogt's method for marking specific cells of the embryonic surface with vital dyes. (B–D) Dorsal surface views of stain on successively later embryos. (E) Newt embryo dissected in a medial sagittal section to show the stained cells in the interior. (After Vogt 1929.)

One of the best examples of this technique is the construction of chimeric embryos by grafting quail cells inside a chick embryo. Chicks and quail embryos develop in a similar manner (especially during the early stages development), and the grafted quail cells become integrated into the chick embryo and participate in the construction of the various organs. The substitution of quail cells for chick cells can be performed on an embryo while it is still inside the egg, and the chick that hatches will have quail cells in particular sites, depending on where the graft was placed. Quail cells differ from chick cells in two important ways. First, the quail nucleus has condensed DNA (*heterochromatin*) concentrated around the nucleoli, making quail nuclei easily distinguishable from chick nuclei. Second, cell-specific antigens that are quail-specific can be used to find individual quail cells, even if they are "hidden" within a large population of chick cells. In this way, fine-structure maps of the chick brain and skeletal system have been produced (Figure 1.15; Le Douarin 1969; Le Douarin and Teillet 1973).

In addition, the chick-quail chimeras dramatically confirmed the extensive cell migrations taken by neural crest cells during vertebrate development. Mary Rawles (1940) showed that the pigment cells (**melanocytes**) of the chick originate in the **neural crest**, a transient band of cells that

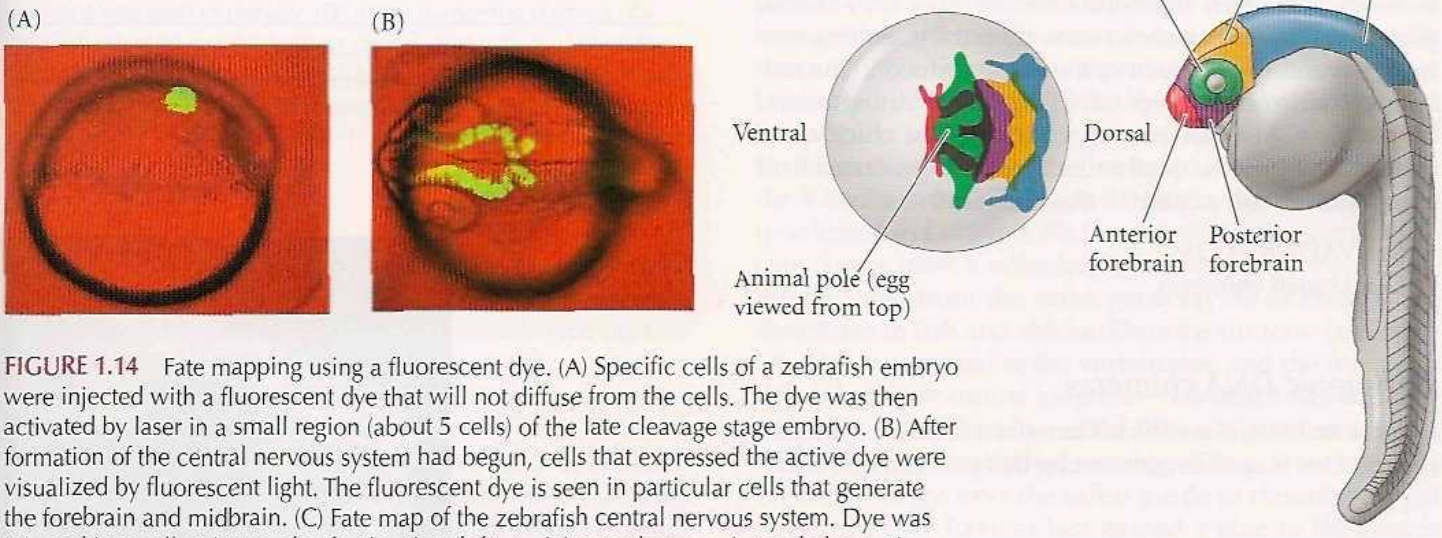
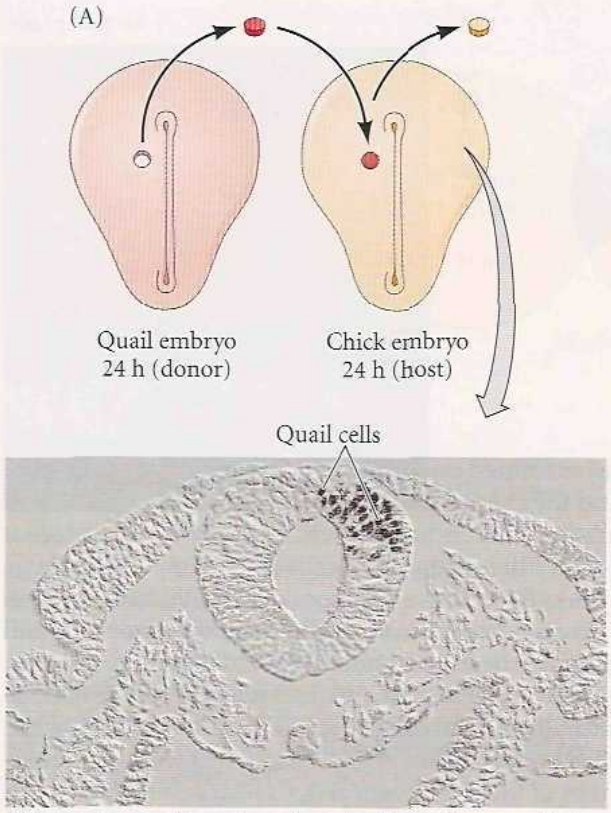


FIGURE 1.14 Fate mapping using a fluorescent dye. (A) Specific cells of a zebrafish embryo were injected with a fluorescent dye that will not diffuse from the cells. The dye was then activated by laser in a small region (about 5 cells) of the late cleavage stage embryo. (B) After formation of the central nervous system had begun, cells that expressed the active dye were visualized by fluorescent light. The fluorescent dye is seen in particular cells that generate the forebrain and midbrain. (C) Fate map of the zebrafish central nervous system. Dye was injected into cells 6 hours after fertilization (left), and the results are color-coded onto the hatched fish (right). Overlapping colors indicate that cells from these regions of the 6-hour embryo contribute to two or more regions. (A,B from Kozlowski et al. 1998, photographs courtesy of E. Weinberg; C after Woo and Fraser 1995.)



Chick embryo with region of quail cells on the neural tube

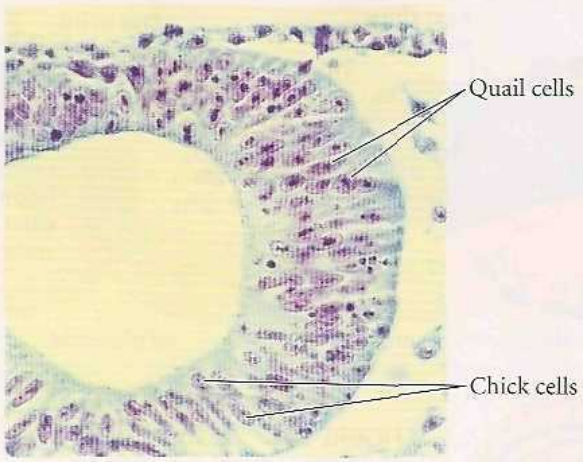


FIGURE 1.15 Genetic markers as cell lineage tracers. (A) Grafting experiment wherein the cells from a particular region of a 1-day quail embryo have been placed into a similar region of a 1-day chick embryo. After several days, the quail cells can be seen by using an antibody to quail-specific proteins. This region of the 3-day embryo produces cells that populate the neural tube. (B) Chick and quail cells can also be distinguished by the heterochromatin of their nuclei. The quail cells have a single large nucleolus (dense purple), distinguishing them from the diffuse nuclei of the chick. (From Darnell and Schoenwolf 1997, courtesy of the authors.)

joins the neural tube to the epidermis. When she transplanted small regions of neural crest-containing tissue from a pigmented strain of chickens into a similar position in an embryo from an unpigmented strain of chickens, the migrating pigment cells entered the epidermis and later entered the feathers (Figure 1.16). Ris (1941) used similar techniques to show that while almost all of the external pigment of the chick embryo came from the migrating neural crest cells, the pigment of the retina formed in the retina itself and was not dependent on the migrating neural crest cells. This pattern was confirmed in chick-quail hybrids, in which the quail neural crest cells produced their own pigment and pattern in the chick feathers.

See VADE MECUM
Chick-quail chimeras

Transgenic DNA chimeras

In most animals, it is difficult to meld a chimera from two species. One way of circumventing this problem is to transplant cells from a genetically modified organism. In such a technique, the genetic modification can then be traced only to those cells that express it. One version is to infect the cells of an embryo with a virus whose genes have been altered such that they express the gene for a fluorescently active protein such as GFP.* (This type of altered gene is

*GFP—green fluorescent protein—occurs naturally in certain jellyfish. It emits bright green fluorescence when exposed to ultraviolet light and is widely used as a transgenic label for cells in developmental and other research.

FIGURE 1.16 Chick resulting from transplantation of a trunk neural crest region from an embryo of a pigmented strain of chickens into the same region of an embryo of an unpigmented strain. The neural crest cells that gave rise to the pigment migrated into the wing epidermis and feathers. (From the archives of B. H. Willier.)

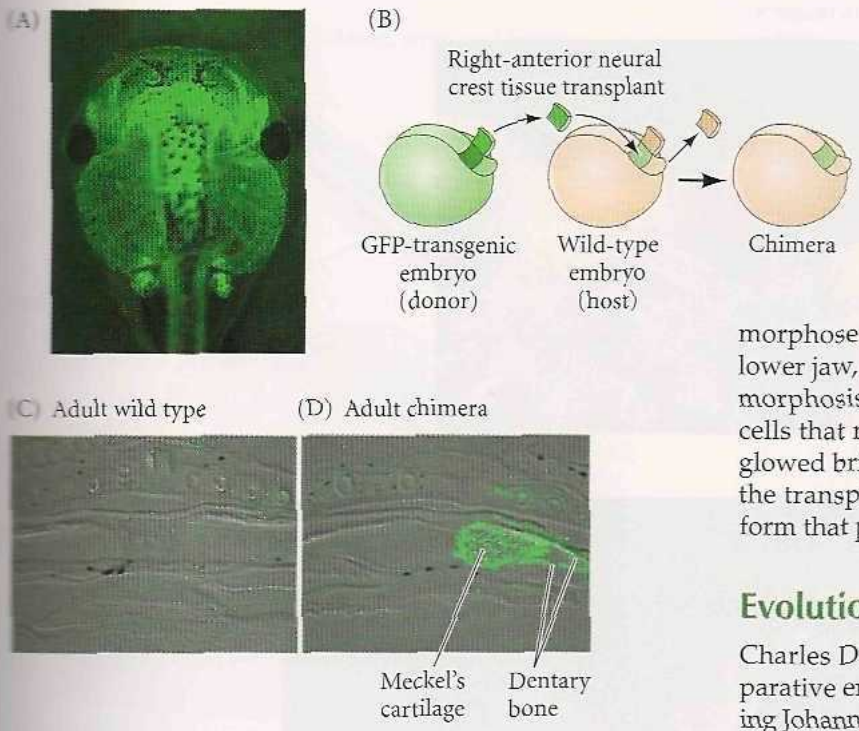


FIGURE 1.17 Fate mapping with transgenic DNA shows that the neural crest is critical in making the bones of the frog jaw. (A) Eggs injected with a virus containing a gene for green fluorescent protein (GFP) on a promoter that is active in all cell types produced tadpoles and frogs that express GFP in every cell (bright green). (B) Frogs labeled in this way were mated, and embryos were obtained that had formed their neural tubes and neural crests. A section from the right-side anterior (head-forming) neural crest was transplanted into the identical region of a wild-type embryo that lacked the *GFP* transgene. (C,D) After transplantation, development through the tadpole stage progressed. Once metamorphosis had occurred and the jawbones formed, the frog was sectioned and viewed under ultraviolet light. (C) In the wild-type, no fluorescence was seen in the jaw. (D) In the transgenic chimera, the right side glowed green wherever neural crest cells had migrated. These areas included Meckel's cartilage, as well as the dentary and other bone structures that formed during metamorphosis. (After Gross et al. 2006; photographs courtesy of J. B. Gross.)

called a *transgene*, because it contains DNA from another species.) When the infected embryonic cells are transplanted into a wild-type host, only the donor cells will express GFP and emit a visible green glow.

Gross and colleagues (2006) used this technique to identify the cells that form the bones of the frog skull. This has been a difficult question, because the tadpole does not have a bony skull; its skull is made of cartilage. So which part of the embryo forms the cells that wait patiently somewhere in the tadpole to become the bony skull of the adult frog? The experimenters first infected the eggs of the frog *Xenopus laevis* with a virus containing an active *GFP* gene. The virus became incorporated into the nuclear DNA of the frog egg, and the viral *GFP* gene was transmitted to every cell of the *Xenopus* embryo. In this way, Gross and colleagues generated tadpoles and adult frogs in which every cell glowed green when placed under ultraviolet light (Figure 1.17A). They then removed cells from the neural tube and neural crest of a GFP-transgenic embryo and transplanted the cells into a similar region of a normal *Xenopus* embryo (Figure 1.17B). Once the host embryos had metamorphosed into adult frogs, the biologists looked at the lower jaw, a structure whose bones form only after metamorphosis (Figure 1.17C,D). In addition to the cartilage cells that remained there, the dentary (lower jaw) bones glowed bright green, indicating that the cells originally in the transplanted region of neural crest had migrated to form that particular bone in the wild-type host's skull.

Evolutionary Embryology

Charles Darwin's theory of evolution restructured comparative embryology and gave it a new focus. After reading Johannes Müller's summary of von Baer's laws in 1842, Darwin saw that embryonic resemblances would be a strong argument in favor of the genetic connectedness of different animal groups. "Community of embryonic structure reveals community of descent," he would conclude in *On the Origin of Species* in 1859. Darwin's evolutionary interpretation of von Baer's laws established a paradigm that was to be followed for many decades—namely, that relationships between groups can be established by finding common embryonic or larval forms.

Even before Darwin, larval forms had been used for taxonomic classification. J. V. Thompson, for instance, demonstrated in the 1830s that larval barnacles were almost identical to larval shrimp, and therefore (correctly) counted barnacles as arthropods rather than molluscs (Figure 1.18; Winsor 1969). Darwin, an expert on barnacle taxonomy, celebrated this finding: "Even the illustrious Cuvier did not perceive that a barnacle is a crustacean, but a glance at the larva shows this in an unmistakable manner." Alexander Kowalevski (1871) made the similar type of discovery (publicized in Darwin's *The Descent of Man*, 1874) that tunicate larvae have a notochord and pharyngeal pouches, which came from the same germ layers as those same structures in fish and chicks. Thus the tunicate (an invertebrate) was related to the vertebrates, and the two great domains of the animal kingdom—invertebrates and vertebrates— were thereby united through larval structures. Darwin (1874) was thrilled, writing, "Thus, if we may rely on embryology, ever the safest guide in classification, it seems that we have at last gained a clue to the source whence the Vertebrata were derived."

Darwin also noted that embryonic organisms sometimes make structures that are inappropriate for their adult form but that show their relatedness to other animals. He pointed out the existence of eyes in embryonic moles, pelvic

(A) Barnacle

(B) Shrimp

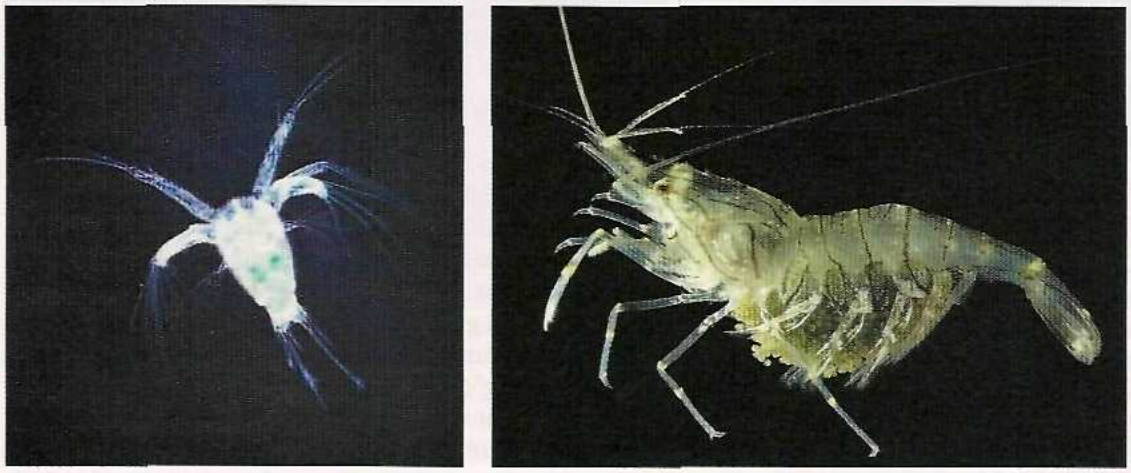


FIGURE 1.18 Larval stages reveal the common ancestry of two crustacean arthropods. (A) Barnacle. (B) Shrimp. Barnacles and shrimp both exhibit a distinctive larval stage (the nauplius) that underscores their common ancestry as crustacean arthropods, even though adult barnacles—once classified as molluscs—are sedentary, differing in body form and lifestyle from the free-swimming adult shrimp. (A © Wim van Egmond/Visuals Unlimited and © Barrie Watts/OSF/Photolibrary.com; B courtesy of U.S. National Oceanic and Atmospheric Administration and © Kim Taylor/Naturepl.com.)

bone rudiments in embryonic snakes, and teeth in baleen whale embryos.

Darwin also argued that adaptations that depart from the "type" and allow an organism to survive in its particular environment develop late in the embryo.* He noted that the differences among species within genera become greater as development persists, as predicted by von Baer's laws. Thus, Darwin recognized two ways of looking at "descent with modification." One could emphasize the common descent by pointing out embryonic similarities between two or more groups of animals, or one could emphasize the modifications by showing how development was altered to produce structures that enabled animals to adapt to particular conditions.

*Moreover, as first noted by Weismann (1875), larvae must have their own adaptations to help them survive. The adult viceroy butterfly mimics the monarch butterfly, but the viceroy caterpillar does not resemble the beautiful larva of the monarch. Rather, the viceroy larva escapes detection by resembling bird droppings (Begon et al. 1986).

Embryonic homologies

One of the most important distinctions made by evolutionary embryologists was the difference between *analogy* and *homology*. Both terms refer to structures that appear to be similar. **Homologous** structures are those organs whose underlying similarity arises from their being derived from a common ancestral structure. For example, the wing of a bird and the forelimb of a human are homologous, both having evolved from the forelimb bones of a common

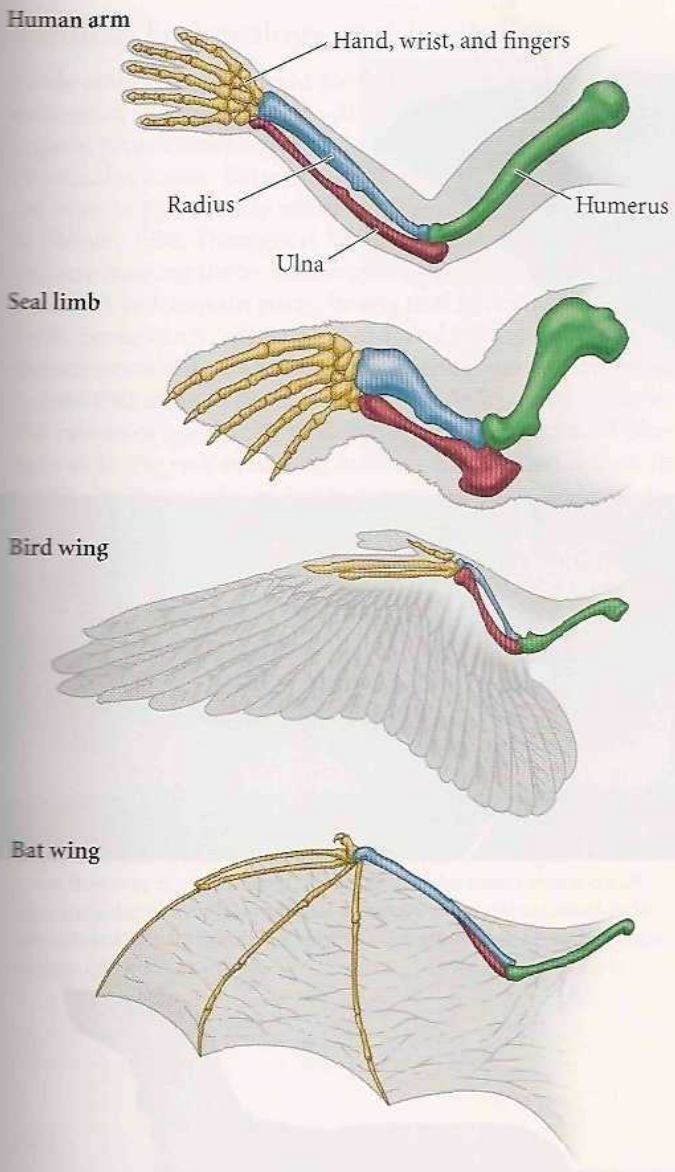


FIGURE 1.19 Homologies of structure among a human arm, a seal forelimb, a bird wing, and a bat wing; homologous supporting structures are shown in the same color. All four were derived from a common tetrapod ancestor and thus are homologous as forelimbs. The adaptations of bird and bat forelimbs to flight, however, evolved independently of each other, long after the two lineages diverged from their common ancestor. Therefore, as wings they are not homologous, but analogous.

ancestor. Moreover, their respective parts are homologous (Figure 1.19).

Analogous structures are those whose similarity comes from their performing a similar function rather than their arising from a common ancestor. For example, the wing of a butterfly and the wing of a bird are analogous; the two share a common function (and thus both are called wings), but the bird wing and insect wing did not arise from a common ancestral structure that became modified through evolution into bird wings and butterfly wings.*

As we will see in Chapter 19, evolutionary change is based on developmental change. The bat wing, for instance, is made in part by (1) maintaining a rapid growth rate in the cartilage that forms the fingers and (2) preventing the cell death that normally occurs in the webbing between the fingers (Figure 1.20). In human development, we start off with webbing between our digits. This webbing is important for creating the anatomical distinctions between our fingers (see Figure 13.26). Once the webbing has serve that function, genetic signals cause its cells to die, leaving us with free digits that can grasp and manipulate objects. Bats, however, use their fingers for flight—a feat accomplished by changing the genes that are activated in the webbing. The genes activated in embryonic bat webbing encode proteins that *prevent* cell death as well as accelerating finger elongation (Cretekos et al. 2005; Sears et al. 2006; Weatherbee et al. 2006). Thus, homologous anatomical structures (in this case, the human hand and the bat wing) can differentiate by altering development.

Changes in development provide the variations needed for evolutionary change. Darwin looked at artificial selection in pigeon and dog breeds, and these examples remain valuable resources for observing selectable variation. For instance, the short legs of dachshunds (Figure 1.21A) were selected by breeders who wanted to use these dogs to hunt badgers (German *Dachs*, "badger" + *Hund*, "dog"). The mutation that causes the short legs involves an extra copy of the gene *Fgf4*, which tells the cartilage precursor cells that they have divided enough and can start differentiating. With this extra copy of *Fgf4*, the cartilage cells are told too early that they should stop dividing, so the legs stop growing (Parker et al. 2009). Similarly, long-haired dachshunds (Figure 1.21B) differ from their short-haired relatives in having a mutation in the *Fgf5* gene† (Cadieu et al. 2009). This gene is involved in hair production and allows each follicle to make a longer hair shaft (Ota et al. 2002). Thus, mutations in genes controlling developmental processes can generate selectable variation.

*Homologies must always refer to the level of organization being compared. For instance, bird and bat wings are homologous as forelimbs but not as wings. In other words, they share an underlying structure of forelimb bones because birds and mammals share a common ancestor that possessed such bones. Bats, however, descended from a long line of non-winged mammals, while bird wings evolved independently, from the forelimbs of ancestral reptiles. As we will see, the structure of a bat's wing is markedly different from that of a bird's wing.

†The FGF genes will be discussed throughout this book as they regulate construction of numerous organs. Independently acquired mutations in the *Fgf5* gene are also responsible for the long-haired phenotype of Persian cats (Drögemüller et al. 2007; Kehler et al. 2007). However, *Fgf5* is not considered a good candidate to explain the woolliness of mammoths: the sequence of the *Fgf5* gene extracted from the DNA of extinct woolly mammoths appears virtually identical to that of the gene in modern elephants (Roca et al. 2009).

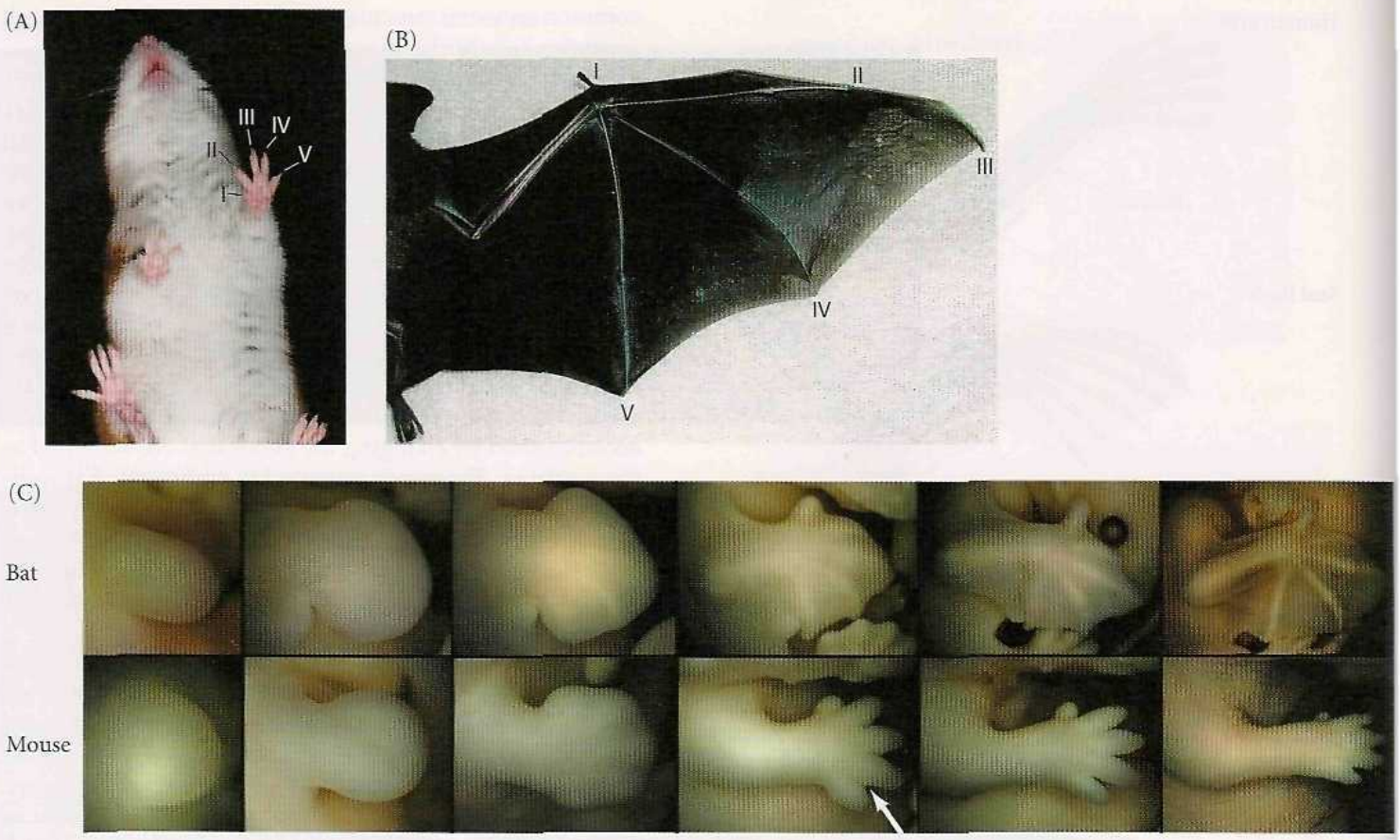


FIGURE 1.20 Development of bat and mouse forelimbs. (A, B) Mouse and bat torsos, showing the elongated fingers and the prominent webbing in the bat wing. A mouse forelimb is shown in the inset, and the digit numbers (I, thumb; V, "pinky") are on both sets. (C) Comparison of mouse and bat forelimb morphogenesis. Both limbs start as webbed appendages, but the webbing between the mouse's digits dies at embryonic day 14 (arrow). The webbing in the bat forelimb does not die, but is sustained as the fingers grow. (A courtesy of David McIntyre; B,C from Cretekos et al. 2008, courtesy of C. J. Cretekos.)

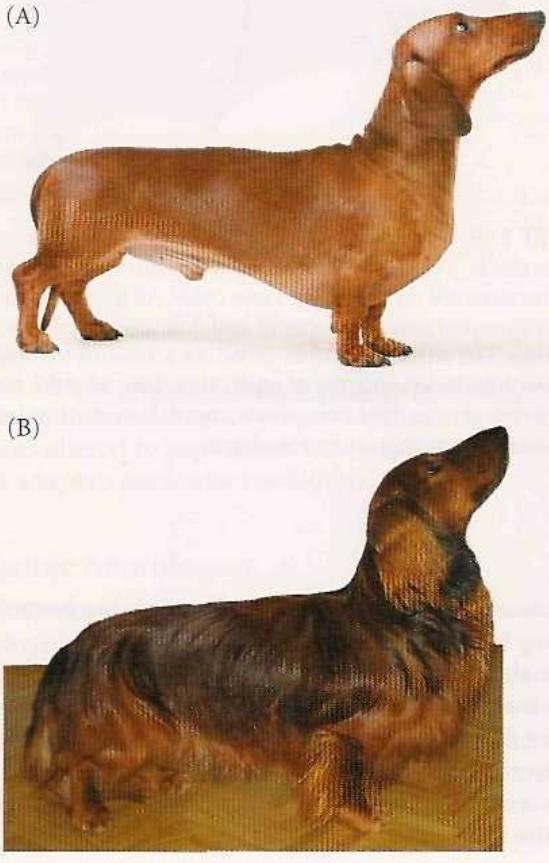


FIGURE 1.21 Selectable variation through mutations of genes that work during developmental. (A) The dachshund has been selected by breeders for its small legs, which enable it to seek badgers in their tunnels. The small legs are a result of premature cessation of cell division in the limb cartilage precursor cells. This premature end to cell division is caused by early activation of the cartilage FGF receptor protein, because the dachshund genome has an extra copy of the *Fgf4* gene. (B) Long-haired dachshunds have an additional mutation, a truncated *Fgf5* gene, which alters the hair follicle cycle, thereby allowing the hair growth beyond the wild-type levels. (A © Alex Potemkin/istockphoto.com; B courtesy of K. Lilleväli.)

Medical Embryology and Teratology

While embryologists could look at embryos to describe the evolution of life and how different animals form their organs, physicians became interested in embryos for more practical reasons. Between 2% and 5% of human infants are born with a readily observable anatomical abnormality (Winter 1996; Thorogood 1997). These abnormalities may include missing limbs, missing or extra digits, cleft palate, eyes that lack certain parts, hearts that lack valves, and so forth. Some birth defects are produced by mutant genes or chromosomes, and some are produced by environmental factors that impede development. Physicians need to know the causes of specific birth defects in order to counsel parents as to the risk of having another malformed infant. In addition, the study of birth defects can tell us how the human body is normally formed. In the absence of experimental data on human embryos, nature's "experiments" sometimes offer important insights into how the human body becomes organized.*

*The word *monster*, used frequently in textbooks prior to the mid-twentieth century to describe malformed infants, comes from the Latin *monstrare*, "to show or point out." This is also the root of the English word *demonstrate*. In the 1830s, J. F. Meckel realized that syndromes of congenital anomalies demonstrated certain principles about normal development. Parts of the body that were affected together must have some common developmental origin or mechanism that was being affected. It should also be noted that a condition considered a developmental anomaly in one situation may be considered advantageous in another. The short legs of dachshunds is only one such example.

Genetic malformations and syndromes

Abnormalities caused by genetic events (gene mutations, chromosomal aneuploidies, and translocations) are called **malformations**. Malformations often appear as **syndromes** (Greek, "running together"), in which several abnormalities occur concurrently. For instance, a human malformation called piebaldism, shown in Figure 1.22A, is due to a dominant mutation in a gene (*KIT*) on the long arm of chromosome 4 (Spritz et al. 1992). The piebald syndrome includes anemia, sterility, unpigmented regions of the skin and hair, deafness, and the absence of the nerves that cause peristalsis in the gut. The common feature underlying these conditions is that the *KIT* gene encodes a protein that is expressed in the neural crest cells and in the precursors of blood cells and germ cells. The Kit protein enables these cells to proliferate. Without this protein, the neural crest cells—which generate the pigment cells, certain ear cells, and the gut neurons—do not multiply as extensively as they should (resulting in underpigmentation, deafness, and gut malformations), nor do the precursors of the blood cells (resulting in anemia) or the germ cells (resulting in sterility).

Developmental biologists and clinical geneticists often study human syndromes (and determine their causes) by studying animals that display the same syndrome. These are called **animal models** of the disease; the mouse model for piebaldism is shown in Figure 1.22B. It has a phenotype very similar to that of the human condition, and it is caused by a mutation in the *Kit* gene of the mouse.*

*The mouse *Kit* and human *KIT* genes are considered homologous by their structural similarities and their presumed common ancestry. Human genes are usually italicized and written in all capitals. Mouse genes are italicized, but only the first letter is usually capitalized. Gene products—proteins—are not italicized. If the protein has no standard biochemical or physiological name, it is usually represented with the name of the gene in Roman type, with the first letter capitalized. These rules are frequently bent, however. One is reminded of Cohen's (1982) dictum that "Academicians are more likely to share each other's toothbrush than each other's nomenclature."

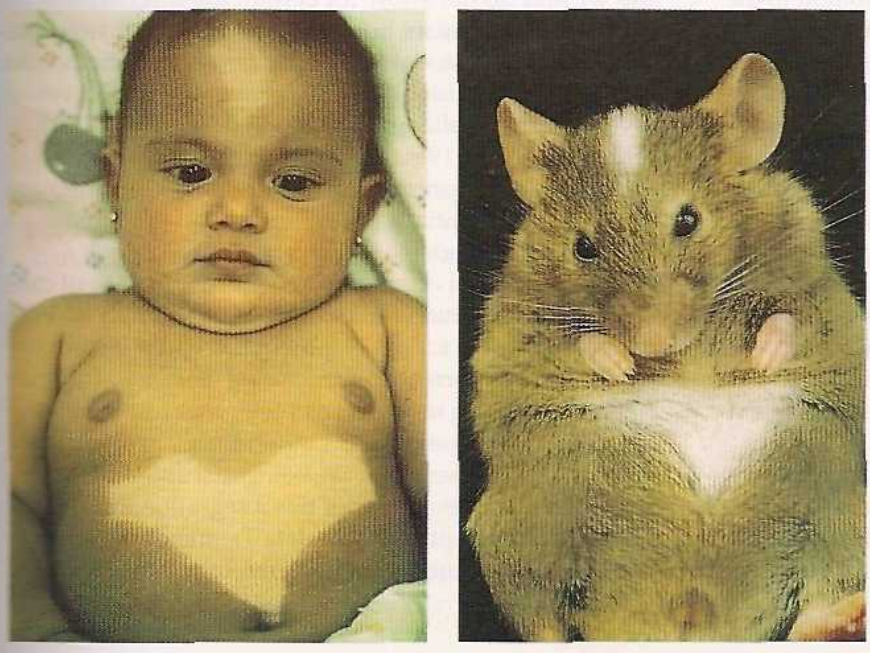


FIGURE 1.22 Developmental anomalies caused by genetic mutation. (A) Piebaldism in a human infant. This genetically produced condition results in sterility, anemia, and underpigmented regions of the skin and hair, along with defective development of gut neurons and the ear. Piebaldism is caused by a mutation in the *KIT* gene. The Kit protein is essential for the proliferation and migration of neural crest cells, germ cell precursors, and blood cell precursors. (B) A piebald mouse with a *Kit* mutation. Mice provide important models for studying human developmental diseases. (Photographs courtesy of R. A. Fleischman.)

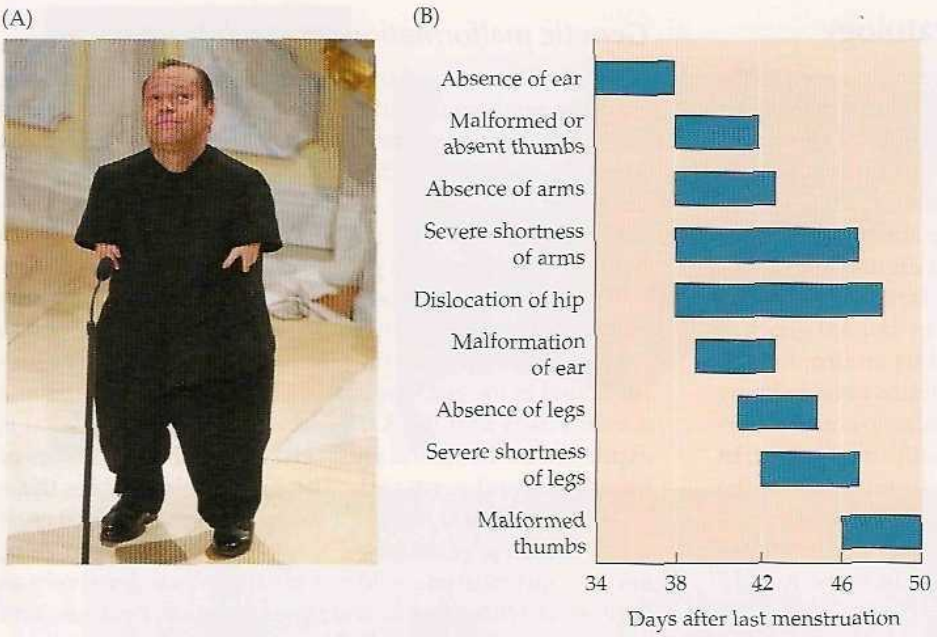


FIGURE 1.23 Developmental anomalies caused by an environmental agent. (A) Phocomelia, the lack of proper limb development, was the most visible of the birth defects that occurred in many children born in the early 1960s whose mothers took the drug thalidomide during pregnancy. These children are now adults; this photograph is a recent one of Grammy-nominated German singer Thomas Quasthoff. (B) Thalidomide disrupts different structures at different times of human development. (A © AP Photo; B after Nowack 1965.)

Disruptions and teratogens

Abnormalities caused by exogenous agents (certain chemicals or viruses, radiation, or hyperthermia) are called **disruptions**. The agents responsible for these disruptions are called **teratogens** (Greek, "monster-formers"), and the study of how environmental agents disrupt normal development is called *teratology*. Teratogens were brought to the attention of the public in the early 1960s. In 1961, Lenz and McBride independently accumulated evidence that the drug thalidomide, prescribed as a mild sedative to many pregnant women, caused an enormous increase in a previously rare syndrome of congenital anomalies. The most noticeable of these anomalies was *phocomelia*, a condition in which the long bones of the limbs are deficient or absent (Figure 1.23A). More than 7000 affected infants were born to women who took thalidomide, and a woman need only have taken one tablet for her child to be born with all four limbs deformed (Lenz 1962, 1966; Toms 1962). Other abnormalities induced by the ingestion of this drug included heart defects, absence of the external ears, and malformed intestines.

Nowack (1965) documented the period of susceptibility during which thalidomide caused these abnormalities (Figure 1.23B). The drug was found to be teratogenic only during days 34–50 after the last menstruation (i.e., 20–36 days postconception). From days 34 to 38, no limb abnormalities are seen, but during this period, thalidomide can cause the absence or deficiency of ear components. Malformations of the upper limbs are seen before those of the lower limbs, because during development the arms form slightly before the legs.

The only animal models for thalidomide are primates, and we still do not know for certain the mechanisms by which this drug causes human developmental disruptions (although it may work by blocking certain molecules from the developing mesoderm, thus preventing blood vessel development). Thalidomide was withdrawn from the market in November 1961. However, the drug is once more beginning to be prescribed (although not to pregnant women) as a potential anti-tumor and anti-autoimmunity drug (Raje and Anderson 1999).

The integration of anatomical information about congenital malformations with our new knowledge of the genes responsible for development has had a revolutionary effect and is currently restructuring medicine. This integration is allowing us to discover the genes responsible for inherited malformations, and it permits us to identify the steps in development that are being disrupted by teratogens. We will see examples of this integration throughout this text, and Chapter 17 will detail some of the remarkable new discoveries in human teratology.

Snapshot Summary: *Developmental Anatomy*

1. The life cycle can be considered a central unit in biology; the adult form need not be paramount. The basic animal life cycle consists of fertilization, cleavage, gastrulation, germ layer formation, organogenesis, metamorphosis, adulthood, and senescence.
2. In gametogenesis, the germ cells (i.e., those cells that will become sperm or eggs) undergo meiosis. Eventually, usually after adulthood is reached, the mature gametes are released to unite during fertilization. The resulting new generation then begins development.
3. Epigenesis happens. New organisms are created de novo each generation from the relatively disordered cytoplasm of the egg.
4. Preformation is not found in the anatomical structures themselves, but in the genetic instructions that instruct their formation. The inheritance of the fertilized egg includes the genetic potentials of the organism. These preformed nuclear instructions include the ability to respond to environmental stimuli in specific ways.
5. The three germ layers give rise to specific organ systems. The ectoderm gives rise to the epidermis, nervous system, and pigment cells; the mesoderm generates the kidneys, gonads, muscles, bones, heart, and blood cells; and the endoderm forms the lining of the digestive tube and the respiratory system.
6. Karl von Baer's principles state that the general features of a large group of animals appear earlier in the embryo than do the specialized features of a smaller group. As each embryo of a given species develops, it diverges from the adult forms of other species. The early embryo of a "higher" animal species is not like the adult of a "lower" animal.
7. Labeling cells with dyes shows that some cells differentiate where they form, whereas others migrate from their original sites and differentiate in their new locations. Migratory cells include neural crest cells and the precursors of germ cells and blood cells.
8. "Community of embryonic structure reveals community of descent" (Charles Darwin, *On the Origin of Species*).
9. Homologous structures in different species are those organs whose similarity is due to sharing a common ancestral structure. Analogous structures are those organs whose similarity comes from serving a similar function (but which are not derived from a common ancestral structure).
10. Congenital anomalies can be caused by genetic factors (mutations, aneuploidies, translocations) or by environmental agents (certain chemicals, certain viruses, radiation).
11. Syndromes consist of sets of developmental abnormalities that "run together."
12. Organs that are linked in developmental syndromes share either a common origin or a common mechanism of formation.

For Further Reading

Complete bibliographical citations for all literature cited in this chapter can be found at the free-access website **www.devbio.com**

Cadieu, E. and 19 others. 2009. Coat variation in the domestic dog is governed by variants in three genes. *Science* 326: 150–153.

Cebra-Thomas, J. A., E. Betters, M. Yin, C. Plafkin, K. McDow and S. F. Gilbert. 2007. Evidence that a late-emerging population of trunk neural crest cells forms the plastron bones in the turtle *Trachemys scripta*. *Evol. Dev.* 9: 267–277.

Larsen, E. and H. McLaughlin. 1987. The morphogenetic alphabet: Lessons for simple-minded genes. *BioEssays* 7: 130–132.

Le Douarin, N. M. and M.-A. Teillet. 1973. The migration of neural crest cells to the wall of the digestive tract in the avian embryo. *J. Embryol. Exp. Morphol.* 30: 31–48.

Nishida, H. 1987. Cell lineage analysis in ascidian embryos by intracellular injection of a tracer enzyme. III. Up to the tissue-restricted stage. *Dev. Biol.* 121: 526–541.

Pinto-Correia, C. 1997. *The Ovary of Eve: Egg and Sperm and Preformation*. University of Chicago Press, Chicago.

Weatherbee, S. D., R. R. Behringer, J. J. Rasweiler 4th and L. A. Niswander. 2006. Interdigital webbing retention in bat wings illustrates genetic changes underlying amniote limb diversification. *Proc. Natl. Acad. Sci. USA* 103: 15103–15107.

Winter, R. M. 1996. Analyzing human developmental abnormalities. *BioEssays* 18: 965–971.

Woo, K. and S. E. Fraser. 1995. Order and coherence in the fate map of the zebrafish embryo. *Development* 121: 2595–2609.

Go Online

WEBSITE 1.1 Conklin's art and science. The plates from Conklin's remarkable 1905 paper are online. Looking at them, one can see the precision of his observations and how he constructed his fate map of the tunicate embryo.

VADE MECUM

Chick-quail chimeras. We are fortunate to present here a movie made by Dr. Nicole Le Douarin of her chick-quail grafts. You will be able to see how these grafts are actually done.

The compound microscope. The compound microscope has been the critical tool of developmental anatomists. Mastery of microscopic techniques allows one to enter an entire world of form and pattern.

Histotechniques. Most cells must be stained in order to see them; different dyes stain different types of molecules. Instructions on staining cells to observe particular structures (such as the nucleus) are given here.

Developmental Genetics

2

But whatever the immediate operations of the genes turn out to be, they most certainly belong to the category of developmental processes and thus belong to the province of embryology.

C. H. WADDINGTON (1956)

We have entered the cell, the mansion of our birth, and have started the inventory of our acquired wealth.

ALBERT CLAUDE (1974)

CYTOLOGICAL STUDIES DONE AT THE TURN OF THE TWENTIETH CENTURY established that the chromosomes in each cell of an organism's body are the mitotic descendants of the chromosomes established at fertilization (Wilson 1896; Boveri 1904). In other words, each somatic cell nucleus has the same chromosomes—and therefore the same set of genes—as all the other somatic nuclei. This fundamental concept is called **genomic equivalence**. Given this concept, one of the major questions facing biologists of the early twentieth century was how nuclear genes could direct development when these genes are the same in every cell type (Harrison 1937; Just 1939). If every cell in the body contains the genes for hemoglobin and insulin proteins, why is it that hemoglobin proteins are made only in red blood cells, insulin proteins are made only in certain pancreas cells, and neither protein is made in the kidneys or nervous system?

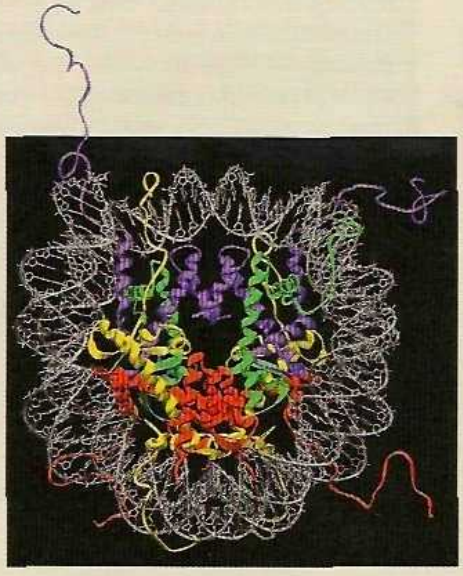
Based on the embryological evidence for genomic equivalence (as well as on bacterial models of gene regulation), a consensus emerged in the 1960s that the answer to this question lies in **differential gene expression**. The three postulates of differential gene expression are:

- Every cell nucleus contains the complete genome established in the fertilized egg. In molecular terms, the DNAs of all differentiated cells are identical.
- The unused genes in differentiated cells are neither destroyed nor mutated, but retain the potential for being expressed.
- Only a small percentage of the genome is expressed in each cell, and a portion of the RNA synthesized in each cell is specific for that cell type.

Gene expression can be regulated at several levels such that different cell types synthesize different sets of proteins:

- **Differential gene transcription** regulates which of the nuclear genes are transcribed into nuclear RNA.
- **Selective nuclear RNA processing** regulates which of the transcribed RNAs (or which parts of such a nuclear RNA) are able to enter into the cytoplasm and become messenger RNAs.
- **Selective messenger RNA translation** regulates which of the mRNAs in the cytoplasm are translated into proteins.
- **Differential protein modification** regulates which proteins are allowed to remain and/or function in the cell.

Some genes (such as those coding for the globin proteins of hemoglobin) are regulated at all these levels.



Evidence for Genomic Equivalence

Until the mid-twentieth century, genomic equivalence was not so much proved as it was assumed (because every cell is the mitotic descendant of the fertilized egg. One of the first tasks of developmental genetics was to determine whether every cell of an organism indeed does have the same *genome*—that is, the same set of genes—as every other cell.

Evidence that every cell in the body has the same genome originally came from the analysis of *Drosophila* chromosomes, in which the DNA of certain larval tissues undergoes numerous rounds of DNA replication without separation, such that the structure of the chromosomes can be seen. In these **polytene** (Gr. "many strands") **chromosomes**, no structural differences were seen between cells; but different regions were seen to be "puffed up" at different times and in different cell types, suggesting that these areas were actively making RNA (Beerman 1952).

See WEBSITE 2.1
Does the genome or the cytoplasm direct development?

See WEBSITE 2.2
The origins of developmental genetics

When Giemsa dyes allowed such observations to be made in mammalian chromosomes, it was also found that no chromosomal regions were lost in most cells. These observations, in turn, were confirmed by nucleic acid hybridization studies, which (for instance) found globin genes in pancreatic tissue, which does not make globin proteins.

But the ultimate test of whether the nucleus of a differentiated cell has undergone irreversible functional restriction is to have that nucleus generate every other type of differentiated cell in the body. If each cell's nucleus is identical to the zygote nucleus, then each cell's nucleus should also be capable of directing the entire development of the organism when transplanted into an activated enucleated egg. As early as 1895, the embryologist Yves Delage predicted that "If, without deterioration, the egg nucleus could be replaced by the nucleus of an ordinary embryonic cell, we should probably see this egg developing without changes" (Delage 1895, p. 738).

In 1952, Briggs and King demonstrated that blastula cell nuclei could direct the development of complete tadpoles when transferred into the cytoplasm of an activated enucleated frog egg. This procedure is called **somatic nuclear transfer** or, more commonly, **cloning**. Nuclei from adult frogs, however, were not able to generate adult frogs. However, adult nuclei (from skin cells, for instance) were

SIDELIGHTS & SPECULATIONS

The Basic Tools of Developmental Genetics

DNA analysis

Embryologist Theodor Boveri (1904) wrote that to discover the mechanisms of development, it was "not cell nuclei, not even individual chromosomes, but certain parts of certain chromosomes from certain cells that must be isolated and collected in enormous quantities for analysis." This analysis was finally made possible by the techniques of gene cloning, DNA sequencing, Southern blotting, gene knockouts, and enhancer traps. In addition, techniques for showing which enhancers and promoters are methylated and which are unmethylated have become more important, as investigations of differential gene transcription have focused on these elements.

For discussions of these techniques, see Website 2.3.

RNA analysis

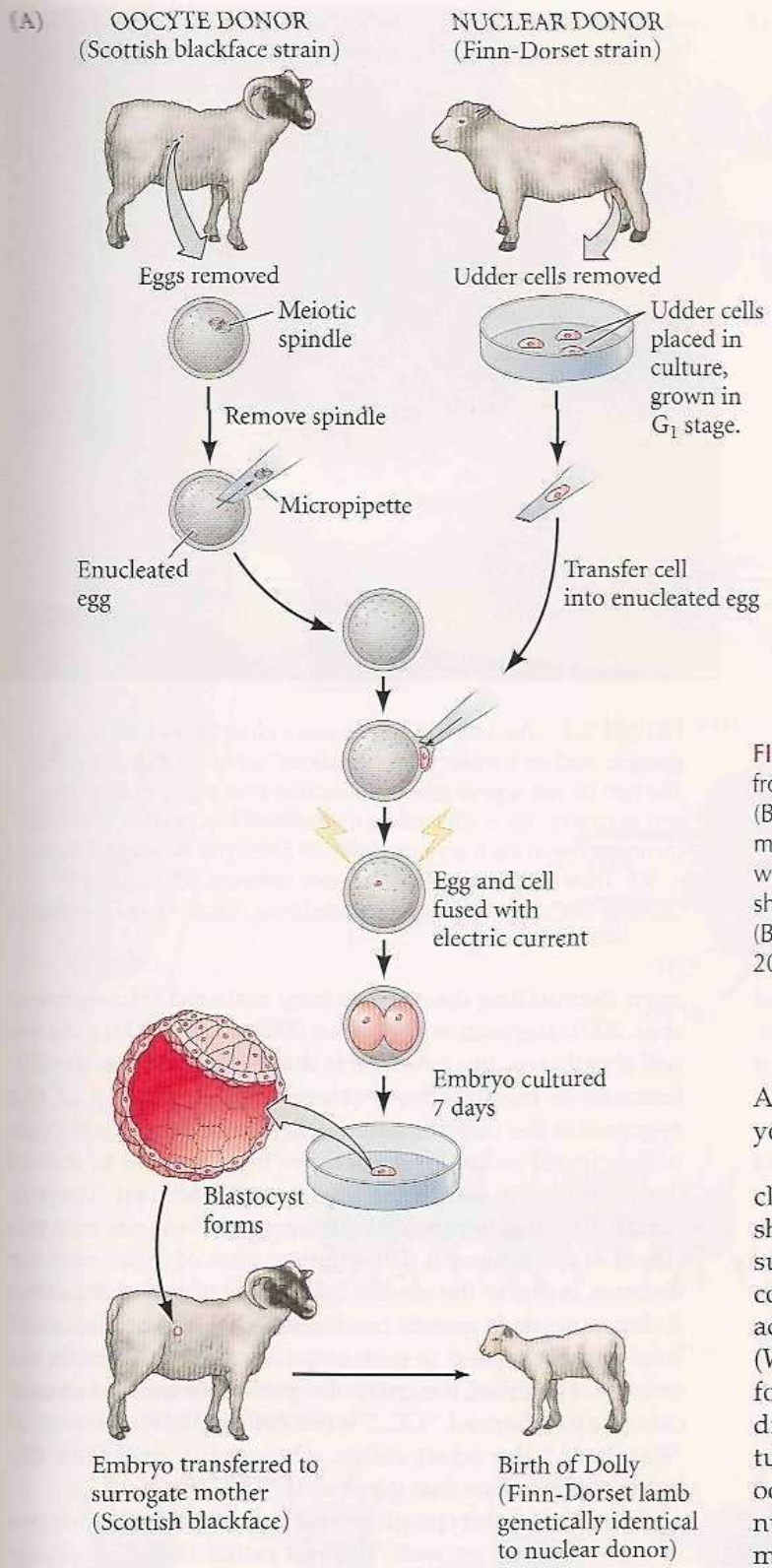
Differential gene transcription is critical in development. In order to know the time of gene expression and the place of gene expression, one has to be able to have procedures that actually locate a particular type of messenger RNA. These techniques include northern blots, RT-PCR, in situ hybridization, and array technology. To ascertain the function of these mRNAs, new techniques have been formulated, which include antisense and RNA interference (which destroy messages), Cre-lox analysis (which allows the message to be made or destroyed in particular cell types) and ChIP-on-Chip techniques (which enable one to localize active chromatin).

For discussions of these techniques, see Website 2.4.

Bioinformatics

Modern developmental genetics often involves comparing DNA sequences (especially regulatory units such as enhancers and 3' UTRs) and looking at specific genomes to determine how genes are being regulated. "High-throughput" RNA analysis by micro- and macroarrays enables researchers to compare thousands of mRNAs, and computer-aided synthetic techniques can predict interactions between proteins and mRNAs. Various free websites enable researchers to use the tools that allow such comparisons. Other sites are organism- or organ-specific and are used by researchers studying that particular organ or organism.

For more about these sites and links to them, see Website 2.5



(B)

FIGURE 2.1 Cloned mammals have been created using nuclei from adult somatic cells. (A) Procedure used for cloning sheep. (B) Dolly, the adult sheep on the left, was derived by fusing a mammary gland cell nucleus with an enucleated oocyte, which was then implanted in a surrogate mother (of a different breed of sheep) that gave birth to Dolly. Dolly later gave birth to a lamb (Bonnie, at right) by normal reproduction. (A after Wilmut et al. 2000; B, photograph by Roddy Field, © Roslin Institute.)

able to direct the development of all the organs of the tadpoles (Gurdon et al. 1975). Although the tadpoles all died prior to feeding, their existence showed that a single differentiated cell nucleus still retained incredible potencies. A nucleus of a skin cell could produce all the cells of a young tadpole.

In 1997, Ian Wilmut announced that a sheep had been cloned from a somatic cell nucleus from an adult female sheep. This was the first time an adult vertebrate had been successfully cloned from another adult. Wilmut and his colleagues had taken cells from the mammary gland of an adult (6-year-old) pregnant ewe and put them into culture (Wilmut et al. 1997; Figure 2.1A). The culture medium was formulated to keep the nuclei in these cells at the intact diploid stage (G1) of the cell cycle. This cell-cycle stage turned out to be critical. The researchers then obtained oocytes from a different strain of sheep and removed their nuclei. These oocytes had to be in the second meiotic metaphase (the stage at which they are usually fertilized). Fusion of the donor cell and the enucleated oocyte was accomplished by bringing the two cells together and sending electric pulses through them, destabilizing the cell membranes and allowing the cells to fuse. The same electric pulses that fused the cells activated the egg to begin development. The resulting embryos were eventually transferred into the uteri of pregnant sheep.

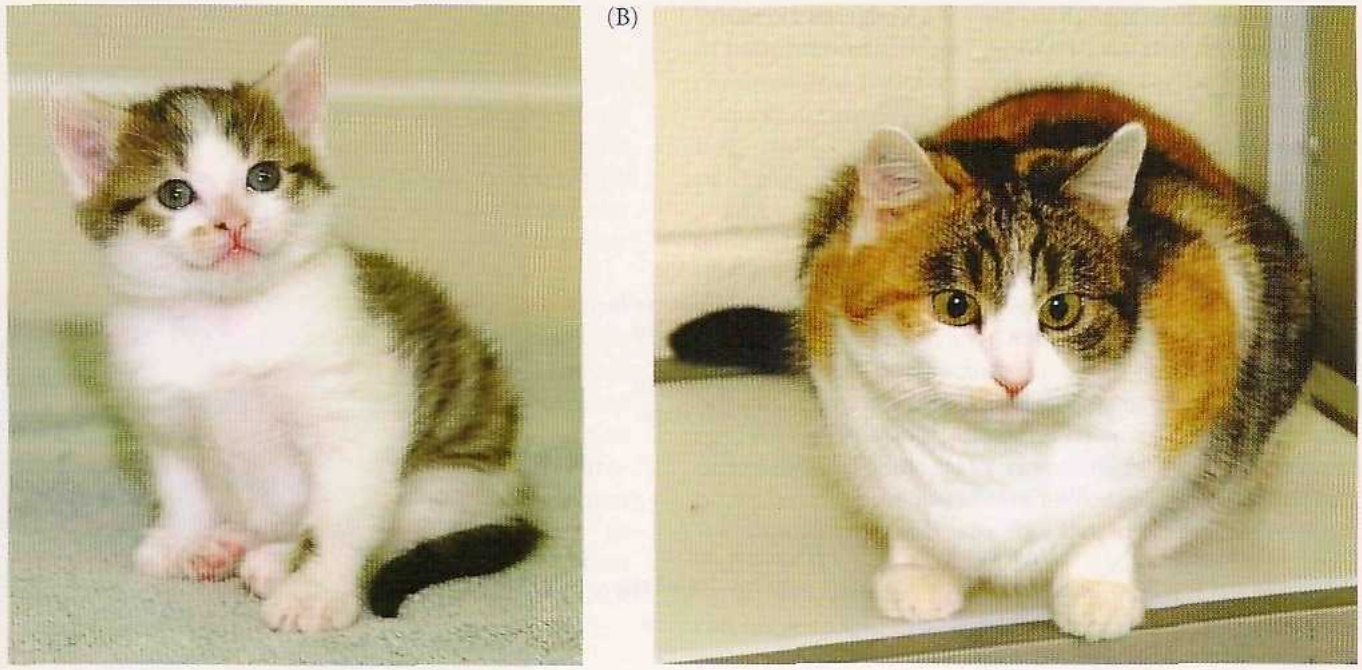


FIGURE 2.2 The kitten "CC" (A) was a clone produced using somatic nuclear transfer from "Rainbow," a female calico cat (B). The two do not appear identical because coat pigmentation pattern in calico cats is affected by the random inactivation of one X chromosome in each somatic cell (see Sidelights & Speculations, p. 50). Their behaviors were also quite different. (Photographs courtesy of College of Veterinary Medicine, Texas A&M University.)

Of the 434 sheep oocytes originally used in this experiment, only one survived: Dolly* (Figure 2.1B). DNA analysis confirmed that the nuclei of Dolly's cells were derived from the strain of sheep from which the donor nucleus was taken (Ashworth et al. 1998; Signer et al. 1998). Cloning of adult mammals has been confirmed in guinea pigs, rabbits, rats, mice, dogs, cats, horses, and cows. In 2003, a cloned mule became the first sterile animal to be so reproduced (Woods et al. 2003). Thus it appears that the nuclei of vertebrate adult somatic cells contain all the genes needed to generate an adult organism. No genes necessary for development have been lost or mutated in the somatic cells.†

Certain caveats must be applied, however. First, although it appears that all the organs were properly formed in the cloned animals, many of the clones developed debilitating diseases as they matured (Humphreys et al. 2001; Jaenisch and Wilmut 2001; Kolata 2001). As we will shortly see, this problem is due in large part to the differences in methylation between the chromatin of the zygote and the differentiated cell. Second, the phenotype of the cloned animal is sometimes not identical to that of the animal from which the nucleus was derived. There is variability due to random chromosomal events and the effects of environment. The pigmentation of calico cats, for instance, is due to the *random inactivation* of one or the other X chromosome (a genetic mechanism that will be discussed later in this chapter) in each somatic cell of the female cat embryo. Therefore, the coat color pattern of the first cloned cat, a calico named "CC," were different from those of "Rainbow," the adult calico whose cells provided the implanted nucleus that generated "CC" (Figure 2.2).

The same genotype gives rise to multiple phenotypes in cloned sheep as well. Wilmut noted that four sheep cloned from blastocyst nuclei from the same embryo "are genetically identical to each other and yet are very different in size and temperament, showing emphatically that an animal's genes do not 'determine' every detail of its physique and personality" (Wilmut et al. 2000, p. 5). Wilmut concludes that for this and other reasons, the "resurrection" of lost loved ones by cloning is not feasible.

See WEBSITE 2.6 Cloning and nuclear equivalence

*The creation of Dolly was the result of a combination of scientific and social circumstances. These circumstances involved job security, people with different areas of expertise meeting each other, children's school holidays, international politics, and who sits near whom in a pub. The complex interconnections giving rise to Dolly are told in *The Second Creation* (Wilmut et al. 2000), a book that should be read by anyone who wants to know how contemporary science actually works. As Wilmut acknowledged (p. 36), "The story may seem a bit messy, but that's because life is messy, and science is a slice of life."

†Although cloning humans does not seem feasible at present, each cell of the human body (with just a few exceptions, such as lymphocytes) does appear to contain the same genome as every other cell. As we will see in Chapter 17, adding certain activated transcription factors to ordinary skin fibroblasts will convert them into embryonic stem cells that are indeed capable of generating entire embryos, at least in mice.

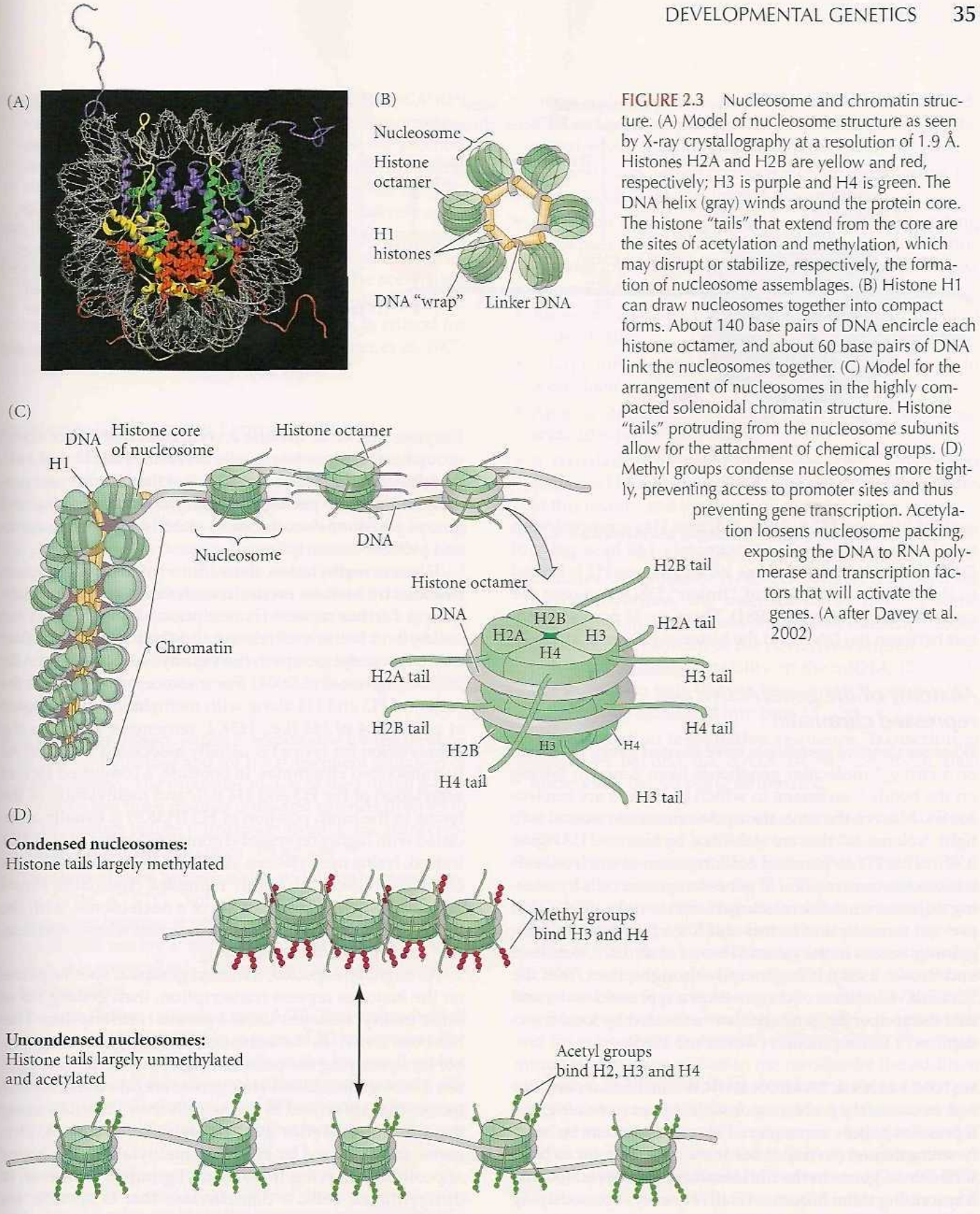


FIGURE 2.3 Nucleosome and chromatin structure. (A) Model of nucleosome structure as seen by X-ray crystallography at a resolution of 1.9 Å. Histones H2A and H2B are yellow and red, respectively; H3 is purple and H4 is green. The DNA helix (gray) winds around the protein core. The histone "tails" that extend from the core are the sites of acetylation and methylation, which may disrupt or stabilize, respectively, the formation of nucleosome assemblages. (B) Histone H1 can draw nucleosomes together into compact forms. About 140 base pairs of DNA encircle each histone octamer, and about 60 base pairs of DNA link the nucleosomes together. (C) Model for the arrangement of nucleosomes in the highly compacted solenoidal chromatin structure. Histone "tails" protruding from the nucleosome subunits allow for the attachment of chemical groups. (D) Methyl groups condense nucleosomes more tightly, preventing access to promoter sites and thus preventing gene transcription. Acetylation loosens nucleosome packing, exposing the DNA to RNA polymerase and transcription factors that will activate the genes. (A after Davey et al. 2002)

Differential Gene Transcription

So how does the same genome give rise to different cell types? To understand this, one needs to understand the anatomy of the genes. One of the fundamental differences distinguishing most eukaryotic genes from prokaryotic genes is that eukaryotic genes are contained within a complex of DNA and protein called **chromatin**. The protein component constitutes about half the weight of chromatin and is composed largely of **histones**. The **nucleosome** is the basic unit of chromatin structure (Figure 2.3). It is composed of an octamer of histone proteins (two molecules

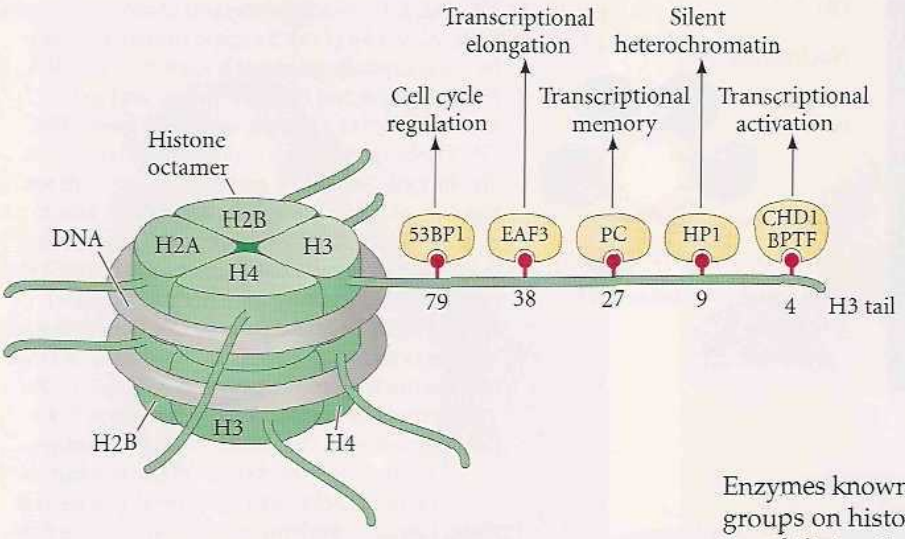


FIGURE 2.4 Histone methylations on histone H3. The tail of histone H3 (its aminomost sequence, at the beginning of the protein) sticks out from the nucleosome and is capable of being methylated or acetylated. Here, lysines can be methylated and recognized by particular proteins. Methylated lysine residues at positions 4, 38, and 79 are associated with gene activation, whereas methylated lysines at positions 9 and 27 are associated with repression. The proteins binding these sites (not shown to scale) are represented above the methyl group. (After Kouzarides and Berger 2007.)

each of histones H2A, H2B, H3, and H4) wrapped with two loops containing approximately 140 base pairs of DNA (Kornberg and Thomas 1974). Histone H1 is bound to the 60 or so base pairs of "linker" DNA between the nucleosomes (Weintraub 1984). There are 14 points of contact between the DNA and the histones (Luger et al. 1997).

Anatomy of the gene: Active and repressed chromatin

Whereas classical geneticists have likened genes to "beads on a string," molecular geneticists liken genes to "string on the beads," an image in which the beads are nucleosomes. Most of the time, the nucleosomes are wound into tight "solenoids" that are stabilized by histone H1 (Figure 2.3C). This H1-dependent conformation of nucleosomes inhibits the transcription of genes in somatic cells by packing adjacent nucleosomes together into tight arrays that prevent transcription factors and RNA polymerases from gaining access to the genes (Thoma et al. 1979; Schlissel and Brown 1984). It is generally thought, then, that the "default" condition of chromatin is a repressed state, and that tissue-specific genes become activated by local interruption of this repression (Weintraub 1985).

HISTONES AS AN ACTIVATION SWITCH The histones are critical because they are responsible for maintaining the repression of gene expression. This repression can be locally strengthened (so that it becomes very difficult to transcribe those genes in the nucleosomes) or relieved (so that transcribing them becomes relatively easy) by modifying the histones (Figure 2.3D). Repression and activation are controlled to a large extent by modifying the tails of histones H3 and H4 with two small organic groups: methyl (CH_3) and acetyl ($COCH_3$) residues. In general, **histone acetylation**—the addition of negatively charged acetyl groups to histones—neutralizes the basic charge of lysine and loosens the histones. This activates transcription. Enzymes known as **histone acetyltransferases** place acetyl groups on histones (especially on lysines in H3 and H4), destabilizing the nucleosomes so that they come apart easily. As might be expected, then, enzymes that *remove* acetyl groups—**histone deacetylases**—stabilize the nucleosomes and prevent transcription.

Histone methylation, the addition of methyl groups to histones by **histone methyltransferases**, can either activate or further repress transcription, depending on the amino acid being methylated and the presence of other methyl or acetyl groups in the vicinity (see Strahl and Allis 2000; Cosgrove et al. 2004). For instance, acetylation of the "tails" of H3 and H4 along with methylation of the lysine at position 4 of H3 (i.e., H3K4; remember that K is the abbreviation for lysine) is usually associated with actively transcribed chromatin. In contrast, a combined lack of acetylation of the H3 and H4 tails and methylation of the lysine in the ninth position of H3 (H3K9) is usually associated with highly repressed chromatin (Norma et al. 2001). Indeed, lysine methylations at H3K9, H327, and H4K20 are often associated with highly repressed chromatin. Figure 2.4 shows a schematic drawing of a nucleosome, with the histone H3 tail having on it some residues whose modification can regulate transcription.

As might be expected, if methyl groups at specific places on the histones repress transcription, then getting rid of these methyl moieties should permit transcription. This has been shown in the activation of those genes responsible for specifying the posterior halves of vertebrate bodies. These genes, called Hox genes, encode transcription factors that are critical in giving cells their identities along the anterior-posterior axis. In early development, Hox genes are repressed by H3K27 trimethylation (the lysine at position 27 having three methyl groups). However, in differentiated cells, a demethylase that is specific for H3K27me3 is recruited to these promoters and enables the gene to be transcribed (Agger et al. 2007; Lan et al. 2007).

The effects of methylation in controlling gene transcription are extensive. So far, we have documented transcriptional regulation by *histone* methylation. Later in this chapter we will discuss the exciting research on the control of transcription by *DNA* methylation.

HISTONE REGULATION OF TRANSCRIPTIONAL ELONGATION In addition to regulating the initiation of the transcriptional complex (i.e., getting RNA polymerase on the promoter), nucleosomes also appear to regulate the progression of RNA polymerase and the elongation of the mRNA. Indeed, recent evidence suggests that it is relatively common for RNA polymerase to be poised at the promoters, ready to go. For transcription to occur, these nucleosomes need to be modified, and it is possible that the acetylation of histone H3 at positions 9 and 14, coupled with the trimethylation of that histone at position 4, is critical for allowing elongation of the message (Guenther et al. 2007; Li et al. 2007).

Anatomy of the gene: Exons and introns

The second difference between prokaryotic and eukaryotic genes is that eukaryotic genes are not co-linear with their peptide products. Rather, the single nucleic acid strand of eukaryotic mRNA comes from noncontiguous regions on the chromosome. Between **exons**—the regions of DNA that code for a protein*—are intervening sequences called **introns** that have nothing whatsoever to do with the amino acid sequence of the protein. The structure of a typical eukaryotic gene can be illustrated by the human β-globin gene (Figure 2.5). This gene, which encodes part of the hemoglobin protein of the red blood cells, consists of the following elements:

- A **promoter region**, which is responsible for the binding of RNA polymerase and for the subsequent initiation of transcription. The promoter region of the human β-globin gene has three distinct units and extends from 95 to 26 base pairs before ("upstream from")† the transcription initiation site (i.e., from –95 to –26).
- The **transcription initiation site**, which for human β-globin is ACATTTG. This site is often called the **cap sequence** because it represents the 5′ end of the RNA, which will receive a "cap" of modified nucleotides soon after it is transcribed. The specific cap sequence varies among genes.
- The **translation initiation site**, ATG. This codon (which becomes AUG in mRNA) is located 50 base pairs after the transcription initiation site in the human β-globin gene (although this distance differs greatly among different genes). The sequence of 50 base pairs intervening between the initiation points of transcription and translation is the **5′ untranslated region**, often called the **5′ UTR** or **leader sequence.** The 5′ UTR can determine the rate at which translation is initiated.
- The first exon, which contains 90 base pairs coding for amino acids 1–30 of human β-globin protein.
- An intron containing 130 base pairs with no coding sequences for β-globin. However, the structure of this intron is important in enabling the RNA to be processed into mRNA and exit from the nucleus.
- An exon containing 222 base pairs coding for amino acids 31–104.
- A large intron—850 base pairs—having nothing to do with globin protein structure.
- An exon containing 126 base pairs coding for amino acids 105–146 of the protein.
- A **translation termination codon**, TAA. This codon becomes UAA in the mRNA. The ribosome dissociates at this codon, and the protein is released.
- A **3′ untranslated region** (**3′ UTR**) that, although transcribed, is not translated into protein. This region includes the sequence AATAAA, which is needed for **polyadenylation**, the insertion of a "tail" of some 200–300 adenylate residues on the RNA transcript, about 20 bases downstream of the AAUAAA sequence. This polyA tail (1) confers stability on the mRNA, (2) allows the mRNA to exit the nucleus, and (3) permits the mRNA to be translated into protein.
- A **transcription termination sequence**. Transcription continues beyond the AATAAA site for about 1000 nucleotides before being terminated.

The original transcription product is called **nuclear RNA** (**nRNA**), sometimes called *heterogeneous nuclear RNA* (hnRNA) or *pre-messenger RNA* (pre-mRNA). Nuclear RNA contains the cap sequence, the 5′ UTR, exons, introns, and the 3′ UTR (Figure 2.6). Both ends of these transcripts are modified before these RNAs leave the nucleus. A cap consisting of methylated guanosine is placed on the 5′ end of the RNA in opposite polarity to the RNA itself. This means there is no free 5′ phosphate group on the nRNA. The 5′ cap is necessary for the binding of mRNA to the ribosome and for subsequent translation (Shatkin 1976). The 3′ terminus is usually modified in the nucleus by the addition of a polyA tail. The adenylate residues in this tail are put together enzymatically and are added to the transcript; they are not part of the gene sequence. Both the 5′ and 3′ modifications may protect the mRNA from exonucleases that would otherwise digest it (Sheiness and Darnell 1973; Gedamu and Dixon 1978). The modifications thus stabilize the message and its precursor.

As the nRNA leaves the nucleus, its introns are removed and the remaining exons spliced together. In this way the coding regions of the mRNA—i.e., the exons—are brought together to form a single transcript, and this transcript is translated into a protein. The protein can be further modified to make it functional (see Figure 2.6).

*The term *exon* refers to a nucleotide sequence whose RNA "exits" the nucleus. It has taken on the functional definition of a protein-encoding nucleotide sequence. Leader sequences and 3′ UTR sequences are also derived from exons, even though they are not translated into protein.

†By convention, upstream, downstream, 5′, and 3′ directions are specified in relation to the RNA. Thus, the promoter is upstream of the gene, near its 5′ end.

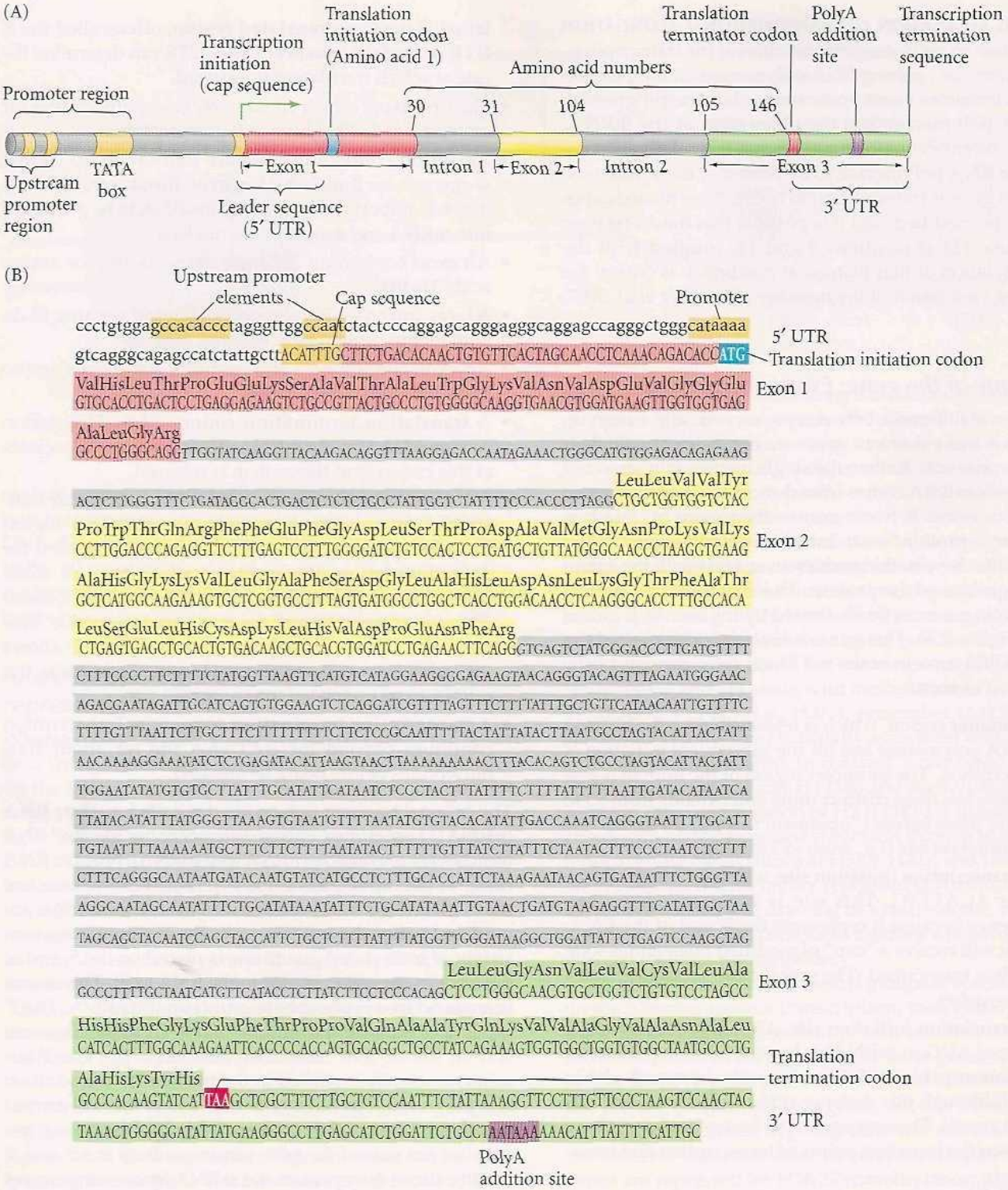


FIGURE 2.5 Nucleotide sequence of the human β-globin gene. (A) Schematic representation of the locations of the promoter region, transcription initiation site (cap sequence), 5′ untranslated region (leader sequence), exons, introns, and 3′ untranslated region. Exons are shown in color; the numbers flanking them indicate the amino acid positions each exon encodes in β-globin. (B) The nucleotide sequence shown from the 5′ end to the 3′ end of the RNA. The colors correspond to their diagrammatic representation in (A). The promoter sequences are boxed, as are the translation initiation and termination codes ATG and TAA. The large capital letters boxed in color are the bases of the exons, with the amino acids for which they code abbreviated above them. Smaller capital letters indicate the intron bases. The codons after the translation termination site exist in β-globin mRNA but are not translated into proteins. Within this group is the sequence thought to be needed for polyadenylation. By convention, only the RNA-like strand of the DNA double helix is shown. (B after Lawn et al. 1980.)

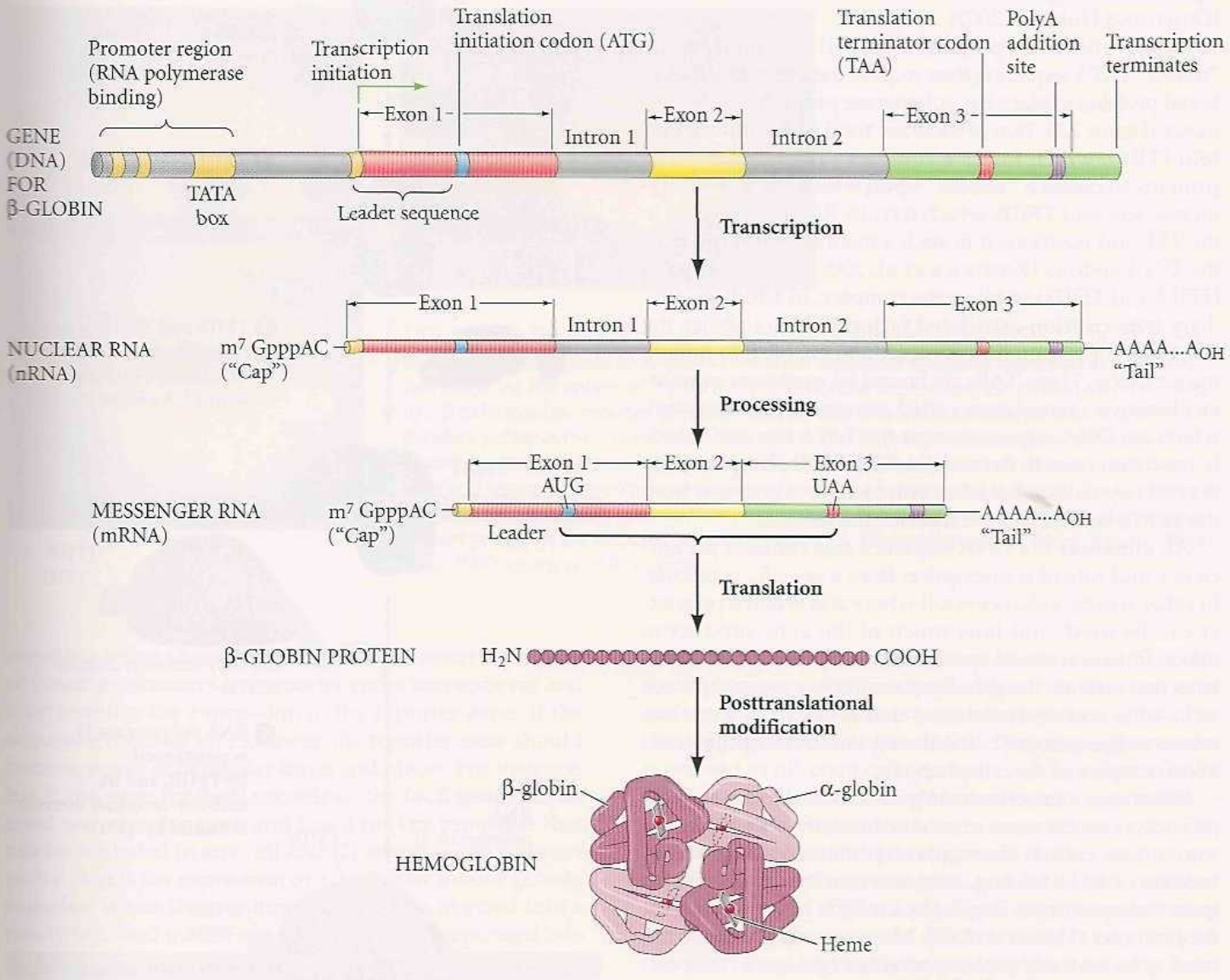


FIGURE 2.6 Summary of steps involved in the production of β-globin and hemoglobin. Transcription of the gene creates a nuclear RNA containing exons and introns, as well as the cap, tail, and 3′ and 5′ untranslated regions. Processing the nuclear RNA into messenger RNA removes the introns. Translation on ribosomes uses the mRNA to encode a protein. The protein is inactive until it is modified and complexed with α-globin and heme to become active hemoglobin (bottom).

Anatomy of the gene: Promoters and enhancers

In addition to the protein-encoding region of the gene, there are regulatory sequences that can be located on either end of the gene (or even within it). These sequences—the promoters and enhancers—are necessary for controlling where and when a particular gene is transcribed.

Promoters are the sites where RNA polymerase binds to the DNA to initiate transcription. Promoters of genes that synthesize messenger RNAs (i.e., genes that encode proteins*) are typically located immediately upstream from the site where the RNA polymerase initiates transcription. Most of these promoters contain the sequence TATA, to which RNA polymerase will be bound. This site, known as the **TATA box**, is usually about 30 base pairs upstream from the site where the first base is transcribed. Since this sequence will appear randomly in the genome at more places than just at promoter sites, other regions flanking it are also important. Many TATA box regions are flanked by

*There are several types of RNA that do *not* encode proteins. These include the ribosomal RNAs and transfer RNAs (which are used in protein synthesis) and the small nuclear RNAs (which are used in RNA processing). In addition, there are regulatory RNAs (such as the microRNAs that we will discuss later in this chapter), which are involved in regulating gene expression and are not translated into peptides.

CpG islands, regions of DNA rich in those two nucleotides (Down and Hubbard 2002).

Eukaryotic RNA polymerases will not bind to the "naked" TATA sequence; they require the presence of additional proteins to place the polymerase properly on the promoter (Figure 2.7). Two of these are the **TATA-binding protein (TBP)**, which forms a complex (TFIID) with other proteins to create a "saddle" upon which the RNA polymerase sits; and **TFIIB**, which recruits RNA polymerase to the TBP and positions it in such a manner that it can read the DNA codons (Kostrewa et al. 2009). Other proteins (TFIIA and TFIIH) stabilize the complex. In addition, auxiliary **transcription-associated factors (TAFs)** stabilize the RNA polymerase on the promoter and enable it to initiate transcription. These TAFs are bound by **upstream promoter elements** (sometimes called *proximal promoter sites*), which are DNA sequences near the TATA box and usually upstream from it. Eventually, TFIIH will phosphorylate the carboxy terminal of RNA polymerase, releasing it from the saddle so that it can transcribe the mRNA.

An **enhancer** is a DNA sequence that controls the efficiency and rate of transcription from a specific promoter. In other words, enhancers tell where and when a promoter can be used, and how much of the gene product to make. Enhancers bind specific **transcription factors**, proteins that activate the gene by (1) recruiting enzymes (such as histone acetyltransferases) that break up the nucleosomes in the area or (2) stabilizing the transcription initiation complex as described above.

Enhancers can activate only *cis*-linked promoters (i.e., promoters on the same chromosome*); therefore they are sometimes called ***cis*-regulatory elements**. However, because of DNA folding, enhancers can regulate genes at great distances (some as great as a million bases away) from the promoter (Visel et al. 2009). Moreover, enhancers do not need to be on the 5′ (upstream) side of the gene; they can be at the 3′ end, or even in the introns (Maniatis et al. 1987). The human β-globin gene has an enhancer in its 3′ UTR. This enhancer sequence is necessary for the temporal- and tissue-specific expression of the β-globin gene in adult red blood cell precursors (Trudel and Constantini 1987).

One of the principal methods of identifying enhancer sequences is to clone DNA sequences flanking the gene of interest and fuse them to **reporter genes** whose products are both readily identifiable and not usually made in the

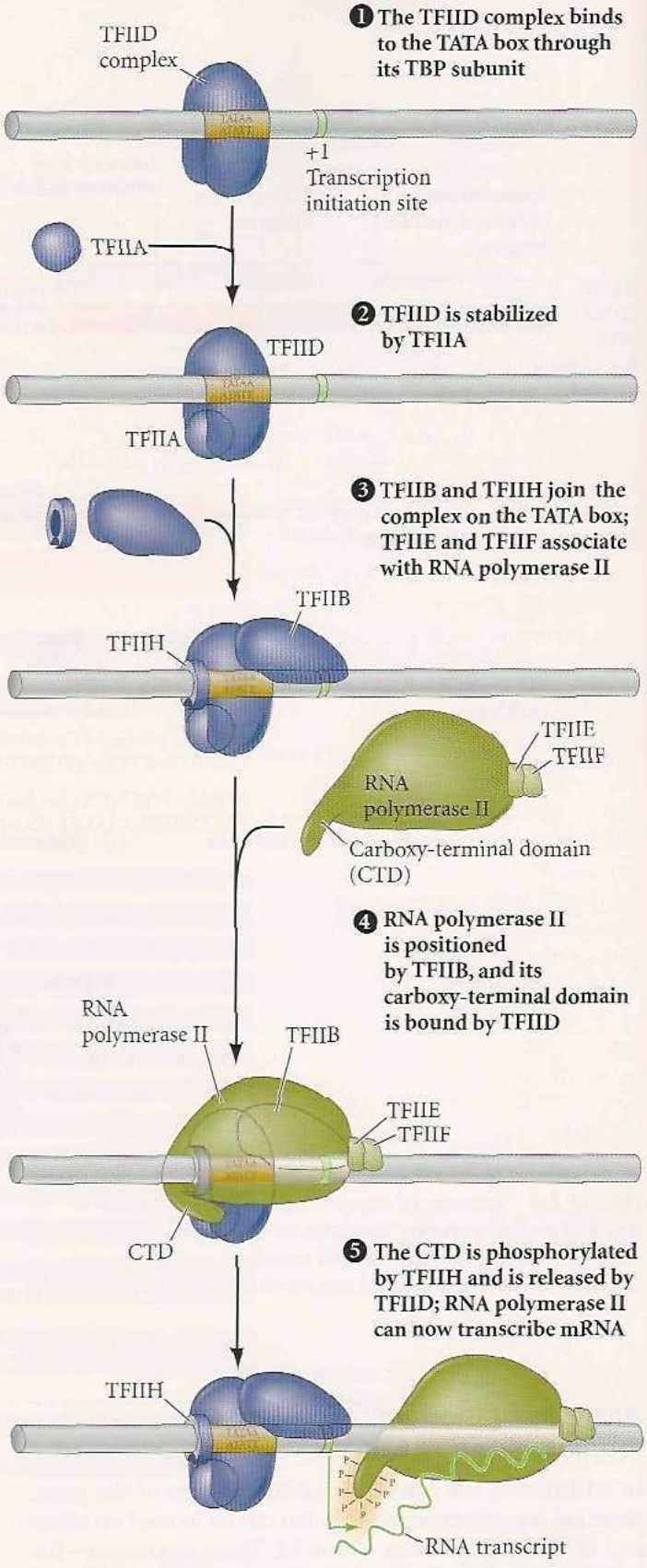


FIGURE 2.7 Formation of the active eukaryotic transcription preinitiation complex. The diagrams represent the formation of the complex that recruits and stabilizes RNA polymerase onto the promoter. *TF* stands for transcription factor; *II* indicates that the factor was first identified as being needed for RNA polymerase II (the RNA polymerase that transcribes protein-encoding genes); and the letters designate the particular active fraction from the phosphocellulose columns used to purify it.

*Cis- and *trans*-regulatory elements are so named by analogy with *E. coli* genetics and organic chemistry. There, *cis*-elements are regulatory elements that reside on the same strand of DNA (*cis*-, "on the same side as"), while *trans*-elements are those that could be supplied from another chromosome (*trans*-, "on the other side of"). The term *cis*-regulatory elements now refers to those DNA sequences that regulate a gene on the same stretch of DNA (i.e., the promoters and enhancers). *Trans*-regulatory factors are soluble molecules whose genes are located elsewhere in the genome and which bind to the *cis*-regulatory elements. They are usually transcription factors or microRNAs.

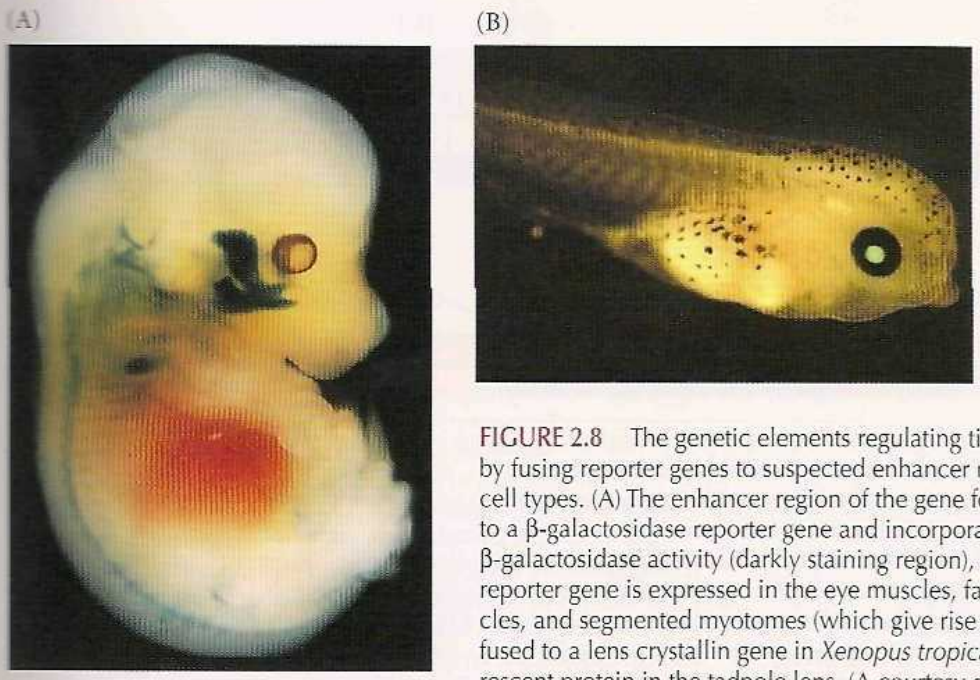


FIGURE 2.8 The genetic elements regulating tissue-specific transcription can be identified by fusing reporter genes to suspected enhancer regions of the genes expressed in particular cell types. (A) The enhancer region of the gene for the muscle-specific protein Myf-5 is fused to a β-galactosidase reporter gene and incorporated into a mouse embryo. When stained for β-galactosidase activity (darkly staining region), the 13.5-day mouse embryo shows that the reporter gene is expressed in the eye muscles, facial muscles, forelimb muscles, neck muscles, and segmented myotomes (which give rise to the back musculature). (B) The *GFP* gene is fused to a lens crystallin gene in *Xenopus tropicalis*. The result is the expression of green fluorescent protein in the tadpole lens. (A courtesy of A. Patapoutian and B. Wold; B from Offield et al. 2000, courtesy of R. Grainger.)

organism being studied. Researchers can insert constructs of possible enhancers and reporter genes into embryos and then monitor the expression of the reporter gene. If the sequence contains an enhancer, the reporter gene should become active at particular times and places. For instance, the *E. coli* gene for β-galactosidase (the ***lacZ* gene**) can be used as a reporter gene and fused to (1) a promoter that can be activated in any cell and (2) an enhancer that normally directs the expression of a particular mouse gene in muscles. When the resulting transgene is injected into a newly fertilized mouse egg and becomes incorporated into its DNA, β-galactosidase will be expressed in the mouse muscle cells. By staining for the presence of β-galactosidase, the expression pattern of that muscle-specific gene can be seen (Figure 2.8A).

Similarly, a sequence flanking a lens crystallin protein in *Xenopus* was shown to be an enhancer. When this sequence was fused to a reporter gene for green fluorescent protein (see Figure 1.17), GFP was expressed only in the lens (Figure 2.8B; Offield et al. 2000). GFP reporter genes are very useful because they can be monitored in live embryos and because the changes in gene expression can be seen in single cells.

ENHANCER MODULARITY The enhancer sequences on the DNA are the same in every cell type; what differs is the combination of transcription factor proteins the enhancers bind. Once bound to enhancers, transcription factors are able to enhance or suppress the ability of RNA polymerase to initiate transcription. Enhancers can bind several transcription factors, and it is the specific *combination* of transcription factors present that allows a gene to be active in a particular cell type. That is, the same transcription factor, in conjunction with different other factors, will activate different promoters in different cells. Moreover, the same gene can have several enhancers, with each enhancer binding transcription factors that enable that same gene to be expressed in different cell types.

Figure 2.9 illustrates this phenomenon for expression of the mouse *Pax6* gene in the cornea and pancreas. The mouse *Pax6* gene (which is expressed in the lens and retina of the eye, in the neural tube, and in the pancreas) has several enhancers (Figure 2.9A). The 5′ regulatory regions of the mouse *Pax6* gene were discovered by taking regions from its 5′ flanking sequence and introns and fusing them to a *lacZ* reporter gene. Each of these transgenes was then microinjected into newly fertilized mouse pronuclei, and the resulting embryos were stained for β-galactosidase (Figure 2.9B; Kammandel et al. 1998; Williams et al. 1998). Analysis of the results revealed that the enhancer farthest upstream from the promoter contains the regions necessary for *Pax6* expression in the pancreas, while a second enhancer activates *Pax6* expression in surface ectoderm (lens, cornea, and conjunctiva). A third enhancer resides in the leader sequence; it contains the sequences that direct *Pax6* expression in the neural tube. A fourth enhancer sequence, located in an intron shortly downstream of the translation initiation site, determines the expression of *Pax6* in the retina. The *Pax6* gene illustrates the principle of enhancer modularity, wherein having multiple, separate enhancers allows a protein to be expressed in several different tissues while not being expressed at all in others.

COMBINATORIAL ASSOCIATION While enhancers are modular *between* enhancers, there are co-dependent units *within* each enhancer. Enhancers contain regions of DNA that

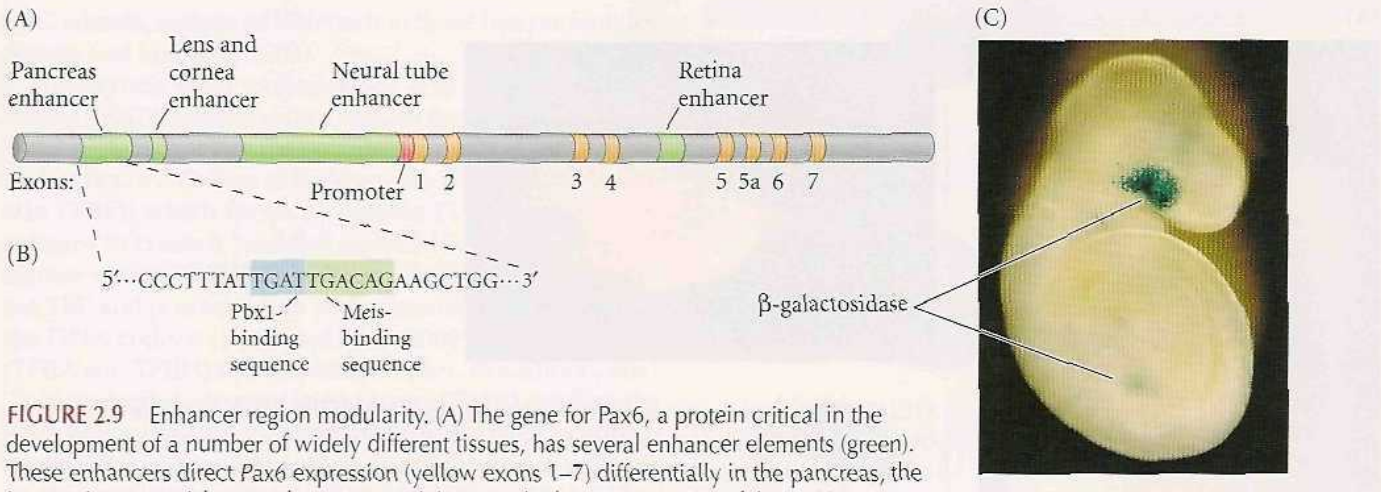


FIGURE 2.9 Enhancer region modularity. (A) The gene for Pax6, a protein critical in the development of a number of widely different tissues, has several enhancer elements (green). These enhancers direct Pax6 expression (yellow exons 1–7) differentially in the pancreas, the lens and cornea of the eye, the retina, and the neural tube. (B) A portion of the DNA sequence of the pancreas-specific enhancer element. This sequence has binding sites for the Pbx1 and Meis transcription factors; both must be present in order to activate the *Pax6* gene in the pancreas. (C) When the gene for bacterial β-galactosidase is fused to *Pax6* enhancers for expression in the pancreas and the lens/cornea, this enzyme (which is easily stained) can be seen in those tissues. (C from Williams et al. 1998, courtesy of R. A. Lang.)

bind transcription factors, and it is this combination of transcription factors that activates the gene. For instance, the pancreas-specific enhancer of the *Pax6* gene has binding sites for the Pbx1 and Meis transcription factors (see Figure 2.9A). Both need to be present in order for the enhancer to activate *Pax6* in the pancreas cells (Zang et al. 2006).

Moreover, the product of the *Pax6* gene encodes a transcription factor that works in combinatorial partnerships with other transcription factors. Figure 2.10 shows two gene regulatory regions that bind Pax6. The first is that of the chick δ1 lens *crystallin* gene (Figure 2.10A; Cvekl and Piatigorsky 1996; Muta et al. 2002). This gene encodes crystallin, a lens protein that is transparent and allows light to reach the retina. A promoter within the *crystallin* gene contains a site for TBP binding, and an upstream promoter element that binds Sp1 (a general transcriptional activator found in all cells). The gene also has an enhancer in its third intron that controls the time and place of crystallin expression. This enhancer has two Pax6-binding sites. The Pax6 protein works with the Sox2 and L-Maf transcription factors to activate the *crystallin* gene only in those head cells that are going to become lens. As we will see in Chapter 10, this involves the cell being head ectoderm (which has Pax6), being in the region of the ectoderm likely to form eyes (L-Maf), and being in contact with the future retinal cells (which induce Sox2 expression; Kamachi et al. 1998).

Meanwhile, another set of regulatory regions that use Pax6 are the enhancers regulating the transcription of the genes for insulin, glucagon, and somatostatin in the pancreas (Figure 2.10B). Here, Pax6 is also essential for gene expression, and it works in cooperation with other transcription factors such as Pdx1 (specific for the pancreatic region of the endoderm) and Pbx1 (Andersen et al. 1999; Hussain and Habener 1999). In the absence of Pax6 (as in the homozygous *small eye* mutation in mice and rats), the endocrine cells of the pancreas do not develop properly and the production of hormones by those cells is deficient (Sander et al. 1997; Zhang et al. 2002).

There are other genes that are activated by Pax6 binding, and one of them is the *Pax6* gene itself. Pax6 protein can bind to a *cis*-regulatory element of the *Pax6* gene (Plaza et al. 1993). This means that once the *Pax6* gene is turned on, it will continue to be expressed, even if the signal that originally activated it is no longer given.

One can see that the genes for specific proteins use numerous transcription factors in various combinations. Thus, *enhancers are modular* (such that the *Pax6* gene is expressed in the eye, pancreas, and nervous system, as shown in Figure 2.9); but *within each cis-regulatory module, transcription factors work in a combinatorial fashion* (such that Pax6, L-Maf, and Sox2 are all needed for the transcription of crystallin in the lens). The combinatorial association of transcription factors on enhancers leads to the spatiotemporal output of any particular gene (see Davidson 2006; Zinzen et al. 2009).

Transcription factor function

Natalie Angier (1992) has written, "A series of new discoveries suggests that DNA is more like a certain type of politician, surrounded by a flock of protein handlers and advisers that must vigorously massage it, twist it, and on occasion, reinvent it before the grand blueprint of the body can make any sense at all." These "handlers and advisers" are the transcription factors. These factors can be grouped together in families based on similarities in structure (Table 2.1). The transcription factors within such a family share a common framework in their DNA-binding sites, and slight

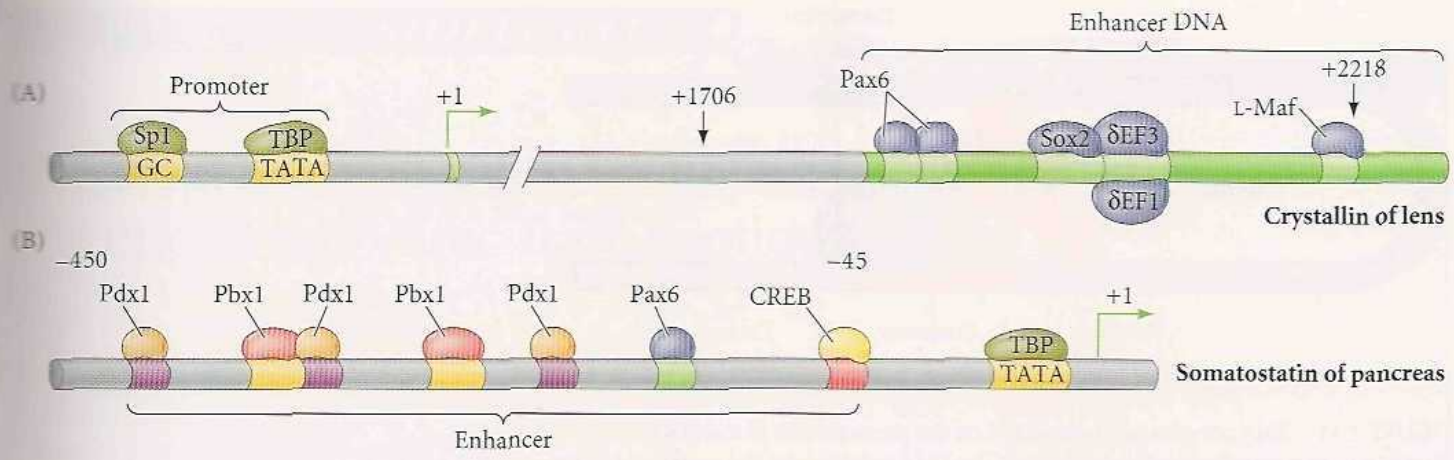


FIGURE 2.10 Modular transcriptional regulatory regions using Pax6 as an activator. (A) Promoter and enhancer of the chick δ1 lens *crystallin* gene. Pax6 interacts with two other transcription factors, Sox2 and L-Maf, to activate this gene. The protein δEF3 binds factors that permit this interaction; δEF1 binds factors that inhibit it. (B) Promoter and enhancer of the rat somatostatin gene. Pax6 activates this gene by cooperating with the Pbx1 and Pdx1 transcription factors. (A after Cvekl and Piatigorsky 1996; B after Andersen et al. 1999.)

differences in the amino acids at the binding site can cause the binding site to recognize different DNA sequences.

As we have already seen, enhancers function by binding transcription factors, and each enhancer can have binding sites for several transcription factors. Transcription factors bind to the enhancer DNA with one part of the protein and use other sites on the protein to interact with one another to recruit histone-modifying enzymes.

For example, the association of the Pax6, Sox2, and L-Maf transcription factors in lens cells recruits a histone acetyltransferase that can transfer acetyl groups to the histones and dissociate the nucleosomes in that area (Yang et al. 2006). Similarly, when MITF, a transcription factor essential for ear development and pigment production, binds to its specific DNA sequence, it also binds a (different) histone acetyltransferase that also facilitates the dissociation of nucleosomes (Ogryzko et al. 1996; Price et al. 1998). And the Pax7 transcription factor that activates muscle-specific genes binds to the enhancer region of these genes within the muscle precursor cells. Pax7 then recruits a histone methyltransferase that methylates the lysine in the fourth position of histone H3 (H3K4), resulting in the trimethylation of this lysine and the activation of transcription (McKinnell et al. 2008). The displacement of nucleosomes along the DNA makes it possible for the transcription fac-

TABLE 2.1 Some major transcription factor families and subfamilies

Family	Representative transcription factors	Some functions
Homeodomain:		
Hox	Hoxa1, Hoxb2, etc.	Axis formation
POU	Pit1, Unc-86, Oct-2	Pituitary development; neural fate
LIM	Lim1, Forkhead	Head development
Pax	Pax1, 2, 3, 6, etc.	Neural specification; eye development
Basic helix-loop-helix (bHLH)	MyoD, MITF, daughterless	Muscle and nerve specification; *Drosophila* sex determination; pigmentation
Basic leucine zipper (bZip)	C/EBP, AP1	Liver differentiation; fat cell specification
Zinc finger:		
Standard	WT1, Krüppel, Engrailed	Kidney, gonad, and macrophage development; *Drosophila* segmentation
Nuclear hormone receptors	Glucocorticoid receptor, estrogen receptor, testosterone receptor, retinoic acid receptors	Secondary sex determination; craniofacial development; limb development
Sry-Sox	Sry, SoxD, Sox2	Bend DNA; mammalian primary sex determination; ectoderm differentiation

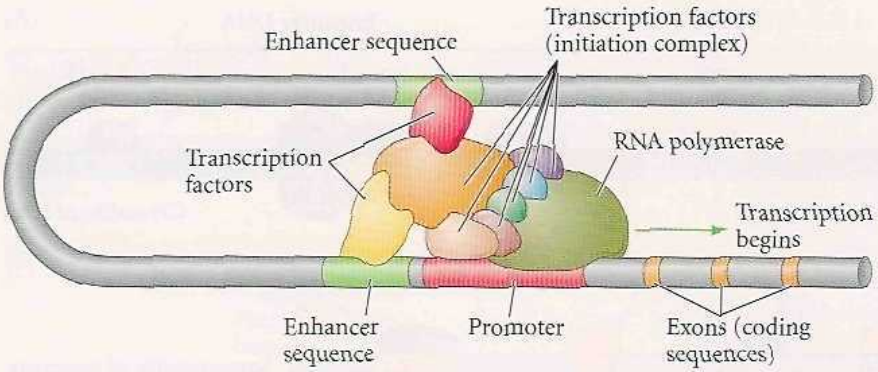


FIGURE 2.11 RNA polymerase is stabilized on the promoter site of the DNA by transcription factors recruited by the enhancers. The TATA sequence at the promoter binds a protein that serves as a "saddle" for RNA polymerase. However, RNA polymerase would not remain bound long enough to initiate transcription were it not for the stabilization by the transcription factors.

tors to find their binding sites (Adkins et al. 2004; Li et al. 2007).

In addition to recruiting nucleosome modifying enzymes, transcription factors can also work by stabilizing the transcription preinitiation complex that enables RNA polymerase to bind to the promoter (Figure 2.11). For instance, MyoD, a transcription factor that is critical for muscle cell development (see Chapter 11), stabilizes TFIIB, which supports RNA polymerase at the promoter site (Heller and Bengal 1998). Indeed, MyoD plays several roles in activating gene expression, since it also can bind histone acetyltransferases that initiate nucleosome remodeling and dissociation (Cao et al. 2006).

One of the important consequences of the combinatorial association of transcription factors is **coordinated gene expression**. The simultaneous expression of many cell-specific genes can be explained by the binding of transcription factors by the enhancer elements. For example, many genes that are specifically activated in the lens contain an enhancer that binds Pax6. This means that all the other transcription factors might be assembled at the enhancer, but until Pax6 binds, they cannot activate the gene. Similarly, many of the co-expressed muscle-specific genes contain enhancers that bind the MEF2 transcription factor; and the enhancers on genes encoding pigment-producing enzymes bind MITF (see Davidson 2006).

TRANSCRIPTION FACTOR DOMAINS Transcription factors have three major domains. The first is a **DNA-binding domain** that recognizes a particular DNA sequence in the enhancer. Figure 2.12 shows a model of such a domain in the Pax6 protein described earlier (see Figure 2.9). The second is a ***trans*-activating domain** that activates or suppresses the transcription of the gene whose promoter or enhancer it has bound. Usually, this *trans*-activating domain enables the transcription factor to interact with the proteins involved in binding RNA polymerase (such as TFIIB or TFIIE; see Sauer et al. 1995) or with enzymes that modify histones. In addition, there may be a **protein-protein interaction domain** that allows the transcription factor's activity to be modulated by TAFs or other transcription factors.

MITF, a transcription factor essential for ear development and pigment production, has a protein-protein interaction domain that enables it to dimerize with another MITF protein (Ferré-D'Amaré et al. 1993). The resulting homodimer (i.e., two identical protein molecules bound

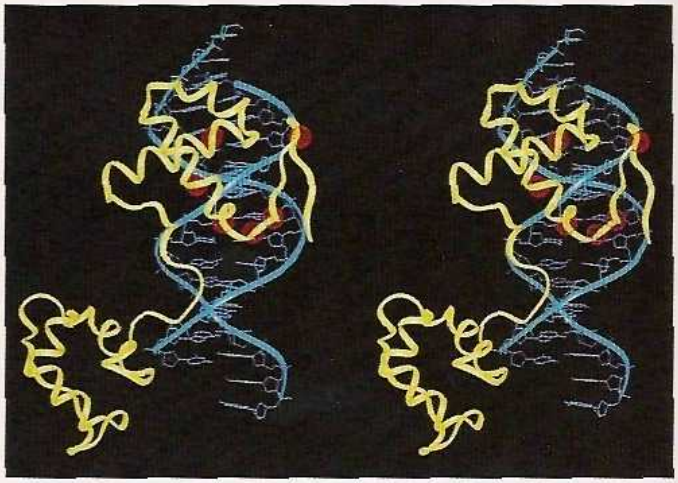


FIGURE 2.12 Stereoscopic model of Pax6 protein binding to its enhancer element in DNA. The DNA-binding region of Pax6 is shown in yellow; the DNA double helix is blue. Red dots indicate the sites of loss-of-function mutations in the *Pax6* gene that give rise to nonfunctional Pax6 proteins. It is worth trying to cross your eyes to see the central three-dimensional figure. (From Xu et al. 1995; photograph courtesy of S. O. Pääbo.)

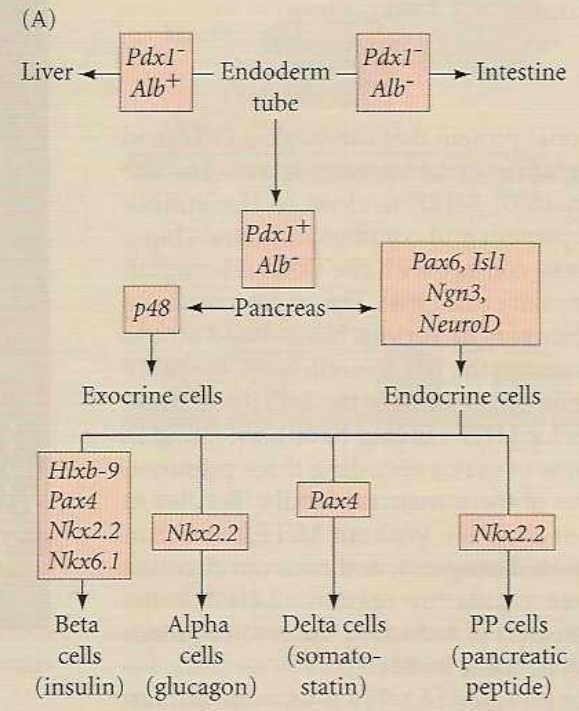
SIDELIGHTS & SPECULATIONS

Reprogramming Cells: Changing Cell Differentiation through Embryonic Transcription Factors

The importance and power of transcription factors were elegantly demonstrated when Zhou and colleagues (2008) used three transcription factors to convert *exocrine* pancreatic cells (which make amylase, chymotrypsin, and other digestive enzymes) into insulin-secreting *endocrine* pancreatic β cells. The researchers infected the pancreases of living 2-month-old mice with harmless viruses containing the genes for three transcription factors: Pdx1, Ngn3, and Mafa.

The Pdx1 protein stimulates the outgrowth of the digestive tube that results in the pancreatic buds. This protein is found throughout the pancreas and is critical in specifying that organ's endocrine cells, as well as in activating genes that encode endocrine proteins (see Figure 2.10B). Ngn3 is a transcription factor found in endocrine, but not exocrine, pancreatic cells. Mafa, a transcription factor regulated by glucose levels, is found only in pancreatic β cells (i.e., those cells that make insulin) and can activate transcription of the insulin gene.

Pdx1, Ngn3, and Mafa activate other transcription factors that work in concert to turn a pancreatic endodermal cell into an insulin-secreting β cell. Zhou and colleagues found that, of all the transcription factor genes tested, these three were the only ones that were crucial for the conversion (Figure 2.13). Converted pancreas cells looked identical to normal β cells, and like normal β cells, they were capable of secreting both insulin and a blood vessel-inducing factor. The converted cells retained their new properties for months. Moreover, mice whose normal β cells were destroyed by a particular drug developed diabetes, just like the Type 1 diabetes seen when human adults lose β cells. This diabetes could be cured by injecting the mice with viruses containing the three transcription factors. When this was done, about 20% of the exocrine pancreatic cells became β cells and secreted insulin.



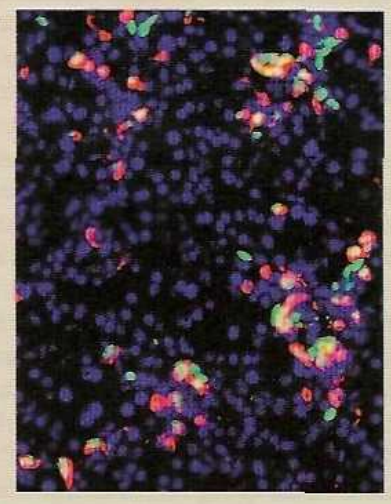


Figure 2.13 Pancreatic lineage and transcription factors. (A) Pdx1 protein is expressed in pancreatic lineages. Several transcription factors, including Ngn3, distinguish the endocrine lineage. Several other transcription factors, including Mafa, are found in the β cells that produce insulin. (B) New pancreatic β cells arise in adult mouse pancreas in vivo after viral delivery of three transcription factors (Ngn3, Pdx1, and Mafa). Virally infected exocrine cells are detected by their expression of nuclear green fluorescent protein. Newly induced β cells are recognized by insulin staining (red). Their overlap produces yellow. The nuclei of all pancreatic cells are stained blue. (B courtesy of D. Melton.)

This study opens the door to an entire new field of regenerative medicine, illustrating the possibilities of changing one adult cell type into another by using the transcription factors that had made the new cell type in the embryo.

See VADE MECUM
Transdetermination in *Drosophila*

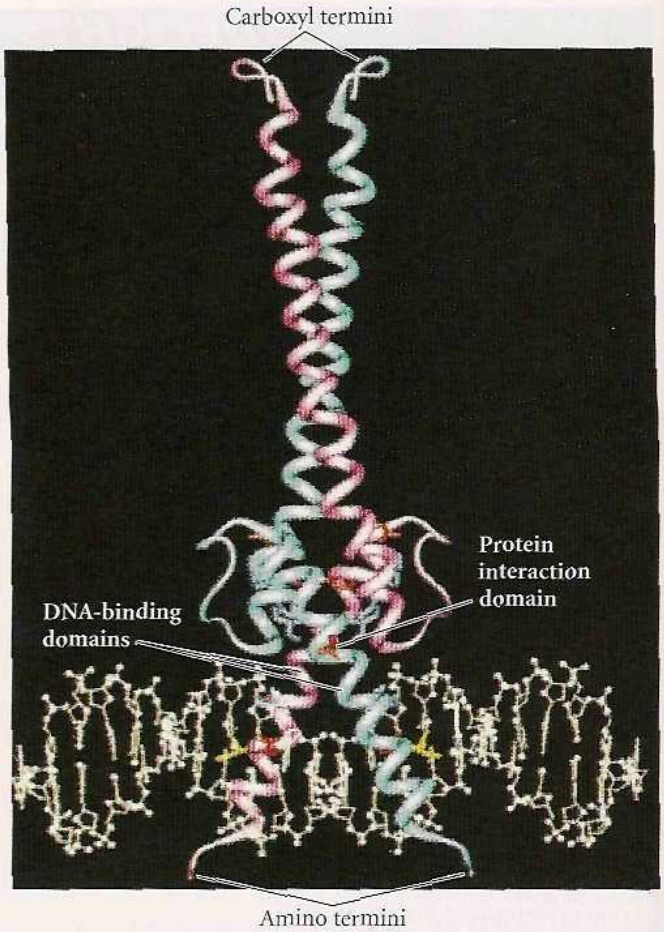


FIGURE 2.14 Three-dimensional model of the homodimeric transcription factor MITF (one protein shown in red, the other in blue) binding to a promoter element in DNA (white). The amino termini are located at the bottom of the figure and form the DNA-binding domain that recognizes an 11-base-pair sequence of DNA having the core sequence CATGTG. The protein-protein interaction domain is located immediately above. MITF has the basic helix-loop-helix structure found in many transcription factors. The carboxyl end of the molecule is thought to be the *trans*-activating domain that binds the p300/CBP co-activator protein. (From Steingrímsson et al. 1994, courtesy of N. Jenkins.)

together) is a functional protein that can bind to DNA and activate the transcription of certain genes (Figure 2.14). The DNA-binding domain of MITF is close to the amino-terminal end of the protein and contains numerous basic amino acids that make contact with the DNA (Hemesath et al. 1994; Steingrímsson et al. 1994). This assignment was confirmed by the discovery of various human and mouse mutations that map within the DNA-binding site for MITF and which prevent the attachment of the MITF protein to the DNA. Sequences for MITF binding have been found in the regulatory regions of genes encoding three pigment-cell-specific enzymes of the tyrosinase family (Bentley et al. 1994; Yasumoto et al. 1997). Without MITF, these proteins are not synthesized properly, and melanin pigment is not made. These *cis*-regulatory regions all contain the same 11-base-pair sequence, including the core sequence (CATGTG) that is recognized by MITF.

The third functional region of MITF is its *trans*-activating domain. This domain includes a long stretch of amino acids in the center of the protein. When the MITF dimer is bound to its target sequence in the enhancer, the *trans*-activating region is able to bind a TAF, p300/CBP. The p300/CPB protein is a histone acetyltransferase enzyme that can transfer acetyl groups to each histone in the nucleosomes (Ogryzko et al. 1996; Price et al. 1998). Acetylation of the nucleosomes destabilizes them and allows the genes for pigment-forming enzymes to be expressed.

EPIGENETIC MEMORY: KEEPING THE RIGHT GENES ON OR OFF

The modifications of histones can also signal the recruitment of the proteins that can retain the memory of transcriptional state from generation to generation through mitosis. These are the proteins of the Trithorax and Polycomb families. When bound to the nucleosomes of active genes, **Trithorax** proteins keep these genes active, whereas **Polycomb** proteins, which bind to condensed nucleosomes, keep the genes in an inactive state.

The Polycomb proteins fall into two categories that act sequentially in repression. The first set has histone methyltransferase activities that methylate lysines H3K27 and H3K9 to repress gene activity. In many organisms, this repressive state is stabilized by the activity of a second set of Polycomb factors, which bind to the methylated tails of histone 3 and keep the methylation active and also methylate adjacent nucleosomes, thereby forming tightly packed repressive complexes (Grossniklaus and Paro 2007; Margueron et al. 2009).

The Trithorax proteins help retain the memory of activation; they act to counter the effect of the Polycomb proteins. Trithorax proteins can modify the nucleosomes or alter their positions on the chromatin, allowing transcription factors to bind to the DNA previously covered by them. Other Trithorax proteins keep the H3K4 lysine trimethylated (preventing its demethylation into a dimethylated, repressive, state; Tan et al. 2008).

PIONEER TRANSCRIPTION FACTORS: BREAKING THE SILENCE

Finding a promoter is not easy, because the DNA is usually so wound up that the promoter sites are not accessible. Indeed, more than *6 feet* of DNA is packaged into chromosomes of each human cell nucleus (Schones and Zhao 2008).

How can a transcription factor find its binding site, given that the enhancer might be covered by nucleosomes? Several studies have identified certain transcription factors that penetrate repressed chromatin and bind to their enhancer DNA sequences (Cirillo et al. 2002;

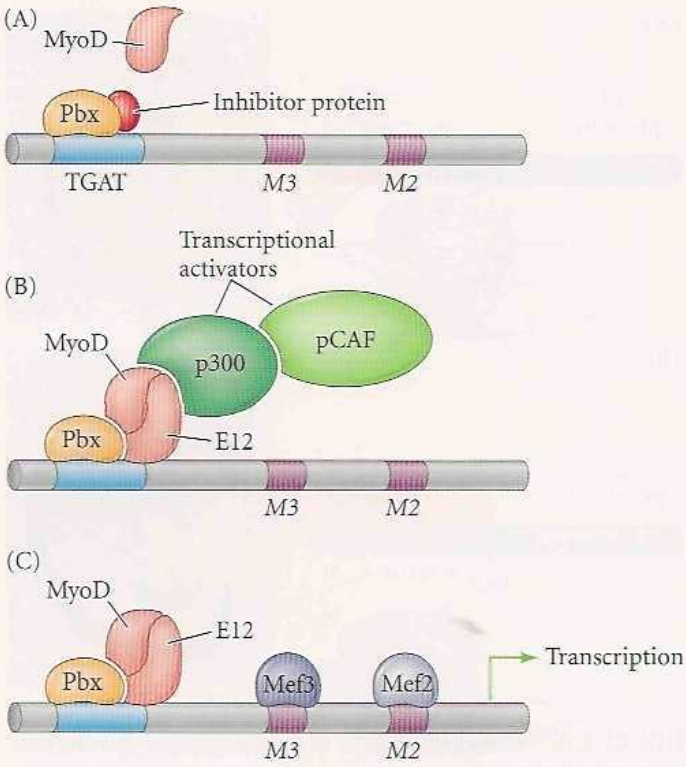


FIGURE 2.15 Model for the role of the "pioneer" transcription factor Pbx in aligning the muscle-specific transcription factor MyoD on DNA. (A) Pbx protein recognizes its DNA binding site (TGAT), even within nucleosome-rich chromatin. Pbx probably binds to transcriptional inhibitors. (B) MyoD, complexed with its E12 cofactor, is able to bind to Pbx, replacing the transcriptional inhibitors. MyoD then binds to its recognition element on the DNA. (C) The MyoD/E12 complex can then recruit the histone acetyltransferases and nucleosome remodeling compounds that make the chromatin accessible to other transcription factors (Mef3 and Mef2) and to RNA polymerase. (After Tapscott 2005.)

Berkes et al. 2004). They have been called "pioneer" transcription factors, and they appear to be critical in establishing certain cell lineages. One of these transcription factors is FoxA1, which binds to certain enhancers and opens up the chromatin to allow other transcription factors access to the promoter (Lupien et al. 2008). FoxA proteins remain bound to the DNA during mitosis, providing a mechanism to re-establish normal transcription in presumptive liver cells (Zaret et al. 2008). Another pioneer transcription factor is the Pax7 protein mentioned above. It activates muscle-specific gene transcription in a population of muscle stem cells by binding to its DNA recognition sequence and being stabilized there by dimethylated H3K4 on the nucleosomes. It then recruits the histone methyltransferase that converts the dimethylated H3K4 into the trimethylated H3K4 associated with active transcription (McKinnell et al. 2008).

Another pioneer transcription factor in muscle development is Pbx. Members of the Pbx family are made in every cell, and they are able to find their appropriate sites even in highly compacted chromatin. Pbx appears to be used as a "molecular beacon" for another muscle-determining transcription factor, MyoD (mentioned earlier). MyoD is critical for initiating muscle development in the embryo, activating hundreds of genes that are involved with establishing the muscle phenotype. However, MyoD is not able to bind to DNA without the help of Pbx proteins, which bind to DNA elements adjacent to the DNA sequence recognized by MyoD (Figure 2.15A). Berkes and colleagues (2004) have shown that MyoD (when complexed with another transcription factor, E12) can bind to the Pbx protein and align itself on its target DNA sequence (Figure 2.15B). Once bound there, the E12 protein recruits histone acetyltransferases and nucleosome remodeling complexes to open up the DNA on those genes (Figure 2.15C).

SILENCERS **Silencers** are DNA regulatory elements that actively repress the transcription of a particular gene. They can be viewed as "negative enhancers." For instance, in the mouse, there is a DNA sequence that prevents a promoter's activation in any tissue *except* neurons. This sequence, given the name **neural restrictive silencer element** (**NRSE**), has been found in several mouse genes whose expression is limited to the nervous system: those encoding synapsin I, sodium channel type II, brain-derived neurotrophic factor, Ng-CAM, and L1. The protein that binds to the NRSE is a zinc finger transcription factor called **neural restrictive silencer factor** (**NRSF**). (It is also called REST). NRSF appears to be expressed in every cell that is *not* a mature neuron (Chong et al. 1995; Schoenherr and Anderson 1995).

To test the hypothesis that the NRSE sequence is necessary for the normal repression of neural genes in non-neural cells, transgenes were made by fusing a β-galactosidase (*lacZ*) gene with part of the *L1* neural cell adhesion gene. (L1 is a protein whose function is critical for brain development, as we will see in later chapters.) In one case, the *L1* gene, from its promoter through the fourth exon, was fused to the *lacZ* sequence. A second transgene was made just like the first, except that the NRSE was deleted from the *L1* promoter. The two transgenes were separately inserted into the pronuclei of fertilized oocytes, and the resulting transgenic mice were analyzed for β-galactosidase expression (Kallunki et al. 1995, 1997). In embryos receiving the complete transgene (which included the NSRE), expression was seen only in the nervous system (Figure 2.16A). In mice whose transgene lacked the NRSE sequence, however, expression was seen in the heart, the limb mesenchyme and limb ectoderm, the kidney mesoderm, the ventral body wall, and the cephalic mesenchyme (Figure 2.16B).

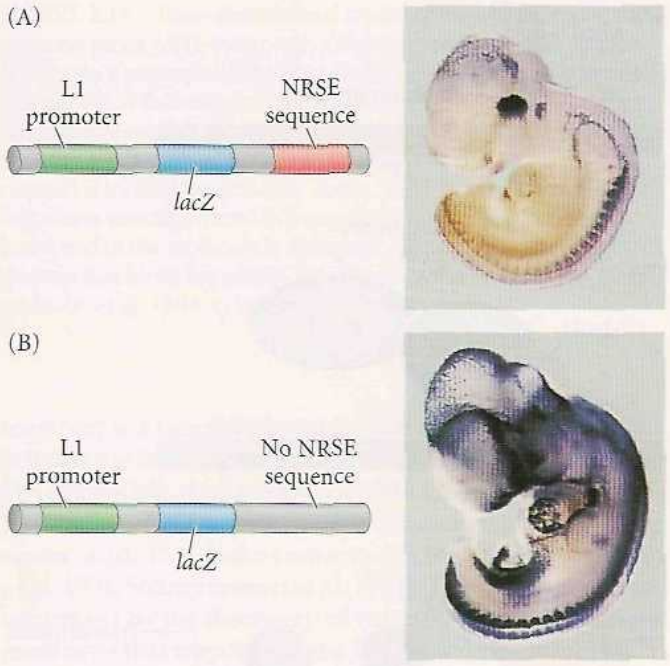


FIGURE 2.16 Silencers. Analysis of β-galactosidase staining patterns in 11.5-day embryonic mice. (A) Embryo containing a transgene composed of the L1 promoter, a portion of the *L1* gene, and a *lacZ* gene fused to the second exon (which contains the NRSE region). (B) Embryo containing a similar transgene but lacking the NRSE sequence. The dark areas reveal the presence of β-galactosidase (the *lacZ* product). (Photographs from Kallunki et al. 1997.)

DNA Methylation and the Control of Transcription

How does a pattern of gene transcription become stable? How can a lens cell continue to remain a lens cell and not activate muscle-specific genes? How can cells undergo rounds of mitosis and still maintain their differentiated characteristics? The answer appears to be **DNA methylation**. We have already discussed *histone* methylation and its importance for transcription. Now we look at how the *DNA itself* can be methylated to regulate transcription. Generally speaking, the promoters of inactive genes become methylated at certain cytosine residues, and the resulting methylcytosine stabilizes nucleosomes and prevents transcription factors from binding.

It is often assumed that a gene contains exactly the same nucleotides whether it is active or inactive; that is, a β-globin gene that is activated in a red blood cell precursor has the same nucleotides as the inactive β-globin gene in a fibroblast or retinal cell of the same animal. However, it turns out there is in fact a subtle difference. In 1948, R. D. Hotchkiss discovered a "fifth base" in DNA, 5-methylcytosine. In vertebrates, this base is made enzymatically after DNA is replicated. At this time, about 5% of the cytosines in mammalian DNA are converted to 5-methylcytosine (Figure 2.17A). This conversion can occur only when the cytosine residue is followed by a guanosine—in other words, at a CpG sequence (as we will soon see, this restriction is important). Numerous studies have shown that the degree to which the cytosines of a gene are methylated can control the level of the gene's transcription. Cytosine methylation appears to be a major mechanism of transcriptional regulation among vertebrates; however, some other species (*Drosophila* and nematodes among them) do not methylate their DNA.

In vertebrates, the presence of methylated cytosines in a gene's promoter correlates with the repression of transcription from that gene. In developing human and chick red blood cells, for example, the DNA of the globin gene promoters is almost completely unmethylated, whereas the same promoters are highly methylated in cells that do not produce globins. Moreover, the methylation pattern changes during development (Figure 2.17B). The cells that produce hemoglobin in the human embryo have unmethylated promoters in the genes encoding the ε-globins ("embryonic globin chains") of embryonic hemoglobin. These promoters become methylated in the fetal tissue, as the genes for fetal-specific γ-globin (rather than the embryonic chains) become activated (van der Ploeg and Flavell 1980; Groudine and Weintraub 1981; Mavilio et al. 1983).

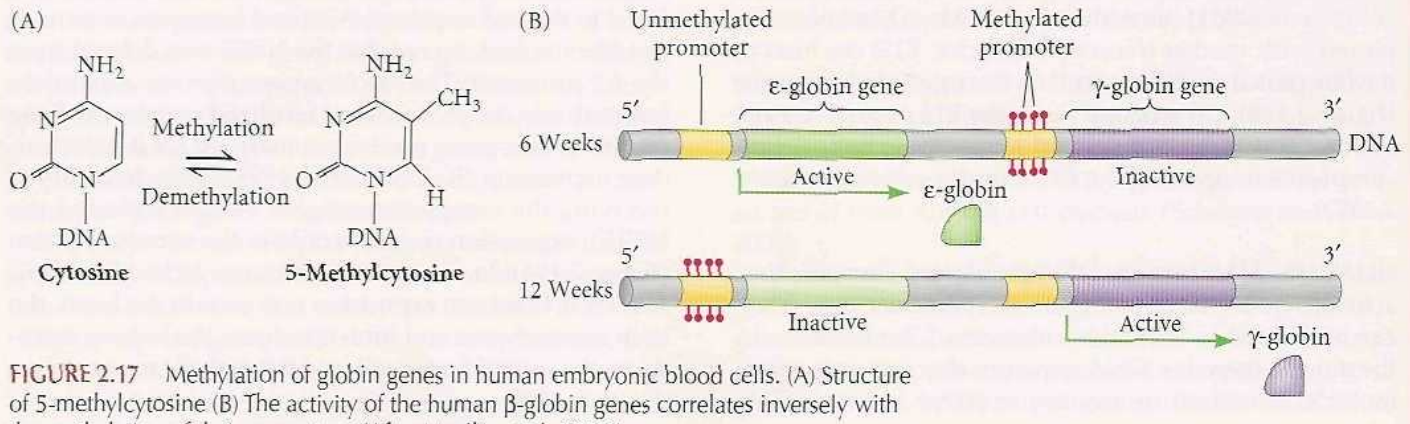


FIGURE 2.17 Methylation of globin genes in human embryonic blood cells. (A) Structure of 5-methylcytosine (B) The activity of the human β-globin genes correlates inversely with the methylation of their promoters. (After Mavilio et al. 1983.)

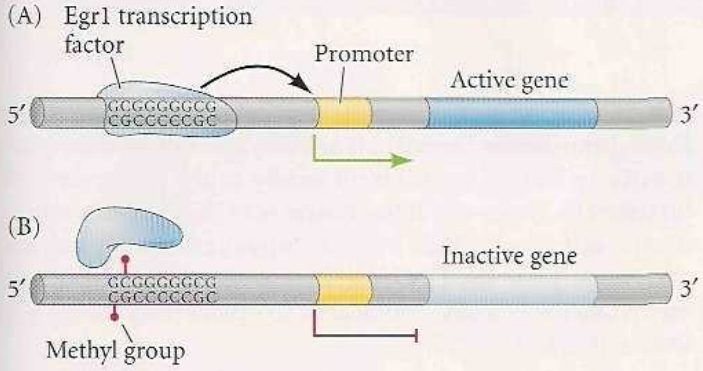


FIGURE 2.18 DNA methylation can block transcription by preventing transcription factors from binding to the enhancer region. (A) The Egr1 transcription factor can bind to specific DNA sequences such as 5′...GCGGGGGCG...3′, helping to activate transcription of those genes. (B) If the first cytosine residue is methylated, however, Egr1 will not bind and the gene will remain repressed. (After Weaver et al. 2005.)

Similarly, when fetal globin gives way to adult (β) globin, promoters of the fetal (γ) globin genes become methylated.

Mechanisms by which DNA methylation blocks transcription

DNA methylation appears to act in two ways to repress gene expression. First, it can block the binding of transcription factors to enhancers. Several transcription factors can bind to a particular sequence of unmethylated DNA, but they cannot bind to that DNA if one of its cytosines is methylated (Figure 2.18). Second, a methylated cytosine can recruit the binding of proteins that facilitate the methylation or deacetylation of histones, thereby stabilizing the nucleosomes. For instance, methylated cytosines in DNA can bind particular proteins such as MeCP2. Once connected to a methylated cytosine, MeCP2 binds to histone deacetylases and histone methyltransferases, which, respectively, remove acetyl groups (Figure 2.19A) and add methyl groups (Figure 2.19B) on the histones. As a result, the nucleosomes form tight complexes with the DNA and don't allow other transcription factors and RNA polymerases to find the genes. Other proteins, such as HP1 and histone H1, will bind and aggregate methylated histones (Fuks 2005; Rupp and Becker 2005). In this way, repressed chromatin becomes associated with regions where there are methylated cytosines.

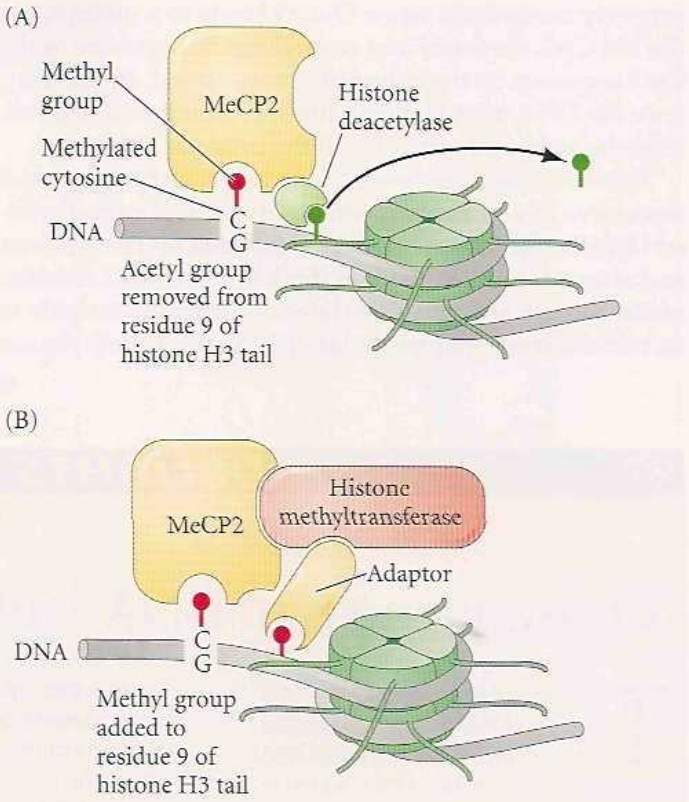


FIGURE 2.19 Modifying nucleosomes through methylated DNA. MeCP2 recognizes the methylated cytosines of DNA. It binds to the DNA and is thereby able to recruit histone deacetylases (which take acetyl groups off the histones) (A) or histone methyltransferases (which add methyl groups to the histones) (B). Both modifications promote the stability of the nucleosome and the tight packing of DNA, thereby repressing gene expression in these regions of DNA methylation. (After Fuks 2003.)

Inheritance and stabilization of DNA methylation patterns

Another enzyme recruited to the chromatin by MeCP2 is DNA methyltransferase-3 (Dnmt3). This enzyme methylates previously unmethylated cytosines on the DNA. In this way, a relatively large region can be repressed. The newly established methylation pattern is then transmitted to the next generation by DNA methyltransferase-1 (Dnmt1). This enzyme recognizes methyl cytosines on one strand of DNA and places methyl groups on the newly synthesized strand opposite it (Figure 2.20; see Bird 2002; Burdge et al. 2007). This is why it is necessary for the C to be next to a G in the sequence. Thus, in each cell division, the pattern of DNA methylation can be maintained. The newly synthesized (unmethylated) strand will become

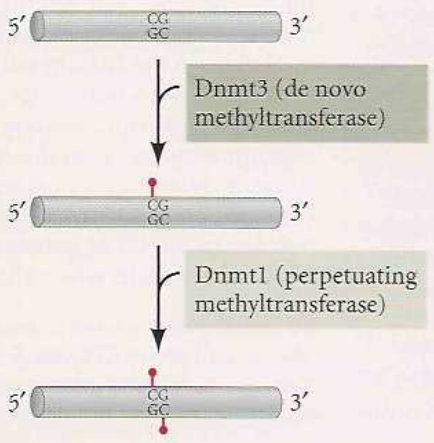


FIGURE 2.20 Two DNA methyltransferases are critically important in modifying DNA. The "de novo" methyltransferase Dnmt3 can place a methyl group on unmethylated cytosines. The "perpetuating" methyltransferase, Dnmt1, recognizes methylated Cs on one strand and methylates the C on the CG pair on the opposite strand.

properly methylated when Dnmt1 binds to a methylC on the old CpG sequence and methylates the cytosine of the CpG sequence on the complementary strand. In this way, once the DNA methylation pattern is established in a cell, it can be stably inherited by all the progeny of that cell.

Reinforcement between repressive chromatin and repressive DNA has also been observed. Just as methylated DNA is able to attract proteins that deacetylate histones and attract H1 linker histones (both of which will stabilize nucleosomes), so repressive states of chromatin are able to recruit enzymes that methylate DNA. DNA methylation patterns during gametogenesis depend in part on the DNA methyltransferase Dnmt3L. It actually has lost its enzymatic activity, but it can still bind avidly to the amino end of histone H3. However, if the lysine at H3K4 is methylated, it will not bind. Once bound, however, it will recruit and/or activate the DNA methyltransferase Dnmt3A2 to methylate the cytosines on nearby CG pairs (Fan et al. 2007; Ooi et al. 2007).

See WEBSITE 2.7
Silencing large blocks of chromatin

SIDELIGHTS & SPECULATIONS

Consequences of DNA Methylation

The control of transcription through DNA methylation has many consequences in addition to cell differentiation. DNA methylation has explained X chromosome inactivation and DNA imprinting. Moreover, as we will see in Chapter 18, improper DNA methylation (when the wrong cytosines are methylated or demethylated) has been associated with aging, cancers, and the poor health of cloned animals.

X chromosome inactivation

In *Drosophila*, nematodes, and mammals, females are characterized as having two X chromosomes per cell, while males are characterized as having a single X chromosome per cell. Unlike the Y chromosome, the X chromosome contains thousands of genes that are essential for cell activity. Yet despite the female's cells having double the number of X chromosomes, male and female cells contain approximately equal amounts of X chromosome-encoded gene products. This equalization phenomenon is called **dosage compensation**, and it can be accomplished in three ways (Migeon 2002). In *Drosophila*, the transcription rate of the male X chromosomes is doubled so that the single male X chromosome makes the same amount of transcript as the two female X chromosomes (Lucchesi and Manning 1987). This is accomplished by acetylation of the nucleosomes throughout the male's X chromosomes, which gives RNA polymerase more efficient access to that chromosome's promoters (Akhtar et al. 2000; Smith et al. 2001). In *C. elegans*, both X chromosomes are partially repressed (Chu et al. 2002) so that the male and female* products of the X chromosomes are equalized.

In mammals, dosage compensation occurs through the inactivation of one X chromosome in each female cell. Thus, each mammalian somatic cell, whether male or female, has only one functioning X chromosome. This phenomenon is called **X chromosome inactivation**. The chromatin of the inactive X chromosome is converted into **heterochromatin**—chromatin that remains condensed throughout most of the cell cycle and replicates later than most of the other chromatin (the **euchromatin**) of the nucleus. This was first shown by Mary Lyon (1961), who observed coat color patterns in mice. If a mouse is heterozygous for an autosomal gene controlling hair pigmentation, then it resembles one of its two parents, or has a color intermediate between the two. In either case, the mouse is a single color. But if a female mouse is heterozygous for a pigmentation gene *on the X chromosome*, a different result is seen: patches of one parental color alternate with patches of the other parental color. This also explains why calico and tortoiseshell cats† are normally female: their coat color alleles (black and orange) are on the X chromosome (Centerwall and Benirscke 1973).

Lyon proposed the following hypothesis to account for these results:

1. Very early in the development of female mammals, both X chromosomes are active. As development proceeds, one X chromosome is inactivated in each cell (Figure 2.21A).
2. This inactivation is random. In some cells, the paternally derived X chromosome is inactivated; in other cells, the maternally derived X chromosome is shut down.
3. This process is irreversible. Once a particular X chromosome (either the one derived from the mother or the one derived from the father) has been inactivated in a cell, the same X chromosome is inactivated in all of that cell's progeny (Figure 2.21B,C). Because X inactivation happens relatively early in development, an entire region of cells derived from a single cell may all have the same X chromosome inactivated. Thus, all tissues in female mammals are mosaics of two cell types.

*As we will see in Chapter 5, the "female" is actually a hermaphrodite capable of making both sperm and eggs.

†Although the terms *calico* and *tortoiseshell* are sometimes used synonymously, tortoiseshell coats are a patchwork of black and orange only; calico cats usually have white patches—i.e., patches with no pigment—as well (see Figure 2.2).

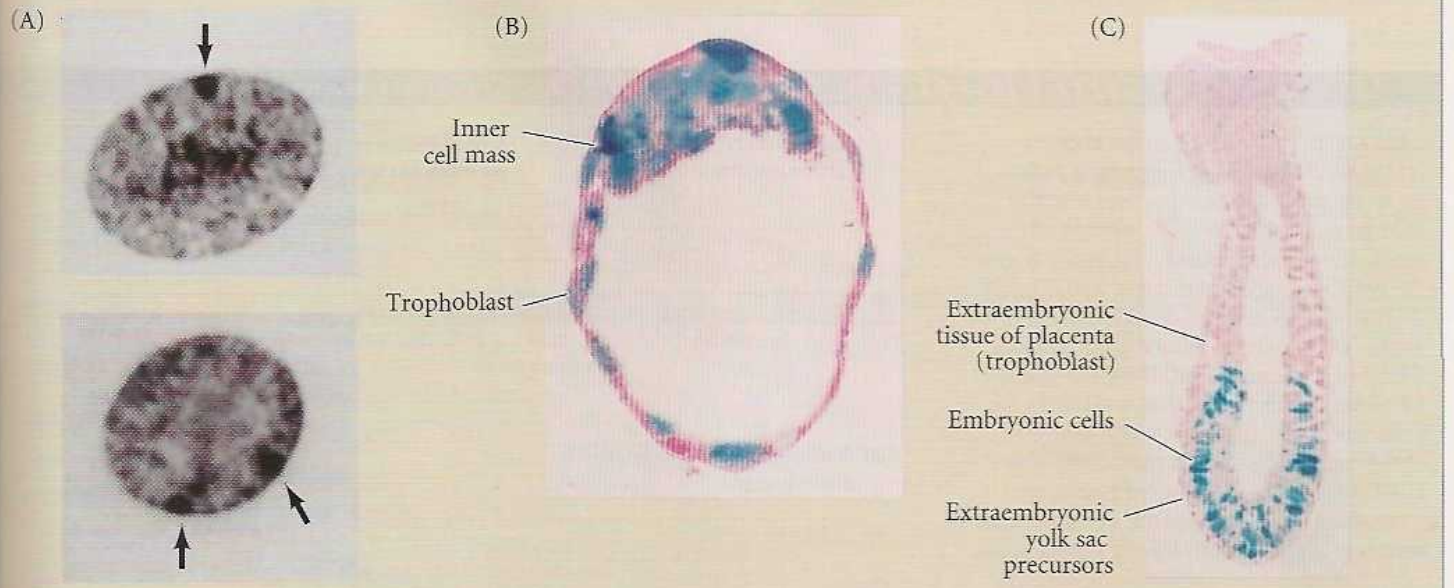


Figure 2.21 X chromosome inactivation in mammals. (A) Inactivated X chromosomes, or *Barr bodies*, in the nuclei of human oral epithelial cells. The top cell is from a normal XX female and has a single Barr body (arrow). In the lower cell, from a female with three X chromosomes, two Barr bodies can be seen. In both cases, only one X chromosome per cell is active. (B,C) The paternally derived X chromosome of this mouse embryo contained a *lacZ* transgene. Those cells in which the chromosome is active make β-galactosidase and stain blue. The other cells are counterstained and appear pink. (B) In the early blastocyst stage (day 4), both X chromosomes are active in all cells. (C) At day 6, random inactivation of one of the chromosomes occurs. Embryonic cells in which the maternal X is active appear pink, while those where the paternal X is active stain blue. In mouse (but not human) trophoblast, the paternally derived X chromosome is preferentially inactivated, so the trophoblast cells are uniformly pink. (A courtesy of M. L. Barr; B,C from Sugimoto et al. 2000, courtesy of N. Takagi.)

The inactivation of the X chromosome is complicated; indeed, it is a bottleneck that many female embryos do not get through (Migeon 2007). The mechanisms of X chromosome inactivation appear to differ between mammalian groups, but these mechanisms converge in that they all inactivate an X chromosome by methylating promoters. In mice and humans, the promoter regions of numerous genes are methylated on the inactive X chromosome and unmethylated on the active X chromosome (Wolf et al. 1984; Keith et al. 1986; Migeon et al. 1991). The memory of this "X inactivation" is transmitted to the progeny of the cells by successive DNA methylation through Dnmt1 (see above).

Genomic Imprinting

The second phenomenon explained by DNA methylation is genomic imprinting. It is usually assumed that the genes one inherits from one's father and the genes one inherits from one's mother are equivalent. In fact, the basis for Mendelian ratios (and the Punnett square analyses used to teach them) is that it does not matter whether the genes came from the sperm or from the egg. But in mammals, there are at least 80 genes for which it *does* matter.* Here, the chromosomes from the male and the female are not equivalent. In these cases, only the sperm-derived or only the egg-derived allele of the gene is expressed. This means that a severe or lethal condition arises if a mutant allele is derived from one parent, but that the same mutant allele will have no deleterious effects if inherited from the other parent. In some of these cases, the nonfunctioning gene has been rendered inactive by DNA methylation. (This means that a mammal must have both a male parent and a female parent. Unlike sea urchins, flies, and frogs, mammals cannot experience parthenogenesis, or "virgin birth.") The methyl groups are placed on the DNA during spermatogenesis and oogenesis by a series of enzymes that first take the existing methyl groups off the chromatin and then place new sex-specific ones on the DNA (Ciccone et al. 2009).

As described earlier in this chapter, methylated DNA is associated with stable DNA silencing, either (1) by interfering with the binding of gene-activating transcription factors or (2) by recruiting repressor proteins that stabilize nucleosomes in a restrictive manner along the gene. The presence of a methyl group in the minor groove of DNA can prevent certain transcription factors from binding to the DNA, thereby preventing the gene from being activated (Watt and Molloy 1988).

For example, during early embryonic development in mice, the *Igf2* gene (for insulin-like growth factor) is active only from the father's chromosome 7. The egg-derived *Igf2* gene does not function during embryonic development. This is because the CTCF protein is an inhibitor that can block the promoter from getting activation signals from enhancers. It binds to a region near the *Igf2* gene in females because this region is not methylated. Once bound, it prevents the maternally derived *Igf2* gene from functioning. In the sperm-derived chromosome 7, the region where CTCF would bind is methylated. CTCF cannot bind and the gene is not inhibited from functioning (Figure 2.22; Bartolomei et al. 1993; Ferguson-Smith et al. 1993; Bell and Felsenfeld 2000).

In humans, misregulation of *Igf2* methylation causes Beckwith-

*A list of imprinted mouse genes is maintained at http://www.har.mrc.ac.uk/research/genomic_imprinting/introduction.html

(Continued on next page)

SIDELIGHTS & SPECULATIONS (Continued)

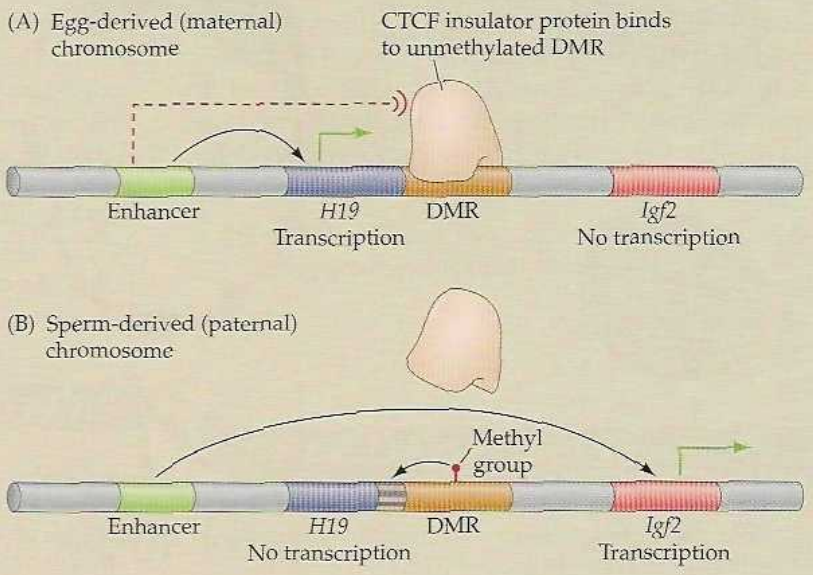


Figure 2.22 Regulation of the imprinted *Igf2* gene in the mouse. This gene is activated by an enhancer element it shares with the *H19* gene. The differentially methylated region (DMR) is a sequence located between the enhancer and the *Igf2* gene, and is found on both sperm- and egg-derived chromosomes. (A) In the egg-derived chromosome, the DMR is unmethylated. The CTCF insulator protein binds to the DMR and blocks the enhancer signal. (B) In the sperm-derived chromosome, the DMR is methylated. The CTCF insulator protein cannot bind to the methylated sequence, and the signal from the enhancer is able to activate *Igf2* transcription.

Wiedemann growth syndrome. Interestingly, although DNA methylation is the mechanism for imprinting this gene in both mice and humans, the mechanisms responsible for the differential *Igf2* methylation between sperm and egg appear to be very different in the two species (Ferguson-Smith et al. 2003; Walter and Paulsen 2003).

Also in humans, the loss of a particular segment of the long arm of chromosome 15 results in different phenotypes, depending on whether the loss is in the male- or the female-derived chromosome (Figure 2.23A). If the chromosome with the defective or missing segment comes from the father, the child is born with Prader-Willi syndrome, a disease associated with mild mental retardation, obesity, small gonads, and short stature. If the defective or missing segment comes from the mother, the child has Angelman syndrome, characterized by severe mental retardation, seizures, lack of speech, and inappropriate laughter (Knoll et al. 1989; Nicholls 1998). The imprinted genes in this region are *SNRPN* and *UBE3A*. In the egg-derived chromosome, *UBE3A* is activated and *SNRPN* is turned off, while in the sperm-derived chromosome, *SNRPN* is activated and *UBE3A* is turned off (Figure 2.23B). The expression of either maternal or paternal loci on human chromosome 15 also depends on methylation differences at specific regions in the chromosome that regulate these genes (Zeschingk et al. 1997; Ferguson-Smith and Surani 2001; Walter and Paulsen 2003).

Differential methylation is one of the most important mechanisms of epigenetic changes. It provides a reminder that an organism cannot be explained solely by its genes. One needs knowledge of developmental parameters as well as genetic ones.

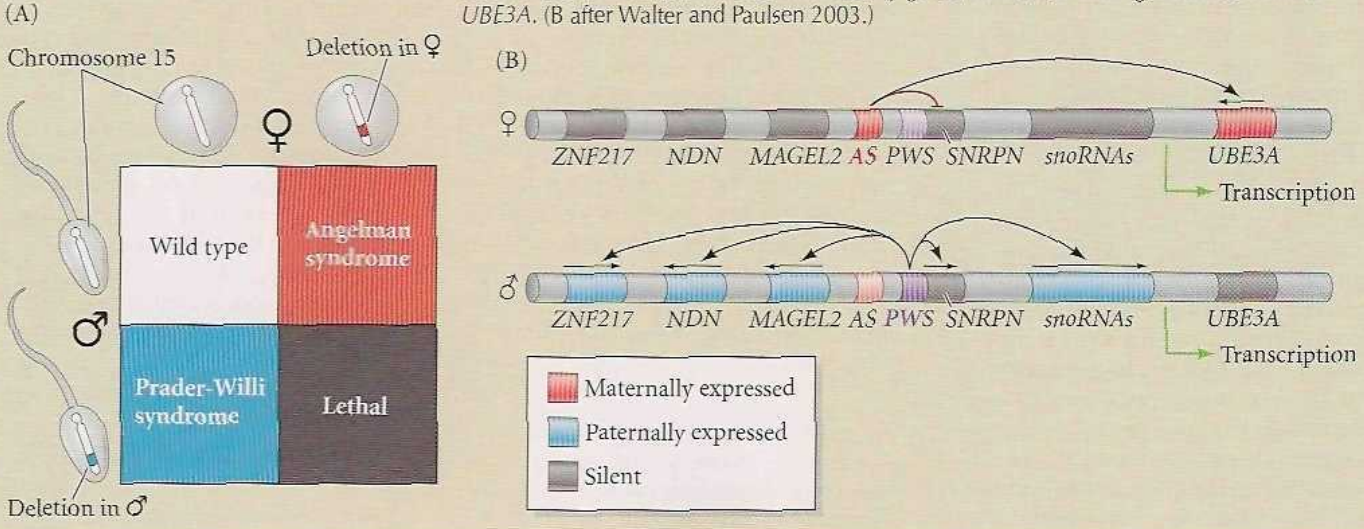
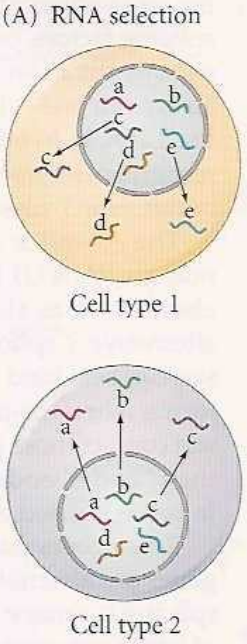


Figure 2.23 Inheritance patterns for Prader-Willi and Angelman syndromes. (A) A region in the long arm of chromosome 15 contains the genes whose absence causes both these syndromes. However, the two conditions are imprinted in reverse fashion. In Prader-Willi syndrome, the paternal genes are active; in Angelman syndrome, the maternal genes are active. (B) Some of the genes and the "inactivation centers" where methylation occurs on this chromosomal region. In the maternal chromosome, the *AS* inactivation center activates *UBE3A* and suppresses *SNRPN*. Conversely, on the paternal chromosome, the *PWS* inactivation center activates *SNRPN* and several other nearby genes, as well as making antisense RNA to *UBE3A*. (B after Walter and Paulsen 2003.)



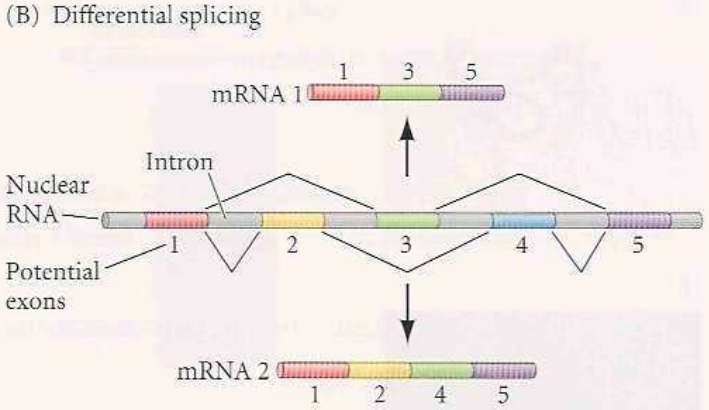


FIGURE 2.24 Roles of differential RNA processing during development. By convention, splicing paths are shown by fine V-shaped lines. (A) RNA selection, whereby the same nuclear RNA transcripts are made in two cell types, but the set that becomes cytoplasmic messenger RNA is different. (B) Differential splicing, whereby the same nuclear RNA is spliced into different mRNAs by selectively using different exons.

Differential RNA Processing

The regulation of gene expression is not confined to the differential transcription of DNA. Even if a particular RNA transcript is synthesized, there is no guarantee that it will create a functional protein in the cell. To become an active protein, the RNA must be (1) processed into a messenger RNA by the removal of introns, (2) translocated from the nucleus to the cytoplasm, and (3) translated by the protein-synthesizing apparatus. In some cases, the synthesized protein is not in its mature form and must be (4) posttranslationally modified to become active. Regulation during development can occur at any of these steps.

The essence of differentiation is the production of different sets of proteins in different types of cells. In bacteria, differential gene expression can be effected at the levels of transcription, translation, and protein modification. In eukaryotes, however, another possible level of regulation exists—namely, control at the level of RNA processing and transport. There are two major ways in which differential RNA processing can regulate development. The first involves "censorship"—selecting which nuclear transcripts are processed into cytoplasmic messages. Different cells select different nuclear transcripts to be processed and sent to the cytoplasm as messenger RNA. Thus, the same pool of nuclear transcripts can give rise to different populations of cytoplasmic mRNAs in different cell types (Figure 2.24A).

The second mode of differential RNA processing is the *splicing* of mRNA precursors into messages that specify different proteins by using different combinations of potential exons. If an mRNA precursor had five potential exons, one cell type might use exons 1, 2, 4, and 5; a different type might use exons 1, 2, and 3; and yet another cell type might use all five (Figure 2.24B). Thus a single gene can produce an entire family of proteins.

Control of early development by nuclear RNA selection

In the late 1970s, numerous investigators found that mRNA was not the primary transcript from the genes. Rather, the initial transcript is a nuclear RNA (nRNA). This nRNA is usually many times longer than the corresponding mRNA because nRNA contains introns that get spliced out during the passage from nucleus to cytoplasm (see Figure 2.6). Originally, investigators thought that whatever RNA was transcribed in the nucleus was processed into cytoplasmic mRNA. But studies of sea urchins showed that different cell types could be *transcribing* the same type of nuclear RNA, but *processing* different subsets of this population into mRNA in different types of cells (Kleene and Humphreys 1977, 1985). Wold and her colleagues (1978) showed that sequences present in sea urchin blastula *messenger* RNA, but absent in gastrula and adult tissue mRNA, were nonetheless present in the *nuclear* RNA of the gastrula and adult tissues.

More genes are transcribed in the nucleus than are allowed to become mRNAs in the cytoplasm. This "censoring" of RNA transcripts has been confirmed by probing for the introns and exons of specific genes. Gagnon and his colleagues (1992) performed such an analysis on the transcripts from the *SpecII* and *CyIIIa* genes of the sea urchin *Strongylocentrotus purpuratus*. These genes encode calcium-binding and actin proteins, respectively, which are expressed only in a particular part of the ectoderm of the sea urchin larva. Using probes that bound to an exon (which is included in the mRNA) and to an intron (which is not included in the mRNA), they found that these genes were being transcribed not only in the ectodermal cells, but also in the mesoderm and endoderm. The analysis of the *CyIIIa* gene showed that the concentration of introns was the same in both the gastrula ectoderm and the mesoderm/endoderm samples, suggesting that this gene was being transcribed at the same rate in the nuclei of all cell types, but was made into cytoplasmic mRNA only in ectodermal cells (Figure 2.25). The unprocessed nRNA for *CyIIIa* is degraded while still in the nuclei of the endodermal and mesodermal cells.

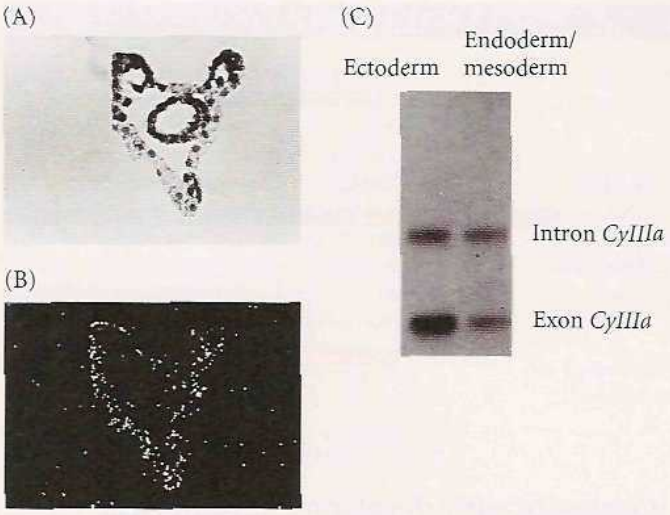


FIGURE 2.25 Regulation of ectoderm-specific gene expression by RNA processing. (A,B) *CyIIIa* mRNA is seen by autoradiography to be present only in the ectoderm. (A) Phase contrast micrograph. (B) In situ hybridization using a probe that binds to a *CyIIIa* exon. (C) The *CyIIIa* nuclear transcript, however, is found in both ectoderm and endoderm/mesoderm. The left lane of the gel represents RNA isolated from the gastrula ectodermal tissue; the right lane represents RNA isolated from endodermal and mesodermal tissues. The upper band is the RNA bound by a probe that binds to an intron sequence (which should be found only in the nucleus) of *CyIIIa*. The lower band represents the RNA bound by a probe complementary to an exon sequence. The presence of the intron indicates that the *CyIIIa* nuclear RNA is being made in both groups of cells, even if the mRNA is seen only in the ectoderm. (From Gagnon et al. 1992, courtesy of R. and L. Angerer.)

Creating families of proteins through differential nRNA splicing

Alternative nRNA splicing is a means of producing a wide variety of proteins from the same gene. The average vertebrate nRNA consists of several relatively short exons (averaging about 140 bases) separated by introns that are usually much longer. Most mammalian nRNAs contain numerous exons. By splicing together different sets of exons, different cells can make different types of mRNAs, and hence, different proteins. Recognizing a sequence of nRNA as either an exon or an intron is a crucial step in gene regulation.

Alternative nRNA splicing is based on the determination of which sequences will be spliced out as introns. This can occur in several ways. Most genes contain "consensus sequences" at the 5′ and 3′ ends of the introns. These sequences are the "splice sites" of the intron. The splicing of nRNA is mediated through complexes known as **spliceosomes** that bind to the splice sites. Spliceosomes are made up of small nuclear RNAs (snRNAs) and proteins called **splicing factors** that bind to splice sites or to the areas adjacent to them. By their production of specific splicing factors, cells can differ in their ability to recognize a sequence as an intron. That is to say, a sequence that is an *exon* in one cell type may be an *intron* in another (Figure 2.26A,B). In other instances, the factors in one cell might recognize different 5′ sites (at the beginning of the intron) or different 3′ sites (at the end of the intron; Figure 2.26C,D).

The 5′ splice site is normally recognized by small nuclear RNA U1 (U1 snRNA) and splicing factor 2 (SF2; also known as alternative splicing factor). The choice of alternative 3′ splice sites is often controlled by which splice site can best bind a protein called U2AF. The spliceosome forms when the proteins that accumulate at the 5′ splice site contact those proteins bound to the 3′ splice site. Once the 5′ and 3′ ends are brought together, the intervening intron is excised and the two exons are ligated together.

Researchers estimate that approximately 92% of human genes are alternatively spliced, and that such alternative splicing is a major way by which the rather limited number of genes can create a much larger array of proteins (Wang et al. 2008). The deletion of certain potential exons in some cells but not in others enables one gene to create a family of closely related proteins. Instead of one gene-one polypeptide, one can have one gene-one family of proteins. For instance, alternative RNA splicing enables the gene for α-tropomyosin to encode brain, liver, skeletal muscle, smooth muscle, and fibroblast forms of this protein (Breitbart et al. 1987). The nuclear RNA for α-tropomyosin contains 11 potential exons, but different sets of exons are used in different cells (Figure 2.27). Such different proteins encoded by the same gene are called **splicing isoforms** of the protein.

In some instances, alternatively spliced RNAs yield proteins that play similar yet distinguishable roles in the same cell. Different isoforms of the WT1 protein perform different functions in the development of the gonads and kidneys. The isoform without the extra exon functions as a transcription factor during kidney development, whereas the isoform containing the extra exon appears to be involved in splicing different nRNAs and may be critical in testis development (Hammes et al. 2001; Hastie 2001).

The *Bcl-x* gene provides a good example of how alternative nRNA splicing can make a huge difference in a protein's function. If a particular DNA sequence is used as an exon, the "large Bcl-X protein," or Bcl-X_L, is made (see Figure 2.26C). This protein inhibits programmed cell death. However, if this sequence is seen as an intron, the "small Bcl-X protein" (Bcl-X_S) is made, and this protein *induces* cell death. Many tumors have a higher than normal amount of Bcl-X_L.

If you get the impression from this discussion that a gene with dozens of introns could create literally thousands of different, related proteins through differential splicing, you are probably correct. The current champion at making multiple proteins from the same gene is the *Drosophila Dscam1* gene. This gene encodes a membrane receptor protein involved in preventing dendrites from the same neuron

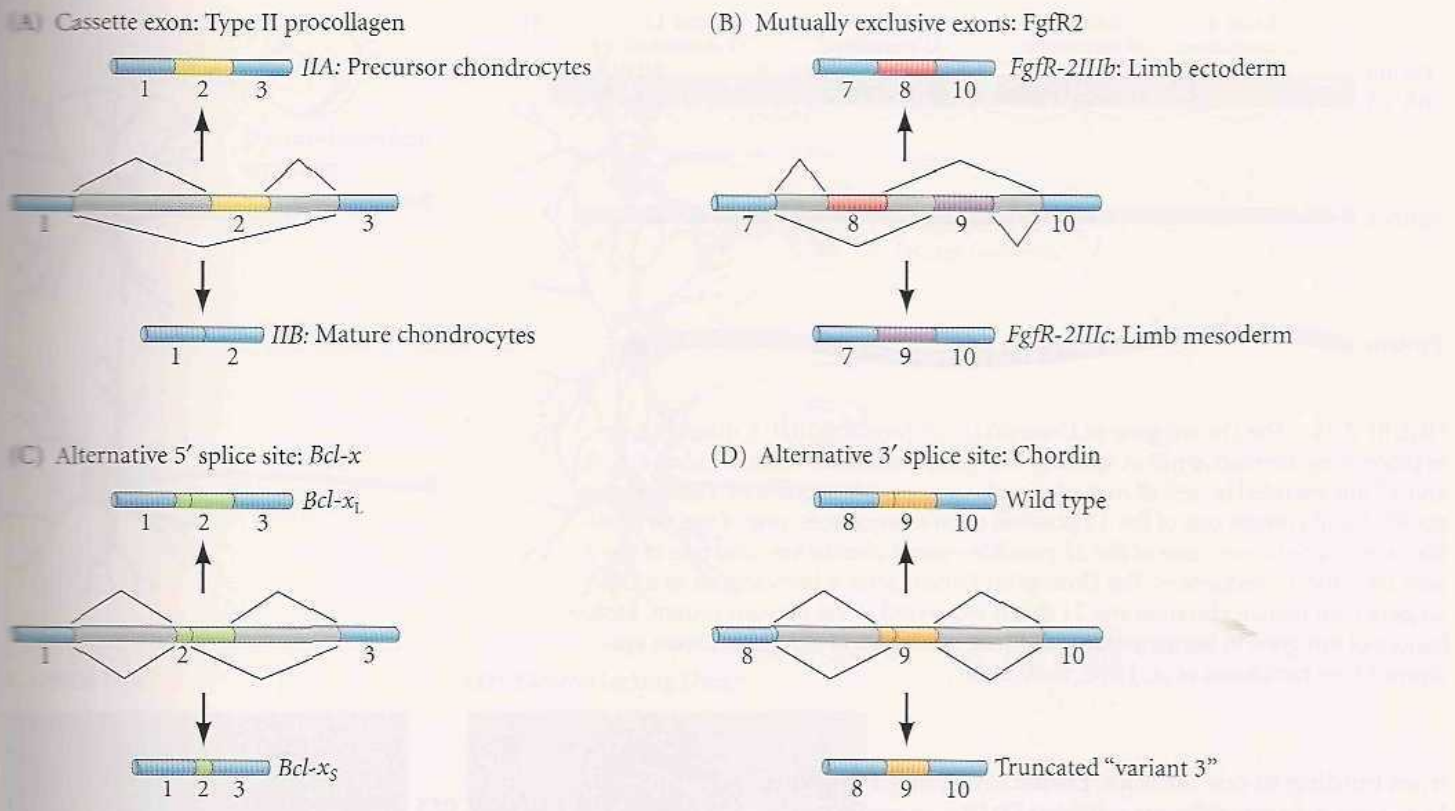


FIGURE 2.26 Some examples of alternative RNA splicing. Blue and colored portions of the bars represent exons; gray represents introns. Alternative splicing patterns are shown with V-shaped lines. (A) A "cassette" (yellow) that can be used as exon or removed as an intron distinguishes the type II collagen types of chondrocyte precursors and mature chondrocytes (cartilage cells). (B) Mutually exclusive exons distinguish fibroblast growth factor receptors found in the limb ectoderm from those found in the limb mesoderm. (C) Alternative 5′ splice site selection, such as that used to create the large and small isoforms of the protein Bcl-X. (D) Alternative 3′ splice sites are used to form the normal and truncated forms of chordin. (After McAlinden et al. 2004.)

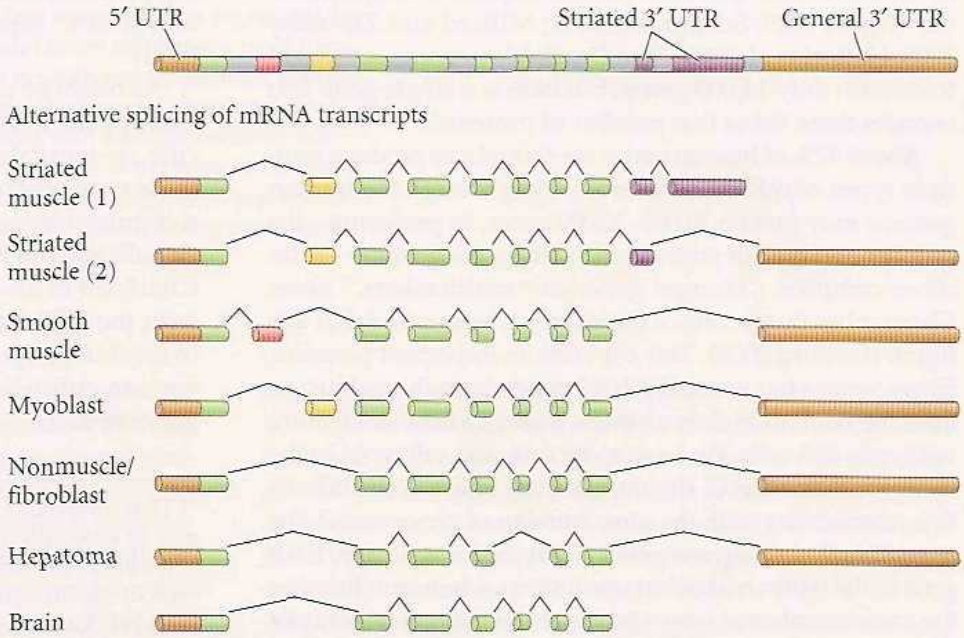


FIGURE 2.27 Alternative RNA splicing to form a family of rat α-tropomyosin proteins. The α-tropomyosin gene is represented on top. The thin lines represent the sequences that become introns and are spliced out to form the mature mRNAs. Constitutive exons (found in all tropomyosins) are shown in green. Those expressed only in smooth muscle are red; those expressed only in striated muscle are purple. Those that are variously expressed are yellow. Note that in addition to the many possible combinations of exons, two different 3′ ends ("striated" and "general") are possible. (After Breitbart et al. 1987.)

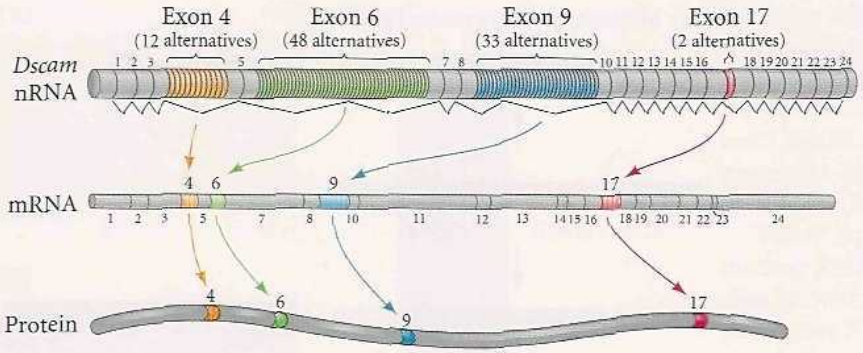


FIGURE 2.28 The *Dscam* gene of *Drosophila* can produce 38,016 different types of proteins by alternative nRNA splicing. The gene contains 24 exons. Exons 4, 6, 9, and 17 are encoded by sets of mutually exclusive possible sequences. Each messenger RNA will contain one of the 12 possible exon 4 sequences, one of the 48 possible exon 6 alternatives, one of the 33 possible exon 9 alternatives, and one of the 2 possible exon 17 sequences. The *Drosophila Dscam* gene is homologous to a DNA sequence on human chromosome 21 that is expressed in the nervous system. Disturbances of this gene in humans may lead to the neurological defects of Down syndrome (After Yamakawa et al. 1998; Saito 2000.)

from binding to one another. *Dscam1* contains 115 exons. However, a dozen different adjacent DNA sequences can be selected to be exon 4. Similarly, more than 30 mutually exclusive adjacent DNA sequences can become exons 6 and 9, respectively (Figure 2.28; Schmucker et al. 2000). If all possible combinations of exons are used, this one gene can produce 38,016 different proteins, and random searches for these combinations indicate that a large fraction of them are in fact made. The nRNA of *Dscam1* has been found to be alternatively spliced in different axons, and when two dendrites from the same axon touch each other, they are repelled. This causes the extensive branching of the dendrites. It appears that the thousands of splicing isoforms are needed to ensure that each neuron acquires a unique identity (Figure 2.29; Schmucker 2007; Millard and Zipursky 2008; Hattori et al. 2009). The *Drosophila* genome is thought to contain only 14,000 genes, but here is a single gene that encodes three times that number of proteins!

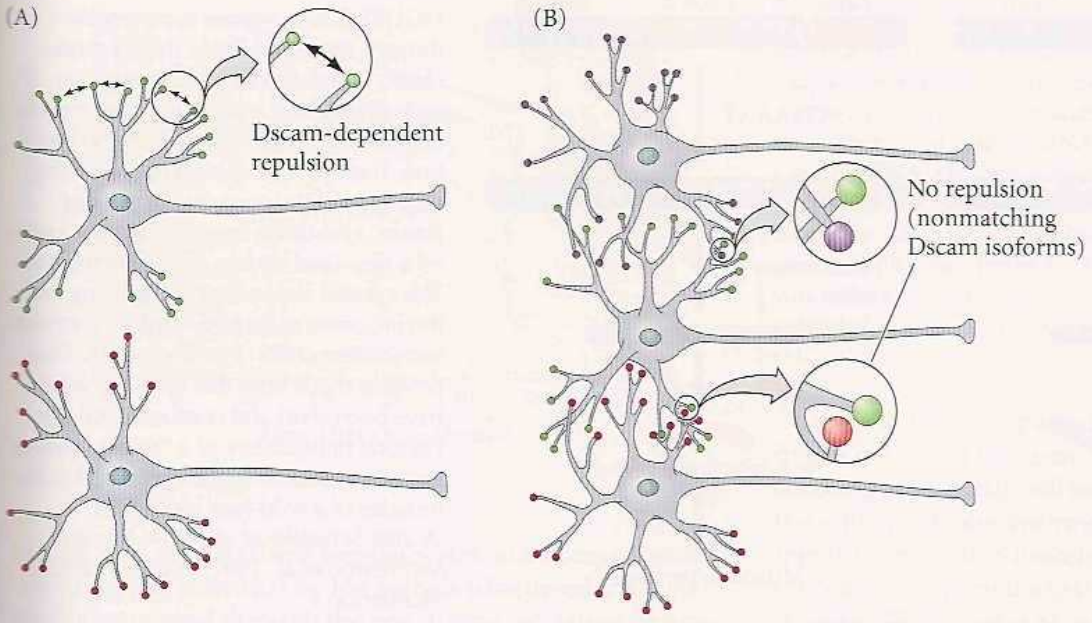
About 92% of human genes are thought to produce multiple types of mRNA. Therefore, even though the human genome may contain 20,000–30,000 genes, its **proteome**—the number and type of proteins encoded by the genome—is far more complex. "Human genes are multitaskers," notes Christopher Burge, one of the scientists who calculated this figure (Ledford 2008). This explains an important paradox. *Homo sapiens* has around 20,000 genes in each nucleus; so does the nematode *Caenorhabditis elegans*, a tubular creature with only 969 cells. We have more cells and cell types in the shaft of a hair than *C. elegans* has in its entire body. What's this worm doing with the same number of genes as us? The answer is that *C. elegans* genes rarely make isoforms. Each gene in the worm makes but one protein, whereas in humans the same number of genes produces an enormous array of different proteins.

Splicing enhancers and recognition factors

The mechanisms of differential RNA processing involve both *cis*-acting sequences on the nRNA and *trans*-acting protein factors that bind to these regions (Black 2003). The *cis*-acting sequences on nRNA are usually close to their potential 5′ or 3′ splice sites. These sequences are called "splicing enhancers," since they promote the assembly of spliceosomes at RNA cleavage sites. Conversely, these same sequences can be "splicing silencers" if they act to exclude exons from an mRNA sequence. These sequences are recognized by *trans*-acting proteins, most of which can recruit spliceosomes to that area. However, some *trans*-acting proteins, such as the polyprimidine tract-binding protein (PTP),* repress spliceosome formation where they bind.

As might be expected, there are some splicing enhancers that appear to be specific for certain tissues. Muscle-specific *cis*-regulatory sequences have been found around those exons characterizing muscle cell messages. These are recognized by certain proteins that are found in the muscle cells early in their development (Ryan and Cooper 1996; Charlet-B et al. 2002). Their presence is able to compete with the PTP that would otherwise prevent the inclusion of the muscle-specific exon into the mature message. In this way, an entire battery of muscle-specific isoforms can be generated.

*PTP is involved in making the correct isoform of tropomyosin and may be especially important in determining the mRNA populations of the brain. PTP is also involved in the mutually exclusive use of exon IIIb or IIIc in the mRNA for fibroblast growth factor 2 (see Figure 2.26B; Carstens et al. 2000; Lilleväli et al. 2001; Robinson and Smith 2006).



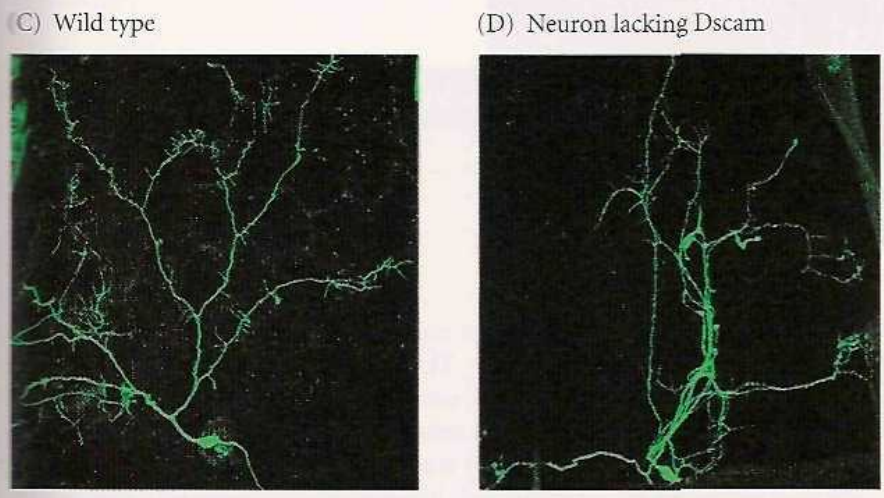


FIGURE 2.29 Dscam protein is specifically required to keep dendrites from the same neuron from adhering to each other. (A) When sister dendrites expressing the same splicing isoform of Dscam touch, the Dscam-Dscam interactions repel them and cause the dendrites to separate. (B) Different neurons express different splicing isoforms that do not interact with one another (and therefore do not trigger repulsion), which allows neurons to interact. (C) Neurons with multiple dendrites normally develop highly branched formations in which none of the branches crosses another. (D) Loss of *Dscam1* in such a neuron abolishes self-repulsion and results in excessive self-crossing and adhesion. (After Schmucker 2007; photographs courtesy of Dietmar Schmucker.)

One might also suspect that mutations of the splicing sites would lead to alternative phenotypes. Most of these splice site mutations lead to nonfunctional proteins and serious diseases. For instance, a single base change at the 5′ end of intron 2 in the human β-globin gene prevents splicing from occurring and generates a nonfunctional mRNA (Baird et al. 1981). This causes the absence of any β-globin from this gene, and thus a severe (and often life-threatening) type of anemia. Similarly, a mutation in the *dystrophin* gene at a particular splice site causes the skipping of that exon and a severe form of muscular dystrophy (Sironi et al. 2001). In at least one such case, the splice site mutation was not dangerous and actually gave the patient greater strength. In a different case, Schuelke and colleagues (2004) described a family in which individuals in four generations had a splice site mutation in the *myostatin* gene (Figure 2.30A). Among the family members were professional athletes and a 4-year-old toddler who was able to hold two 3-kg dumb-

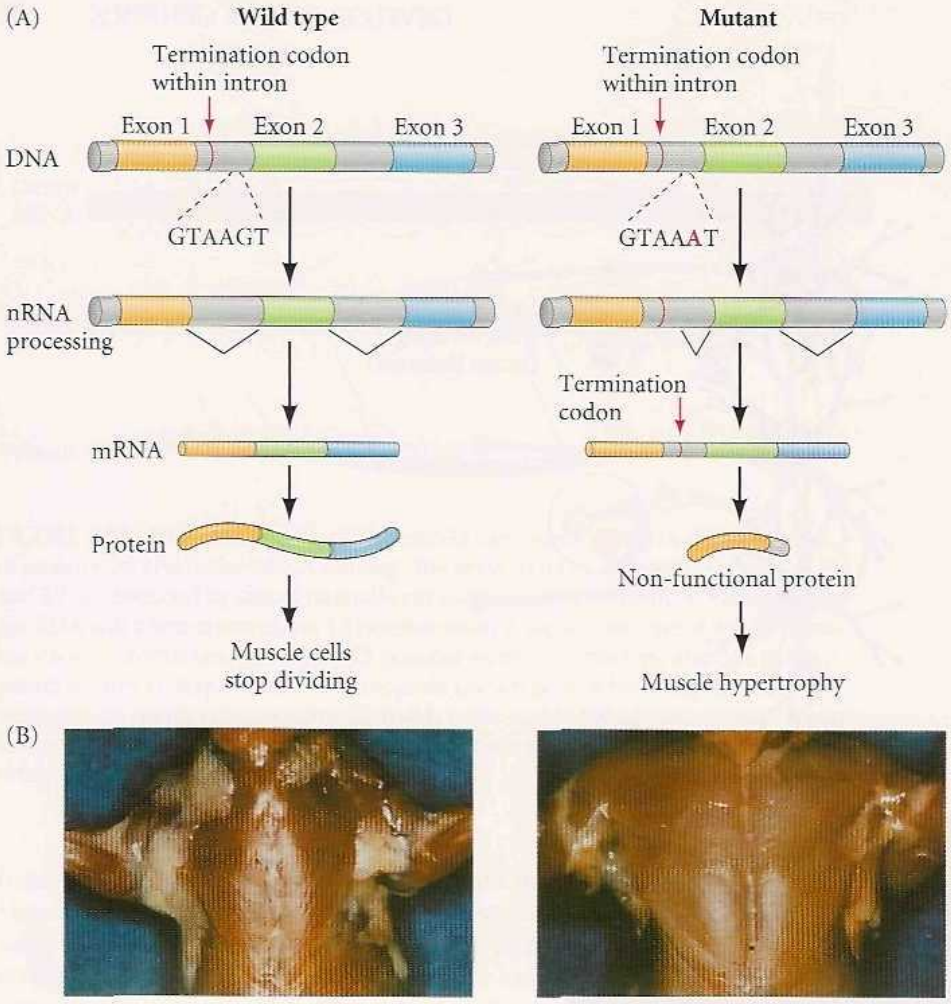


FIGURE 2.30 Muscle hypertrophy through mispliced RNA. This mutation results in a deficiency of the negative growth regulator myostatin in the muscle cells. (A) Molecular analysis of the mutation. There is no mutation in the coding sequence of the gene, but in the first intron, a mutation from a G to an A created a new (and widely used) splicing site. This caused aberrant nRNA splicing and the inclusion of an early protein synthesis termination codon into the mRNA. Thus, proteins made from that message would have been short and nonfunctional. (B) Pectoral musculature of a "mighty mouse" with the mutation (right) compared to the muscles of a wild-type mouse (left). (A after Schuelke et al. 2004; B from McPherron et al. 1997; courtesy of A. C. McPherron.)

bells with his arms fully extended. The *myostatin* gene product is a negative regulator—a factor that tells muscle precursor cells to stop dividing. In mammals with the mutation, the muscles are not told to differentiate until they have undergone many more rounds of cell division, and the result is larger muscles (Figure 2.30B).

Control of Gene Expression at the Level of Translation

The splicing of nuclear RNA is intimately connected with its export through the nuclear pores and into the cytoplasm. As the introns are removed, specific proteins bind to the spliceosome and attach the spliceosome-RNA complex to nuclear pores (Luo et al. 2001; Strässer and Hurt 2001). But once the RNA has reached the cytoplasm, there is still no guarantee that it will be translated. The control of gene expression at the level of translation can occur by many means; some of the most important of these are described below.

Differential mRNA longevity

The longer an mRNA persists, the more protein can be translated from it. If a message with a relatively short half-life were selectively stabilized in certain cells at certain times, it would make large amounts of its particular protein only at those times and places.

The stability of a message often depends on the length of its polyA tail. This, in turn, depends largely on sequences in the 3′ untranslated region, certain of which allow longer polyA tails than others. If these 3′ UTR regions are experimentally traded, the half-lives of the resulting mRNAs are altered: long-lived messages will decay rapidly, while normally short-lived mRNAs will remain around longer (Shaw and Kamen 1986; Wilson and Treisman 1988; Decker and Parker 1995).

In some instances, messenger RNAs are selectively stabilized at specific times in specific cells. The mRNA for casein, the major protein of milk, has a half-life of 1.1 hours in rat mammary gland tissue. However, during periods of lactation, the presence of the hormone prolactin increases this half-life to 28.5 hours (Figure 2.31; Guyette et al. 1979). In the development of the nervous system, a group of proteins called HuD proteins stabilizes a group of mRNAs that stop the neuronal precursor cells from dividing and also stabilizes a second group of mRNAs that are critical for these cells to start neuron differentiation (Okano and Darnell 1997; Deschênes-Furry et al. 2006, 2007).

Selective inhibition of mRNA translation: Stored oocyte mRNAs

Some of the most remarkable cases of translational regulation of gene expression occur in the oocyte. The oocyte often makes and stores mRNAs that will be used only after

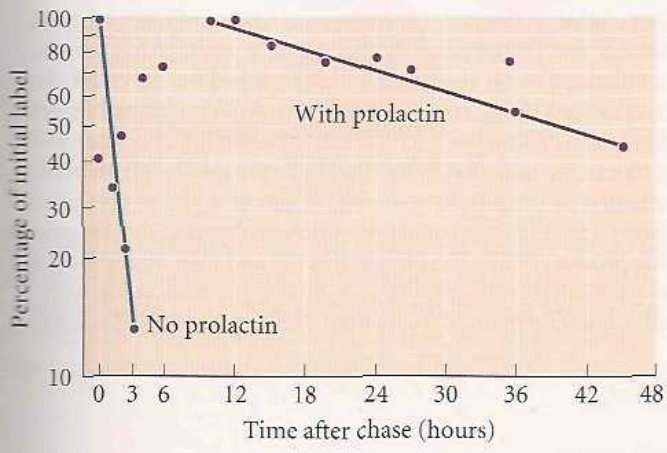


FIGURE 2.31 Degradation of casein mRNA in the presence and absence of prolactin. Cultured rat mammary cells were given radioactive RNA precursors (pulse) and, after a given time, were washed and given nonradioactive precursors (chase). This procedure labeled the casein mRNA synthesized during the pulse time. Casein mRNA was then isolated at different times following the chase and its radioactive label measured. In the absence of prolactin, the labeled (i.e., newly synthesized) casein mRNA decayed rapidly, with a half-life of 1.1 hours. When the same experiment was done in a medium containing prolactin, the half-life was extended to 28.5 hours. (After Guyette et al. 1979.)

fertilization occurs. These messages stay in a dormant state until they are activated by ion signals (discussed in Chapter 4) that spread through the egg during ovulation or fertilization.

Table 2.2 gives a partial list of mRNAs that are stored in the oocyte cytoplasm. Some of these stored mRNAs encode proteins that will be needed during cleavage, when the embryo makes enormous amounts of chromatin, cell membranes, and cytoskeletal components. Some of them encode cyclin proteins that regulate the timing of early cell division (Rosenthal et al. 1980; Standart et al. 1986). Indeed, in many species (including sea urchins and *Drosophila*),

TABLE 2.2 Some mRNAs stored in oocyte cytoplasm and translated at or near fertilization

mRNAs encoding	Function(s)	Organism(s)
Cyclins	Cell division regulation	Sea urchin, clam, starfish, frog
Actin	Cell movement and contraction	Mouse, starfish
Tubulin	Formation of mitotic spindles, cilia, flagella	Clam, mouse
Small subunit of ribonucleotide reductase	DNA synthesis	Sea urchin, clam, starfish
Hypoxanthine phosphoribosyl-transferase	Purine synthesis	Mouse
Vg1	Mesodermal determination(?)	Frog
Histones	Chromatin formation	Sea urchin, frog, clam
Cadherins	Blastomere adhesion	Frog
Metalloproteinases	Implantation in uterus	Mouse
Growth factors	Cell growth; uterine cell growth(?)	Mouse
Sex determination factor FEM-3	Sperm formation	*C. elegans*
PAR gene products	Segregate morphogenetic determinants	*C. elegans*
SKN-1 morphogen	Blastomere fate determination	*C. elegans*
Hunchback morphogen	Anterior fate determination	*Drosophila*
Caudal morphogen	Posterior fate determination	*Drosophila*
Bicoid morphogen	Anterior fate determination	*Drosophila*
Nanos morphogen	Posterior fate determination	*Drosophila*
GLP-1 morphogen	Anterior fate determination	*C. elegans*
Germ cell-less protein	Germ cell determination	*Drosophila*
Oskar protein	Germ cell localization	*Drosophila*
Ornithine transcarbamylase	Urea cycle	Frog
Elongation factor 1a	Protein synthesis	Frog
Ribosomal proteins	Protein synthesis	Frog, *Drosophila*

Compiled from numerous sources.

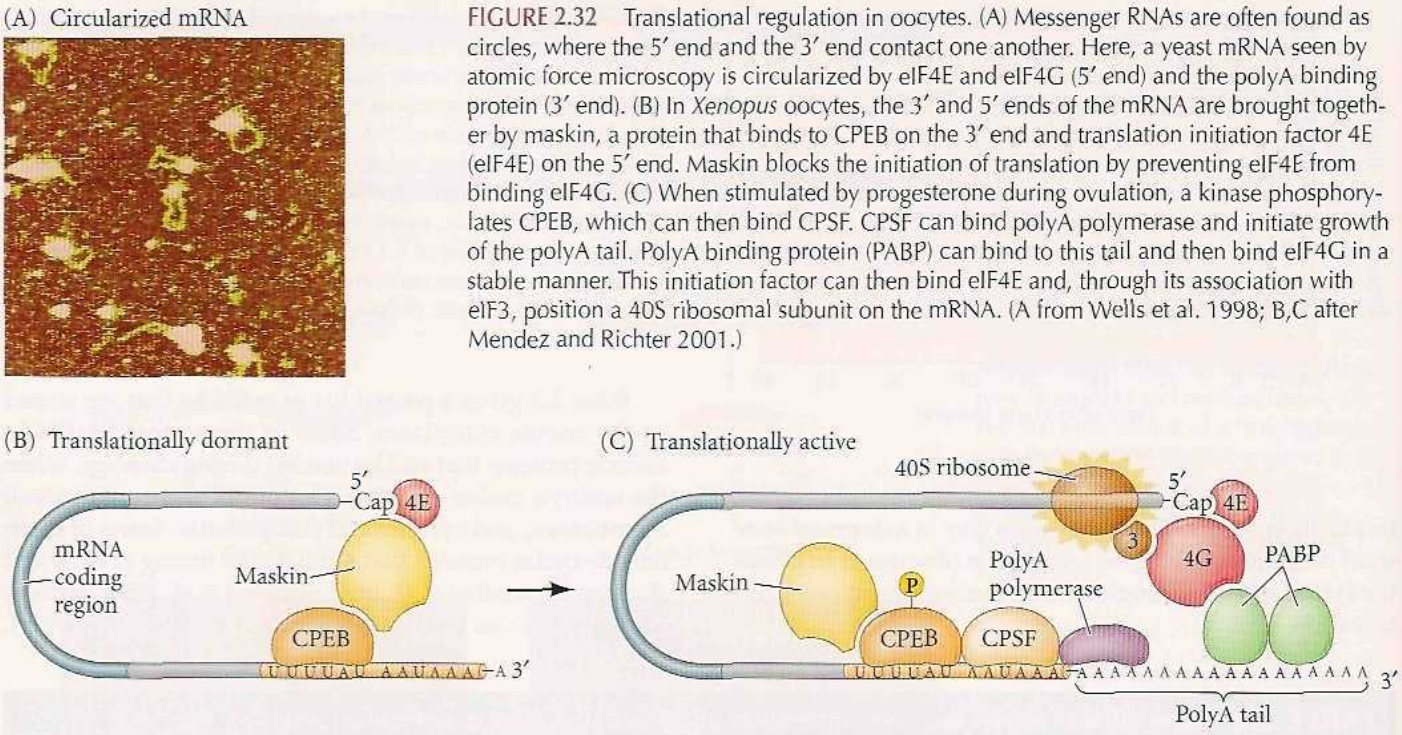


FIGURE 2.32 Translational regulation in oocytes. (A) Messenger RNAs are often found as circles, where the 5′ end and the 3′ end contact one another. Here, a yeast mRNA seen by atomic force microscopy is circularized by eIF4E and eIF4G (5′ end) and the polyA binding protein (3′ end). (B) In *Xenopus* oocytes, the 3′ and 5′ ends of the mRNA are brought together by maskin, a protein that binds to CPEB on the 3′ end and translation initiation factor 4E (eIF4E) on the 5′ end. Maskin blocks the initiation of translation by preventing eIF4E from binding eIF4G. (C) When stimulated by progesterone during ovulation, a kinase phosphorylates CPEB, which can then bind CPSF. CPSF can bind polyA polymerase and initiate growth of the polyA tail. PolyA binding protein (PABP) can bind to this tail and then bind eIF4G in a stable manner. This initiation factor can then bind eIF4E and, through its association with eIF3, position a 40S ribosomal subunit on the mRNA. (A from Wells et al. 1998; B,C after Mendez and Richter 2001.)

maintenance of the normal rate and pattern of early cell divisions does not require a nucleus; rather, it requires continued protein synthesis from stored maternal mRNAs (Wagenaar and Mazia 1978; Edgar et al. 1994). Other stored messages encode proteins that determine the fates of cells. These include the *bicoid*, *caudal*, and *nanos* messages that provide information in the *Drosophila* embryo for the production of its head, thorax, and abdomen.

Most translational regulation in oocytes is negative, as the "default state" of the mRNA is to be available for translation. Therefore, there must be inhibitors preventing the translation of these mRNAs in the oocyte, and these inhibitors must somehow be removed at the appropriate times around fertilization. The 5′ cap and the 3′ untranslated region seem especially important in regulating the accessibility of mRNA to ribosomes. If the 5′ cap is not made or if the 3′ UTR lacks a polyA tail, the message probably will not be translated. The oocytes of many species have "used these ends as means" to regulate the translation of their mRNAs.

It is important to realize that, unlike the usual representations of mRNA, most mRNAs probably form circles, with their 3′ end being brought to their 5′ end (Figure 2.32A). The 5′ cap is bound by **eukaryotic initiation factor-4E (eIF4E)**, a protein that is also bound to eIF4A (a helicase that unwinds double-stranded regions of RNA) and **eIF4G**, a scaffold protein that allows the mRNA to bind to the ribosome through its interaction with eIF4E (Wells et al. 1998; Gross et al. 2003). The polyA binding protein, which sits on the polyA tail of the mRNA, also binds to the eIF4G protein. This brings the 3′ end of the message next to the 5′ end and allows the messenger RNA to be recognized by the ribosome. Thus, the 5′ cap is critical for translation, and some animal's oocytes have used this as a direct means of translational control. For instance, the oocyte of the tobacco hornworm moth makes some of its mRNAs without their methylated 5′ caps. In this state, they cannot be efficiently translated. However, at fertilization, a methyltransferase completes the formation of the caps, and these mRNAs can be translated (Kastern et al. 1982).

In amphibian oocytes, the 5′ and 3′ ends of many mRNAs are brought together by a protein called **maskin** (Stebbins-Boaz et al. 1999; Mendez and Richter 2001). Maskin links the 5′ and 3′ ends into a circle by binding to two other proteins, each at opposite ends of the message. First, it binds to the **cytoplasmic polyadenylation-element-binding protein (CPEB)** attached to the UUUUAU sequence in the 3′ UTR; second, maskin also binds to the eIF4E factor that is attached to the cap sequence. In this configuration, the mRNA cannot be translated (Figure 2.32B). The binding of eIF4E to maskin is thought to prevent the binding of eIF4E to eIF4G, a critically important translation initiation factor that brings the small ribosomal subunit to the mRNA.

Mendez and Richter (2001) have proposed an intricate scenario to explain how mRNAs bound together by maskin become translated at about the time of fertilization. At ovulation (when the hormone progesterone stimulates the last meiotic divisions of the oocyte and the oocyte is released for fertilization), a kinase activated by progesterone phosphorylates the CPEB protein. The phosphorylated CPEB can now bind to CPSF, the cleavage and polyadenylation specificity factor (Mendez et al. 2000; Hodgman et al. 2001). The bound CPSF protein sits on a

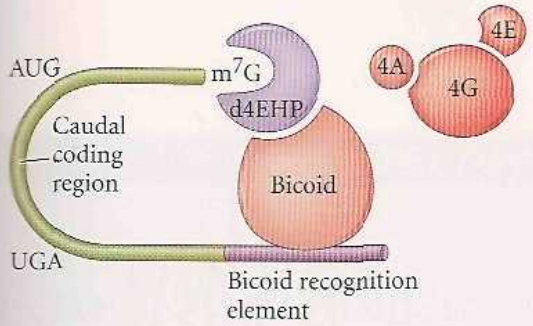


FIGURE 2.33 Protein binding in *Drosophila* oocytes. Bicoid protein binds to a recognition element in the 3′ UTR of the *caudal* message. Bicoid can bind to d4EHP, which prevents the binding of eIF4E to the cap structure. Without eIF4E, the eIF4G cannot bind and initiate translation. (After Cho et al. 2005.)

particular sequence of the 3′ UTR that has been shown to be critical for polyadenylation, and it complexes with a polymerase that elongates the polyA tail of the mRNA. In oocytes, a message having a short polyA tail is not degraded; however, such messages are not translated.

Once the tail is extended, molecules of the polyA binding protein (PABP) can attach to the growing tail. PABP proteins stabilize eIF4G, allowing it to outcompete maskin for the binding site on the eIF4E protein at the 5′ end of the mRNA. The eIF4G protein can then bind eIF3, which can position the small ribosomal subunit onto the mRNA. The small (40S) ribosomal subunit will then find the initiator tRNA, complex with the large ribosomal subunit, and initiate translation (Figure 2.32C).

In the *Drosophila* oocyte, Bicoid can act both as a transcription factor (activating genes such as *hunchback*) and also as a translational inhibitor (see Chapter 6). Bicoid represses the translation of *caudal* mRNA, preventing its transcription in the anterior half of the embryo. (The protein made from the *caudal* message is important in activating those genes that specify the cells to be abdomen precursors.) Bicoid inhibits *caudal* mRNA translation by binding to a "bicoid recognition element," a series of nucleotides in the 3′ UTR of the *caudal* message (Figure 2.33). Once there, Bicoid can bind with and recruit another protein, d4EHP. The d4EHP protein can compete with eIF4E for the cap. Without eIF4E, there is no association with eIF4G and the *caudal* mRNA becomes untranslatable. As a result, the *caudal* message is not translated in the anterior of the embryo (where Bicoid is abundant), but is active in the posterior portion of the embryo.

microRNAs: Specific regulators of mRNA translation and transcription

If proteins can bind to specific nucleic acid sequences to block transcription or translation, you would think that RNA would do the job even better. After all, RNA can be made specifically to bind a particular sequence. Indeed, one of the most efficient means of regulating the translation of a specific message is to make a small RNA complementary to a portion of a particular mRNA. Such a naturally occurring antisense RNA was first seen in *C. elegans* (Lee et al. 1993; Wightman et al. 1993). Here, the *lin-4* gene was found to encode a 21-nucleotide RNA that bound to multiple sites in the 3′ UTR of the *lin-14* mRNA (Figure 2.34). The *lin-14* gene encodes a transcription factor, LIN-14, that is important during the first larval phase of *C. elegans* development. It is not needed afterward, and *C. elegans* is able to inhibit synthesis of LIN-14 from these messages by activating the *lin-4* gene. The binding of *lin-*

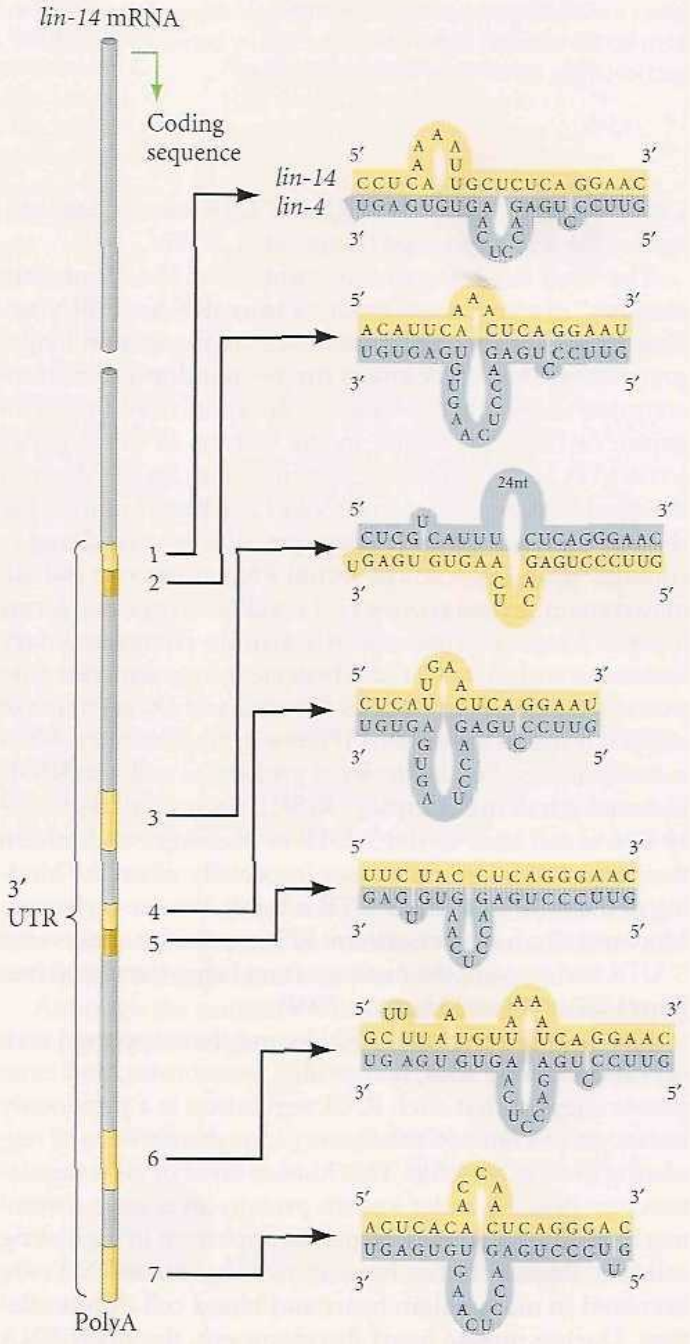


FIGURE 2.34 Hypothetical model of the regulation of *lin-14* mRNA translation by *lin-4* RNAs. The *lin-4* gene does not produce an mRNA. Rather, it produces small RNAs that are complementary to a repeated sequence in the 3′ UTR of the *lin-14* mRNA, which bind to it and prevent its translation. (After Wickens and Takayama 1995.)

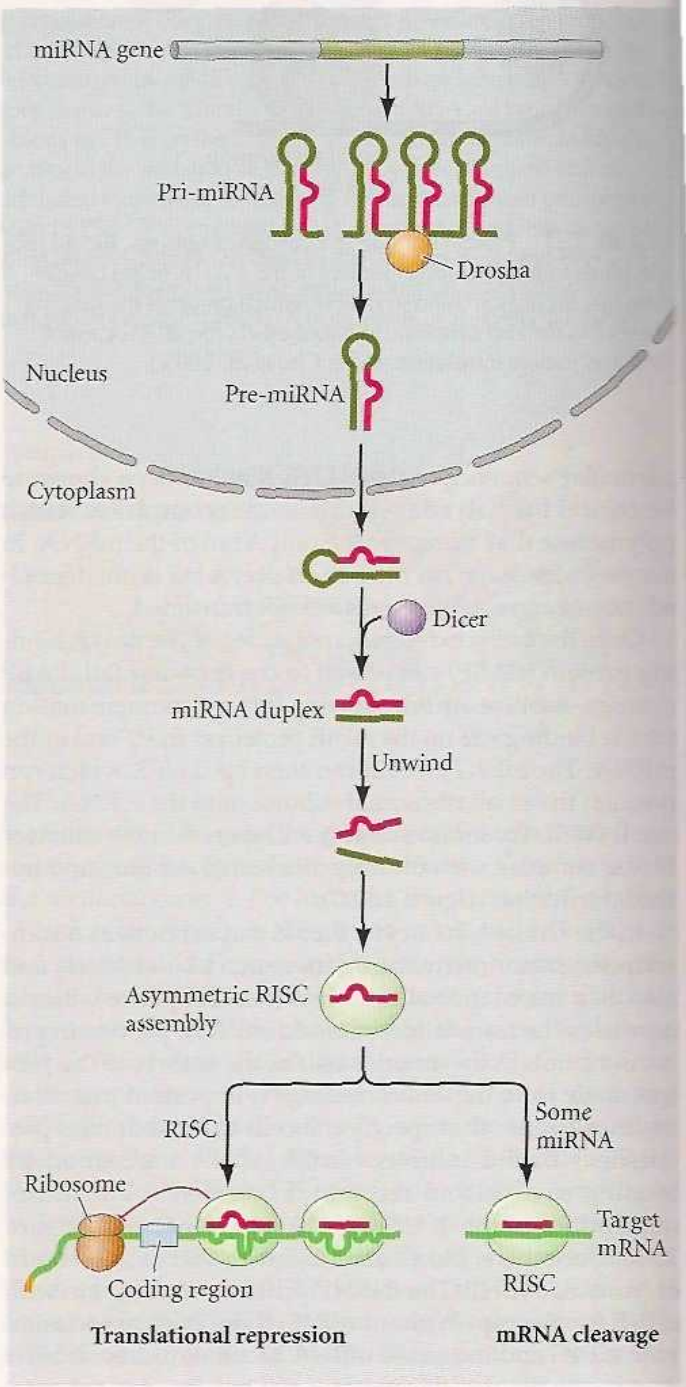


FIGURE 2.35 Current model for the formation and use of microRNAs. The miRNA gene encodes a pri-miRNA that often has several hairpin regions where the RNA finds nearby complementary bases with which to pair. The pri-miRNA is processed into individual pre-miRNA "hairpins" by the Drosha RNAase, and these are exported from the nucleus. Once in the cytoplasm, another RNAase, Dicer, eliminates the non-base-paired loop. Dicer also acts as a helicase to separate the strands of the double-stranded miRNA. One strand (probably recognized by placement of Dicer) is packaged with proteins into the RNA-induced silencing complex (RISC), which subsequently binds to the 3′ UTRs to effect translational suppression or cleavage, depending (at least in part) on the strength of the complementarity between the miRNA and its target. (After He and Hannon 2004.)

4 transcripts to the *lin-14* mRNA 3′ UTR causes degradation of the *lin-14* message (Bagga et al. 2005).

The *lin-4* RNA is now thought to be the "founding member" of a very large group of **microRNAs (miRNAs)**. These RNAs of about 22 nucleotides are made from longer precursors. These precursors can be in independent transcription units (the *lin-4* gene is far apart from the *lin-14* gene), or they can reside in the introns of other genes (Aravin et al. 2003; Lagos-Quintana et al. 2003). Many of the newly discovered microRNAs have been found in the regions between genes (regions previously considered to contain "junk DNA"). The initial RNA transcript (which may contain several repeats of the miRNA sequence) forms hairpin loops wherein the RNA finds complementary structures within its strand. These stem-loop structures are processed by a set of RNases (Drosha and Dicer) to make single-stranded microRNA (Figure 2.35). The microRNA is then packaged with a series of proteins to make an **RNA-induced silencing complex (RISC)**. Such small regulatory RNAs can bind to the 3′ UTR of messages and inhibit their translation. In some cases (especially when the binding of the miRNA to the 3′ UTR is tight), the site is cleaved. More usually, however, several RISCs attach to sites on the 3′ UTR and prevent the message from being translated (see Bartel 2004; He and Hannon 2004).

The abundance of microRNAs and their apparent conservation among flies, nematodes, vertebrates, and even plants suggest that such RNA regulation is a previously unrecognized but potentially very important means of regulating gene expression. This hidden layer of gene regulation parallels the better known protein-level gene control mechanisms, and it may be just as important in regulating cell fate. Recent studies have shown that microRNAs are involved in mammalian heart and blood cell differentiation. During mouse heart development, the microRNA *miR1* can repress the messages encoding the Hand2 transcription factor (Zhao et al. 2005). This transcription factor is critical in the proliferation of ventricle heart muscle cells, and *miR1* may control the balance between ventricle growth and differentiation. The *miR181* miRNA is essential for committing progenitor cells to differentiate into B lymphocytes, and ectopic expression of *miR181* in mice causes a preponderance of B lymphocytes (Figure 2.36; Chen et al. 2004).

MicroRNAs are also used to "clean up" and fine-tune the level of gene products. We mentioned those maternal RNAs that allow early development to occur. How does the embryo get rid of maternal RNAs once they have been used and the embryonic cells are making their own mRNAs? In zebrafish, this cleanup operation is assigned to microRNAs such as *miR430*. This is one of the first genes transcribed by the fish embryonic cells, and there are about 90 copies of this gene in the zebrafish genome. So the level

SIDELIGHTS & SPECULATIONS

microRNAs in Transcriptional Gene Regulation

In addition to its role in the translational regulation of gene expression, microRNAs also appear to be able to silence the transcription of certain genes. Such genes are often located in the heterochromatin—that region of the genome where the DNA is tightly coiled and transcription is inhibited by the packed nucleosomes. Volpe and colleagues (2003) discovered that if they deleted the genes in yeast encoding the appropriate RNases or RISC proteins, the heterochromatin around the centromeres became unpacked, the histones in this region lost their inhibitory methylation, and the centromeric heterochromatin started making RNA. Similar phenomena were seen when these proteins were mutated in *Drosophila* (Pal-Bhadra et al. 2004). Indeed, in *Drosophila*, the *Suppressor-of-stellate* gene on the Y chromosome makes a microRNA that represses the transcription of the *stellate* gene on the X chromosome (Gvozdev et al. 2003). This is important for dosage regulation of the X chromosomes in *Drosophila*.

It appears that microRNAs are able to bind to the nuclear RNA as it is being transcribed, and form a complex with the methylating and deacetylating enzymes, thus repressing the gene (Kato et al. 2005; Schramke et al. 2005). If synthetic microRNA made complementary to specific promoters is added to cultured human cells, that microRNA is able to induce that promoter's DNA to become methylated. Lysine 9 on histone H3 also becomes methylated around the promoter, and transcription from that gene stops (Kawasaki and Taira 2004; Morris et al. 2004).

This appears to be the mechanism by which NRSF (see page 47) functions. NRSF prevents gene expression in nonneural cells by repressing microRNAs that would otherwise recruit histone acetyltransferases to activate genes that promote neuron production. In the presence of NRSF, these miRNAs are not present, and so histone deacetylases and methyltransferases are recruited to the chromatin instead. The resulting methylation produces conglomerations of nucleosomes linked together by heterochromatin protein-1 (HP1), thereby stabilizing the conglomerate and preventing transcription of the neuron-promoting genes "hidden" within it (Ooi and Wood 2007; Yoo et al. 2009.) A single silencer protein bound to the DNA can prevent the gene's expression.

Thus, microRNA directed against the 3′ end of mRNA may be able to shut down gene expression on the translational level, while microRNA directed at the promoters of genes may be able to block gene expression at the transcriptional level. The therapeutic value of these RNAs in cancer therapy is just beginning to be explored (see Gaur and Rossi 2006).

X chromosome inactivation in mammals is also directed by small noncoding RNAs, albeit not the canonical microRNAs. Although DNA methylation is responsible for keeping one of the two X chromosomes inactive, the choice of which X chromosome to activate arises from the physical interactions between the two X chromosomes and their expressing several small, noncoding RNAs. The mechanisms of differential expression of small RNAs between the two X chromosomes is under intensive investigation (Augui et al. 2007; Migeon 2007; Ogawa et al. 2008).

of *miR430* goes up very rapidly. This microRNA has hundreds of targets (about 40% of the maternal RNA types), and when it binds to the 3′ UTR of these target mRNAs, these mRNAs lose their polyA tails and are degraded (Giraldez et al. 2006). Slightly later in development, this same microRNA is used in the fish embryo to fine-tune the expression of *Nodal* mRNA (Choi et al. 2007). The consequence of this latter use of *miR430* is the determination of how many cells become committed to the endoderm and how many become committed to be mesoderm.

Although the microRNA is usually about 22 bases long, it recognizes its target primarily through a "seed" region

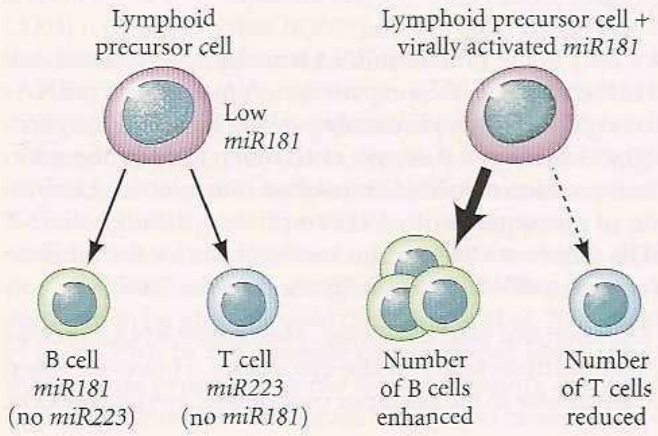


FIGURE 2.36 The lymphoid precursor cell can generate B cells (lymphocytes that make antibodies) or T cells (lymphocytes that kill virally infected cells). This differentiation depends on the organ in which they reside. The regulation of the lineage pathway is controlled in part by levels of the microRNA *miR181*. The lymphocyte precursor cell has little *miR181*. A B cell has high levels of *miR181*, whereas T cells do not appear to have any. If lymphocyte precursor cells are virally transfected with *miR181*, they preferentially generate B cells at the expense of T cells.

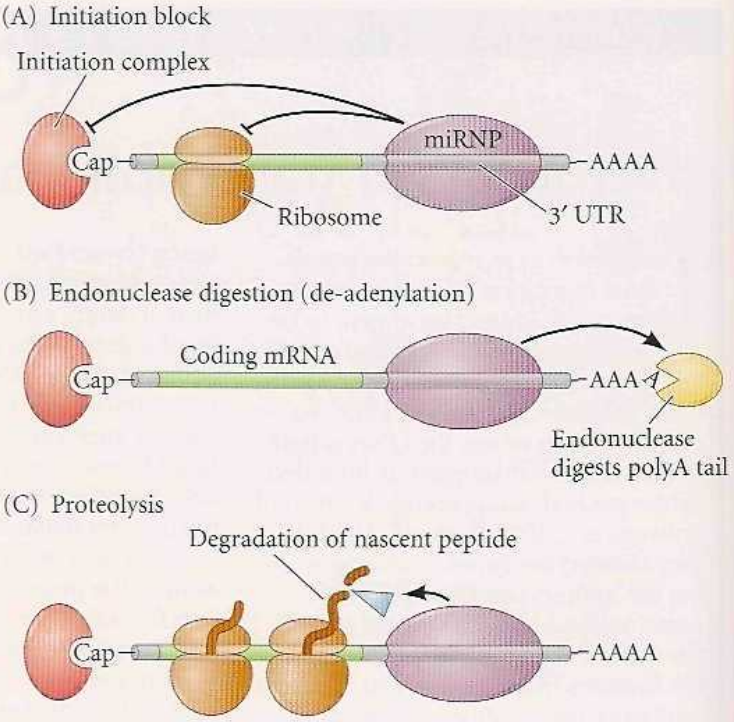


FIGURE 2.37 The miRNA complex, including numerous proteins that bind to the miRNA (miRNP), can block translation in several ways. These include (A) blocking the binding of the mRNA to initiation factors or ribosomes; (B) recruiting endonucleases to chew away the polyA tail of the mRNA, thereby causing its destruction; and (C) recruiting protein-digesting enzymes that destroy the nascent protein. (After Filipowicz et al. 2008.)

of about 5 bases in the 5′ end of the microRNA (usually at positions 2–7). This seed region recognizes targets in the 3′ UTR of the message. What happens, then, if an mRNA has a mutated 3′ UTR? Such a mutation appears to have given rise to the Texel sheep, a breed with a large and well-defined musculature that is the dominant meat-producing sheep in Europe. We have already seen that a mutation in the *myostatin* gene that prevents the proper splicing of the nRNA can produce a large-muscled phenotype. Another way of reducing the levels of myostatin involves a mutation in its 3′UTR sequence (see Figure 2.30). Genetic techniques mapped the basis of the sheep's meaty phenotype to the *myostatin* gene. In the Texel breed, there has been a G-to-A transition in the 3′ UTR of the gene for myostatin, creating a target for the *mir1* and *mir206* microRNAs that are abundant in skeletal muscle (Clop et al. 2006). This mutation causes the depletion of myostatin messages and the increase in muscle mass characteristic of these sheep.

The binding of microRNAs to the 3′ UTR can regulate translation in several ways (Figure 2.37; Filipowicz et al. 2008). First, they can block initiation of translation, preventing the binding of initiation factors or ribosomes. Second, they can recruit endonucleases that digest the mRNA, usually starting with the polyA tail. In a third mechanism, they allow translation to be initiated, but recruit proteolytic enzymes that digest the protein as it is being made. It is also possible that some microRNAs use more than one method, and it has been proposed (Mathonnet et al. 2007) that the microRNAs may first inhibit translation initiation and then consolidate mRNA silencing by causing the digestion of the message.

Control of RNA expression by cytoplasmic localization

Not only is the time of mRNA translation regulated, but so is the place of RNA expression. A majority of mRNAs (about 70% in *Drosophila* embryos) are localized to specific places in the cell (Lécuyer et al. 2007). Just like the selective repression of mRNA translation, the selective localization of messages is often accomplished through their 3′ UTRs. There are three major mechanisms for the localization of an mRNA (Figure 2.38; see Palacios 2007):

- *Diffusion and local anchoring*. Messenger RNAs such as *nanos* diffuse freely in the cytoplasm. However, when they diffuse to the posterior pole of the *Drosophila* oocyte, they are trapped there by proteins that reside particularly in these regions. These proteins also activate the mRNA, allowing it to be translated.
- *Localized protection*. Messenger RNAs such as those encoding the *Drosophila* heat shock protein hsp83 (which helps protect the embryos from thermal extremes) also float freely in the cytoplasm. Like *nanos* mRNA, *hsp83* accumulates at the posterior pole, but its mechanism for getting there is different. Throughout the embryo, the protein is degraded. However, proteins at the posterior pole protect the *hsp83* mRNA from being destroyed.
- *Active transport along the cytoskeleton*. This is probably the most widely used mechanism for mRNA localization. Here, the 3′ UTR of the mRNA is recognized by proteins that can bind these messages to "motor proteins" that travel along the cytoskeleton to their final destination. These motor proteins are usually ATPases such as dynein or kinesin that split ATP for their motive force. For instance, in *Drosophila* oocytes, the *bicoid* messages (which instruct the formation of the head) are localized to one end of the oocyte. The 3′ UTR of *bicoid* mRNA allows its message to bind to the microtubules through its association with two other proteins (Swallow and Staufen). If the *bicoid* 3′ UTR is attached to some other message, that mRNA will also be bound to the anterior pole of the oocyte (Driever and Nüsslein-Volhard 1988a,b; Ferrandon et al. 1994).

The 3′ UTR of the *bicoid* message binds the Staufen protein that connects it to dynein. Dynein travels along the microtubules in the "minus" direction, that is, toward the site where microtubules begin. In this way, the *bicoid* mRNA is localized to the future anterior part of the oocyte. Other mRNAs, such as the *Oskar* message, in contrast, appear to

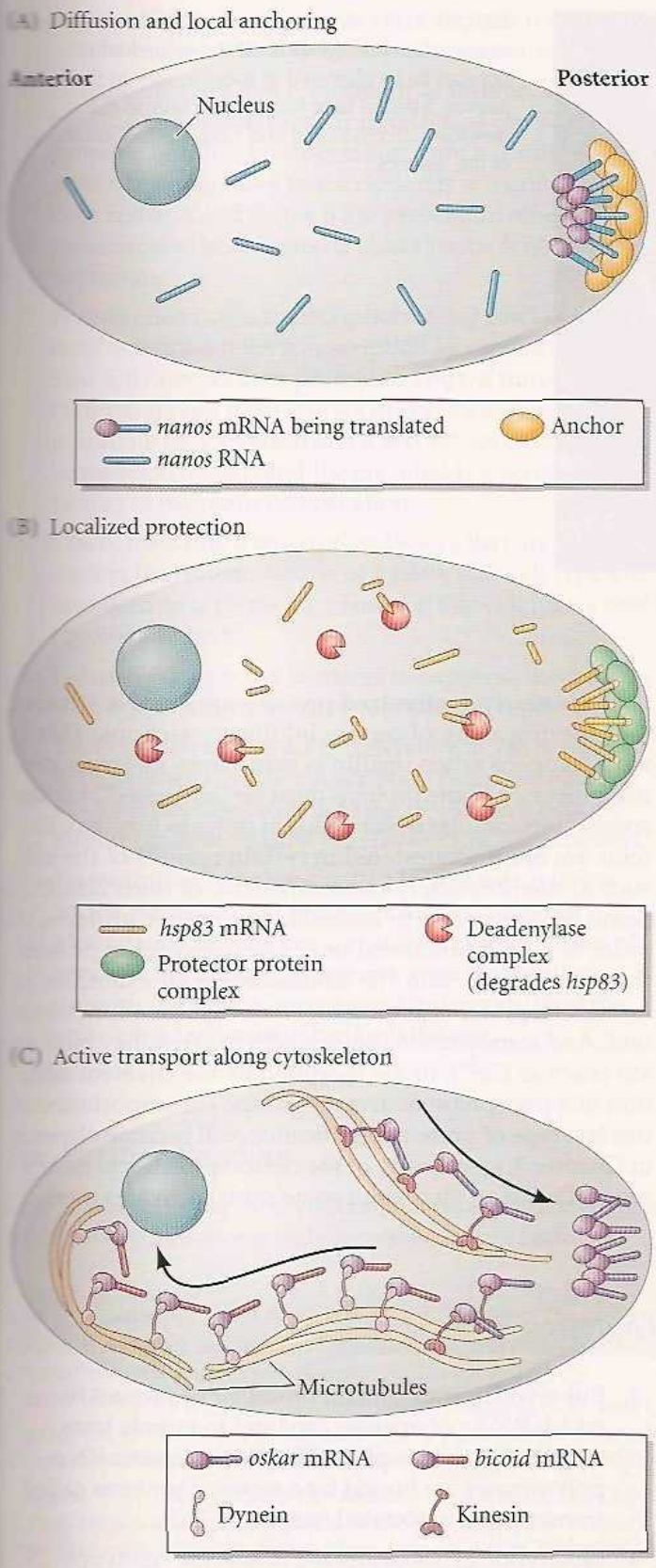


FIGURE 2.38 Localization of mRNAs. (A) Diffusion and local anchoring. *Nanos* mRNA diffuses through the *Drosophila* egg and is bound (in part by the Oskar protein, whose message is described in the text) at the posterior end of the oocyte. This anchoring allows the *nanos* mRNA to be translated. (B) Localized protection. The mRNA for *Drosophila* heat shock protein (hsp83) will be degraded unless it binds to a protector protein (in this case, also at the posterior terminal of the oocyte). (C) Active transport on the cytoskeleton, causing the accumulation of mRNA at a particular site. Here, *bicoid* mRNA is transported to the anterior of the oocyte by dynein and kinesin motor proteins. Meanwhile, *Oskar* mRNA is brought to the posterior pole by transport along microtubules by kinesin ATPases. (After Palacios 2007.)

bind to the kinesin motor protein, and it is taken toward the "plus" end of the microtubules, at the tip of their assembly. It is thereby taken to the posterior end of the *Drosophila* oocyte. Once transported to their destinations, mRNAs often bind to other cytoskeletal proteins (such as actin microfilaments).

Stored mRNAs in brain cells

One of the most important areas of local translational regulation may be in the brain. The storage of long-term memory requires new protein synthesis, and the local translation of mRNAs in the dendrites of brain neurons has been proposed as a control point for increasing the strength of synaptic connections (Martin 2000; Klann et al. 2004; Wang and Tiedge 2004). The ability to increase the strength of the connections between neurons is critical in forming the original architecture of the brain and also in the ability to learn. Indeed, in recent studies of mice, Kelleher and colleagues (2004) have shown that neuronal activity-dependent memory storage depends on the activation of eIF4E and other components of protein synthesis.

Several mRNAs appear to be transported along the cytoskeleton to the dendrites of neurons (the "receiving portion" of the neuron, where synapse connections are formed with the other neurons). These messages include those mRNAs encoding receptors for neurotransmitters (needed to transmit the signals from one neuron to another); activity-regulated enzymes; and the cytoskeletal components needed to build a synapse (Figure 2.39). As we will see in later chapters, one of the proteins responsible for constructing specific synapses is **brain-derived neurotrophic factor**, or **BDNF**. BDNF regulates neural activity and appears to be critical for new synapse formation. Takei and colleagues (2004) have shown that BDNF induces local translation of these neural messages in the dendrites.

Another indication of the importance of dendritic mRNA translation comes from studies of a leading cause of human mental retardation, fragile X syndrome. Fragile X syndrome is caused by loss-of-function mutations in the X-linked *FMR1* gene. The FMR1 protein appears to prevent the translation of several mRNAs that are being transported to the dendrites along microtubules in response to stimulation by glutamic acid (Dictenberg et al. 2008; Wang et al. 2008b). In the absence of functional FMR1, these mRNAs are expressed in the wrong amounts, leading to signaling abnormalities that are believed to cause the prob-

FIGURE 2.39 A brain-specific RNA in a cultured mammalian neuron. BC1 RNA (stained white) appears to be clustered at specific sites in the neuron (stained light blue), especially in the dendrites. (From Wang and Tiedge 2004, courtesy of the authors.)

lems in cognition and learning. Thus, translational regulation in neurons might be important not only for their initial development but also for their continued ability to learn and change.

Posttranslational regulation of gene expression

When a protein is synthesized, the story is still not over. Once a protein is made, it becomes part of a larger level of organization. For instance, it may become part of the structural framework of the cell, or it may become involved in one of the myriad enzymatic pathways for the synthesis or breakdown of cellular metabolites. In any case, the individual protein is now part of a complex "ecosystem" that integrates it into a relationship with numerous other proteins. Thus, several changes can still take place that determine whether or not the protein will be active.

Some newly synthesized proteins are inactive without the cleaving away of certain inhibitory sections. This is what happens when insulin is made from its larger protein precursor. Some proteins must be "addressed" to their specific intracellular destinations in order to function. Proteins are often sequestered in certain regions of the cell, such as membranes, lysosomes, nuclei, or mitochondria. Some proteins need to assemble with other proteins in order to form a functional unit. The hemoglobin protein, the microtubule, and the ribosome are all examples of numerous proteins joining together to form a functional unit. And some proteins are not active unless they bind an ion (such as Ca^{2+}), or are modified by the covalent addition of a phosphate or acetate group. The importance of this last type of protein modification will become obvious in Chapter 3, since many of the critical proteins in embryonic cells just sit there until some signal activates them.

Snapshot Summary: *Developmental Genetics*

1. Differential gene expression from genetically identical nuclei creates different cell types. Differential gene expression can occur at the levels of gene transcription, nuclear RNA processing, mRNA translation, and protein modification. Notice that RNA processing and export occur while the RNA is still being transcribed from the gene.
2. Genes are usually repressed, and activating a gene often means inhibiting its repressor. This fact leads to thinking in double and triple negatives: Activation is often the inhibition of the inhibitor; repression is the inhibition of the inhibitor of the inhibitor.
3. Eukaryotic genes contain promoter sequences to which RNA polymerase can bind to initiate transcription. To accomplish this, the eukaryotic RNA polymerases are bound by a series of proteins called transcription-associated factors, or TAFs.
4. Eukaryotic genes expressed in specific cell types contain enhancer sequences that regulate their transcription in time and space.
5. Specific transcription factors can recognize specific sequences of DNA in the promoter and enhancer

regions. These proteins activate or repress transcription from the genes to which they have bound.

6. Enhancers work in a combinatorial fashion. The binding of several transcription factors can act to promote or inhibit transcription from a certain promoter. In some cases transcription is activated only if both factor A and factor B are present; in other cases, transcription is activated if either factor A or factor B is present.
7. A gene encoding a transcription factor can keep itself activated if the transcription factor it encodes also activates its own promoter. Thus, a transcription factor gene can have one set of enhancer sequences to initiate its activation and a second set of enhancer sequences (which bind the encoded transcription factor) to maintain its activation.
8. Often, the same transcription factors that are used during the differentiation of a particular cell type are also used to activate the genes for that cell type's specific products.
9. Enhancers can act as silencers to suppress the transcription of a gene in inappropriate cell types.
10. Transcription factors act in different ways to regulate RNA synthesis. Some transcription factors stabilize RNA polymerase binding to the DNA; some disrupt nucleosomes, increasing the efficiency of transcription.
11. Transcription correlates with a lack of methylation on the promoter and enhancer regions of genes. Methylation differences can account for examples of genomic imprinting, wherein a gene transmitted through the sperm is expressed differently than the same gene transmitted through the egg.
12. Dosage compensation enables the X chromosome-derived products of males (which have one X chromosome per cell in fruit flies and mammals) to equal the X chromosome-derived products of females (which have two X chromosomes per cell). This compensation is accomplished at the level of transcription, either by accelerating transcription from the lone X chromosome in males (*Drosophila*), decreasing the level of transcription from each X chromosome by 50% (*C. elegans*), or by inactivating a large portion of one of the two X chromosomes in females (mammals).
13. Differential nuclear RNA selection can allow certain transcripts to enter the cytoplasm and be translated while preventing other transcripts from leaving the nucleus.
14. Differential RNA splicing can create a family of related proteins by causing different regions of the nRNA to be read as exons or introns. What is an exon in one set of circumstances may be an intron in another.
15. Some messages are translated only at certain times. The oocyte, in particular, uses translational regulation to set aside certain messages that are transcribed during egg development but used only after the egg is fertilized. This activation is often accomplished either by the removal of inhibitory proteins or by the polyadenylation of the message.
16. MicroRNAs can act as translational inhibitors, binding to the 3′ UTR of the RNA.
17. Many mRNAs are localized to particular regions of the oocyte or other cells. This localization appears to be regulated by the 3′ UTR of the mRNA.

For Further Reading

Complete bibliographical citations for all literature cited in this chapter can be found at the free-access website **www.devbio.com**

Clop, A. and 16 others. 2006. A mutation creating a potential illegitimate microRNA target site in the myostatin gene affects muscularity in sheep. *Nature Genet.* 38: 813–818.

Davidson, E. H. 2006. *The Regulatory Genome.* Academic Press, New York.

Migeon, B. R. 2007. *Females Are Mosaics: X Inactivation and Sex Differences in Disease.* Oxford University Press, New York.

Palacios, I. M. 2007. How does an mRNA find its way? Intracellular localization of transcripts. *Sem. Cell Dev. Biol.* 163–170.

Strahl, B. D. and C. D. Allis. 2000. The language of covalent histone modifications. *Nature* 403: 41–45.

Wilmut, I., K. Campbell and C. Tudge. 2001. *The Second Creation: Dolly and the Age of Biological Control.* Harvard University Press, Cambridge, MA.

Zhou, Q., J. Brown, A. Kanarek, J. Rajagopal and D. A. Melton. 2008. In vivo reprogramming of adult pancreatic exocrine cells to β cells. *Nature* 455: 627–632.

Zinzen, R. P., C. Girardot, J. Gagneur, M. Braun and E. E. Furlong. 2009. Combinatorial binding predicts spatio-temporal *cis*-regulatory activity. *Nature* 462: 65–70.

Go Online

WEBSITE 2.1 Does the genome or the cytoplasm direct development? The geneticists versus the embryologists. Geneticists were certain that genes controlled development, whereas embryologists generally favored the cytoplasm. Both sides had excellent evidence for their positions.

WEBSITE 2.2 The origins of developmental genetics. The first hypotheses for differential gene expression came from C. H. Waddington, Salome Gluecksohn-Waelsch, and other scientists who understood both embryology and genetics.

WEBSITE 2.3 Techniques of DNA analysis. The entries of this website describe crucial laboratory skill including gene cloning, DNA sequencing, Southern blotting, "knock-outs" of specific genes, enhancer traps, and identification of methylated sites.

WEBSITE 2.4 Techniques of RNA analysis. Techniques described here include northern blots, RT-PCR, in situ hybridization, microarray technology, antisense RNA, interference RNA, Cre-lox analysis, and ChIP-on-Chip.

WEBSITE 2.5 Bioinformatics. This entry provides links to various free websites with tools that enable researchers to compare DNA sequences and specific genomes with the aim of further illuminating the various mechanisms of gene regulation.

WEBSITE 2.6 Cloning and nuclear equivalence. The several entries here address the issues of cloning and whether or not the entire genome is the same in each cell of the body. As it turns out, lymphocytes make new genes during their development and their genomes are not identical.

WEBSITE 2.7 Silencing large blocks of chromatin. The inactivation or the elimination of entire chromosomes is not uncommon among invertebrates and is sometimes used as a mechanism of sex determination. Moreover, among mammals, random X chromosome inactivation may provide females with health benefits—as long as the process occurs flawlessly.

WEBSITE 2.8 So you think you know what a gene is? Different scientists have different definitions, and nature has given us some problematic examples of DNA sequences that may or may not be considered genes.

Vade Mecum

Transdetermination in *Drosophila*. These movies describe Ernst Hadorn's discovery of transdetermination and Walter Gehring's pioneering study of homeotic mutants, changing body parts into eyes through transcription factors.

Cell-Cell Communication in Development

3

THE FORMATION OF ORGANIZED ANIMAL BODIES, or *morphogenesis*, has been one of the great sources of wonder for humankind. Indeed, the "miracle of life" seems just that—inanimate matter becomes organized in such a way that it lives. The twelfth-century rabbi and physician Maimonides (1190) framed the question of morphogenesis beautifully when he noted that the pious men of his day believed that an angel of God had to enter the womb to form the organs of the embryo. This, the men say, is a miracle. How much more powerful a miracle would life be, Maimonides asked, if the Deity had made matter such that it could generate such remarkable order without a matter-molding angel having to intervene in every pregnancy? The problem addressed today is the secular version of Maimonides' question: How can matter alone construct the organized tissues of the embryo?

The idea of angelic intervention remained prevalent in the embryology of the Renaissance. By the eighteenth century, however, scientific advances had allowed the learned to dispense with the necessity of involving heavenly beings in human conception and development, even though the process remained a mystery. In 1782, the Enlightenment essayist Denis Diderot posed the question of morphogenesis in the fevered dream of a noted physicist. This character could imagine that the body was formed from myriad "tiny sensitive bodies" that collected together to form an aggregate, but he could not envision how this aggregate could become an animal.

Diderot's "tiny sensitive bodies" are what we call cells, and we can break his problem into at least five questions that confront modern embryologists who study morphogenesis:

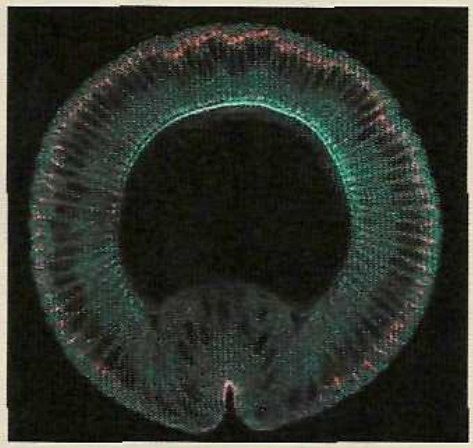
1. *How are separate tissues formed from populations of cells?* For example, how do neural retina cells stick to other neural retina cells rather than becoming part of the pigmented retina or the iris cells next to them? How are the different cell types found in the retina (the three distinct layers of photoreceptors, bipolar neurons, and ganglion cells) arranged such that the retina is functional?
2. *How are organs constructed from tissues?* The retina of the eye forms at a precise distance behind the cornea and the lens. The retina would be useless if it developed behind a bone or in the middle of the kidney. Moreover, neurons from the retina must enter the brain to innervate the regions of the brain cortex that analyze visual information. All these connections must be precisely ordered.
3. *How do organs form in particular locations, and how do migrating cells reach their destinations?* What causes there to be two—and usually only two—kidneys, and how do their ducts form so that they can collect urine made by the filter-

All that you touch
You Change.
All that you Change
Changes you.
The only lasting truth
Is Change.

OCTAVIA BUTLER (1998)

The behaviour of a cell in an embryo depends on the extent to which it listens to its mother or its neighborhood. The size and nature of the noise, the way in which it is heard, and the response are unpredictable and can only be discovered by experimentation.

JONATHAN BARD (1997)



ing tissues of the nephron? Some cells—for instance, the precursors of our pigment cells, germ cells, and blood cells—must travel long distances to reach their final destinations. How are cells instructed to travel along certain routes in our embryonic bodies, and how are they told to stop once they have reached their appropriate destinations?

4. *How do organs and their cells grow, and how is their growth coordinated throughout development?* The cells of all the tissues in the eye must grow in a coordinated fashion if one is to see. Some cells, including most neurons, do not divide after birth. In contrast, the intestine is constantly shedding cells, and new intestinal cells are regenerated each day. The mitotic rate of each tissue must be carefully regulated. If the intestine generated more cells than it sloughed off, it could produce tumorous outgrowths. If it produced fewer cells than it sloughed off, it would soon become nonfunctional. What controls the rate of mitosis in the intestine?
5. *How do organs achieve polarity?* If one were to look at a cross section of the fingers, one would see a certain organized collection of tissues—bone, cartilage, muscle, fat, dermis, epidermis, blood, and neurons. Looking at a cross section of the forearm, one would find the same collection of tissues. But they are arranged very differently. How is it that the same cell types can be arranged in different ways in different parts of the same structure, and that fingers are always at the end of the arm, never in the middle?

Answers to these questions came slowly and are still coming. In the 1850s, Robert Remak (1852, 1855) formulated the cell theory and showed that the fertilized egg divides to produce the myriad "tiny sensitive bodies"—cells—needed to form an embryo. In the mid-twentieth century, E. E. Just (1939) and Johannes Holtfreter (Townes and Holtfreter 1955) predicted that embryonic cells could have differences in their cell membrane components which would enable the formation of organs. In the late twentieth century, these membrane components—the molecules by which embryonic cells adhere to, migrate over, and induce gene expression in neighboring cells—began to be discovered and described. And presently, these pathways are being modeled to understand how the cell integrates the information from its nucleus and from its surroundings to take its place in the community of cells.

As we discussed in Chapter 1 (see Table 1.1), the cells of an embryo are either epithelial cells or mesenchymal cells. The epithelial cells can form tubes and sheets while remaining adhered to one another, whereas the mesenchymal cells often migrate individually and form extensive *extracellular matrices* that keep the individual cells separate. This chapter will discuss the mechanisms of three behaviors requiring cell-cell communication: cell adhesion, cell migration, and cell signaling.

Cell Adhesion

Differential cell affinity

Many of the answers to our five questions about morphogenesis involve the properties of the cell surface. The cell surface looks pretty much the same in all cell types, and many early investigators thought that the cell surface was not even a living part of the cell. We now know that each type of cell has a different set of proteins in its cell membrane, and that some of these differences are responsible for forming the structure of the tissues and organs during development. Observations of fertilization and early embryonic development made by E. E. Just (1939) suggest-

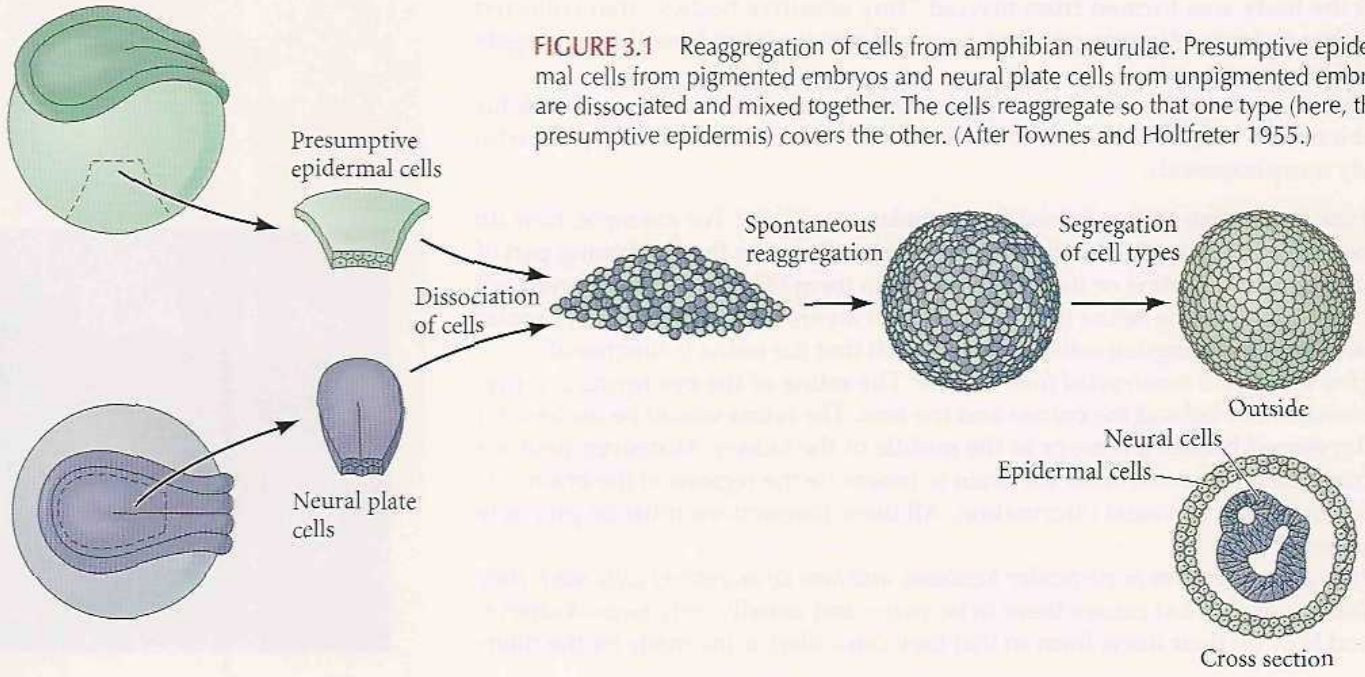


FIGURE 3.1 Reaggregation of cells from amphibian neurulae. Presumptive epidermal cells from pigmented embryos and neural plate cells from unpigmented embryos are dissociated and mixed together. The cells reaggregate so that one type (here, the presumptive epidermis) covers the other. (After Townes and Holtfreter 1955.)

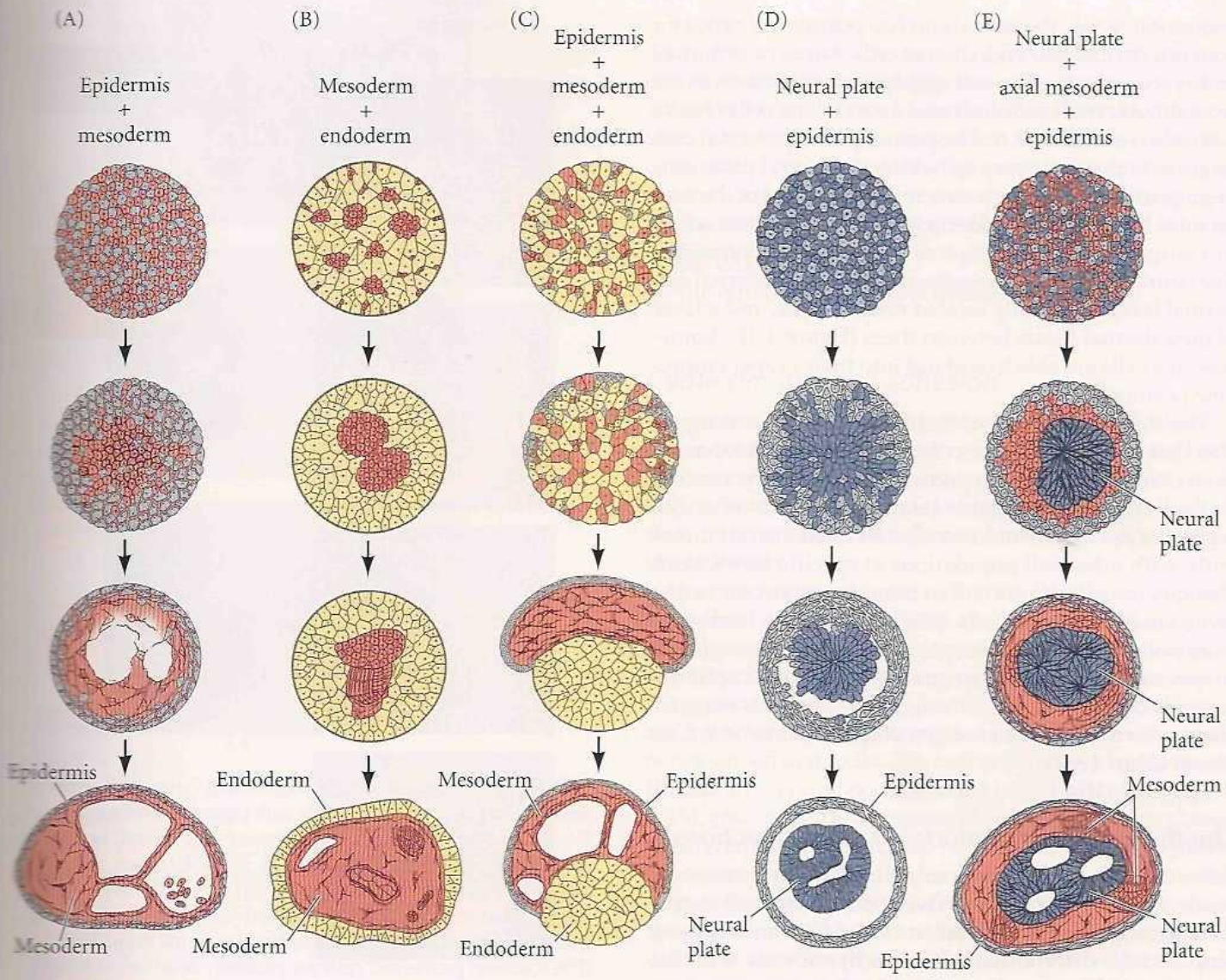


FIGURE 3.2 Sorting out and reconstruction of spatial relationships in aggregates of embryonic amphibian cells. (After Townes and Holtfreter 1955.)

ed that the cell membrane differed among cell types, but the experimental analysis of morphogenesis began with the experiments of Townes and Holtfreter in 1955. Taking advantage of the discovery that amphibian tissues become dissociated into single cells when placed in alkaline solutions, they prepared single-cell suspensions from each of the three germ layers of amphibian embryos soon after the neural tube had formed. Two or more of these single-cell suspensions could be combined in various ways. When the pH of the solution was normalized, the cells adhered to one another, forming aggregates on agar-coated petri dishes. By using embryos from species having cells of different sizes and colors, Townes and Holtfreter were able to follow the behavior of the recombined cells.

The results of their experiments were striking. First, they found that reaggregated cells become spatially segregated. That is, instead of two cell types remaining mixed, each type sorts out into its own region. Thus, when epidermal (ectodermal) and mesodermal cells are brought together in a mixed aggregate, the epidermal cells move to the periphery of the aggregate and the mesodermal cells move to the inside (Figure 3.1). In no case do the recombined cells remain randomly mixed; in most cases, one tissue type completely envelops the other.

Second, the researchers found that the final positions of the reaggregated cells reflect their respective positions in the embryo. The reaggregated mesoderm migrates centrally with respect to the epidermis, adhering to the inner epidermal surface (Figure 3.2A). The mesoderm also migrates centrally with respect to the gut or endoderm (Figure 3.2B). However, when the three germ layers are mixed together, the endoderm separates from the ectoderm and mesoderm and is then enveloped by them (Figure 3.2C). In the final configuration, the ectoderm is on the periphery, the endoderm is internal, and the mesoderm lies in the region between them.

Holtfreter interpreted this finding in terms of **selective affinity**. The inner surface of the ectoderm has a positive affinity for mesodermal cells and a negative affinity for the

endoderm, while the mesoderm has positive affinities for both ectodermal and endodermal cells. Mimicry of normal embryonic structure by cell aggregates is also seen in the recombination of epidermis and neural plate cells (Figure 3.2D; also see Figure 3.1). The presumptive epidermal cells migrate to the periphery as before; the neural plate cells migrate inward, forming a structure reminiscent of the neural tube. When axial mesoderm (notochord) cells are added to a suspension of presumptive epidermal and presumptive neural cells, cell segregation results in an external epidermal layer, a centrally located neural tissue, and a layer of mesodermal tissue between them (Figure 3.2E). Somehow, the cells are able to sort out into their proper embryonic positions.

The third conclusion of Holtfreter and his colleagues was that selective affinities change during development. Such changes should be expected, because embryonic cells do not retain a single stable relationship with other cell types. For development to occur, cells must interact differently with other cell populations at specific times. Such changes in cell affinity are extremely important in the processes of morphogenesis. When tissues from later-stage mammalian and chick embryos were made into single cell suspensions (using the enzyme trypsin, which split the proteins connecting the cells together), the cells reaggregated to form tissuelike arrangements (Moscona 1961; Giudice and Just 1962).

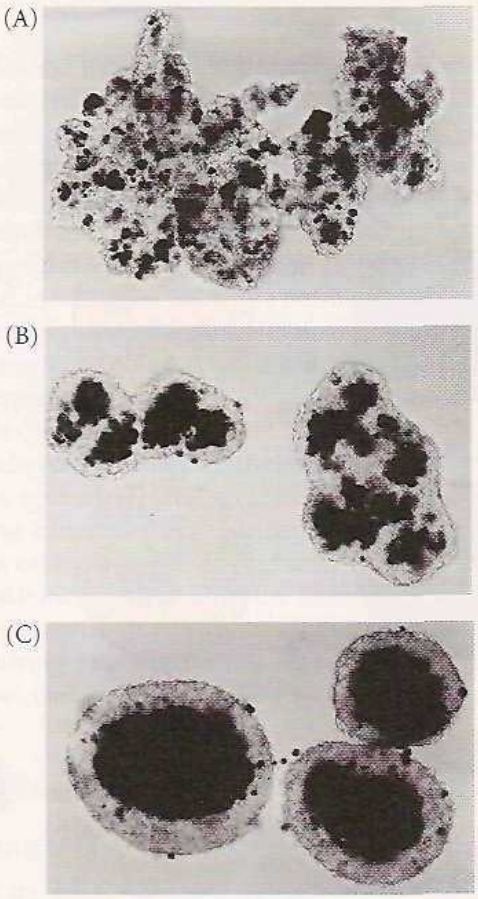


FIGURE 3.3 Aggregates formed by mixing 7-day chick embryo neural retina (unpigmented) cells with pigmented retina cells. (A) Five hours after the single-cell suspensions are mixed, aggregates of randomly distributed cells are seen. (B) At 19 hours, the pigmented retina cells are no longer seen on the periphery. (C) At 2 days, a great majority of the pigmented retina cells are located in a central internal mass, surrounded by the neural retina cells. (The scattered pigmented cells are probably dead cells.) (From Armstrong 1989, courtesy of P. B. Armstrong.)

The thermodynamic model of cell interactions

Cells, then, do not sort randomly, but can actively move to create tissue organization. What forces direct cell movement during morphogenesis? In 1964, Malcolm Steinberg proposed the **differential adhesion hypothesis**, a model that sought to explain patterns of cell sorting based on thermodynamic principles. Using cells derived from trypsinized embryonic tissues, Steinberg showed that certain cell types migrate centrally when combined with some cell types, but migrate peripherally when combined with others. Figure 3.3 illustrates the interactions between pigmented retina cells and neural retina cells. When single-cell suspensions of these two cell types are mixed together, they form aggregates of randomly arranged cells. However, after several hours, no pigmented retina cells are seen on the periphery of the aggregates, and after 2 days, two distinct layers are seen, with the pigmented retina cells lying internal to the neural retina cells. Moreover, such interactions form a hierarchy (Steinberg 1970). If the final position of cell type A is internal to a second cell type B, and the final position of B is internal to a third cell type C, then the final position of A will always be internal to C. For example, pigmented retina cells migrate internally to neural retina cells, and heart cells migrate internally to pigmented retina cells. Therefore, heart cells migrate internally to neural retina cells.

This observation led Steinberg to propose that cells interact so as to form an aggregate with the smallest interfacial free energy. In other words, the cells rearrange themselves into the most thermodynamically stable pattern. If cell types A and B have different strengths of adhesion, and if the strength of A-A connections is greater than the strength of A-B or B-B connections, sorting will occur, with the A cells becoming central. On the other hand, if the strength of A-A connections is less than or equal to the strength of A-B connections, then the aggregate will remain as a random mix of cells. Finally, if the strength of A-A connections is far greater than the strength of A-B connections—in other words, if A and B cells show essentially no adhesivity toward one another—then A cells and B cells will form separate aggregates. According to this hypothesis, the early embryo can be viewed as existing in an equilibrium state until some change in gene activity changes the cell surface molecules. The movements that result seek to restore the cells to a new equilibrium configuration. All that is required for sorting to occur is that cell types differ in the strengths of their adhesion.

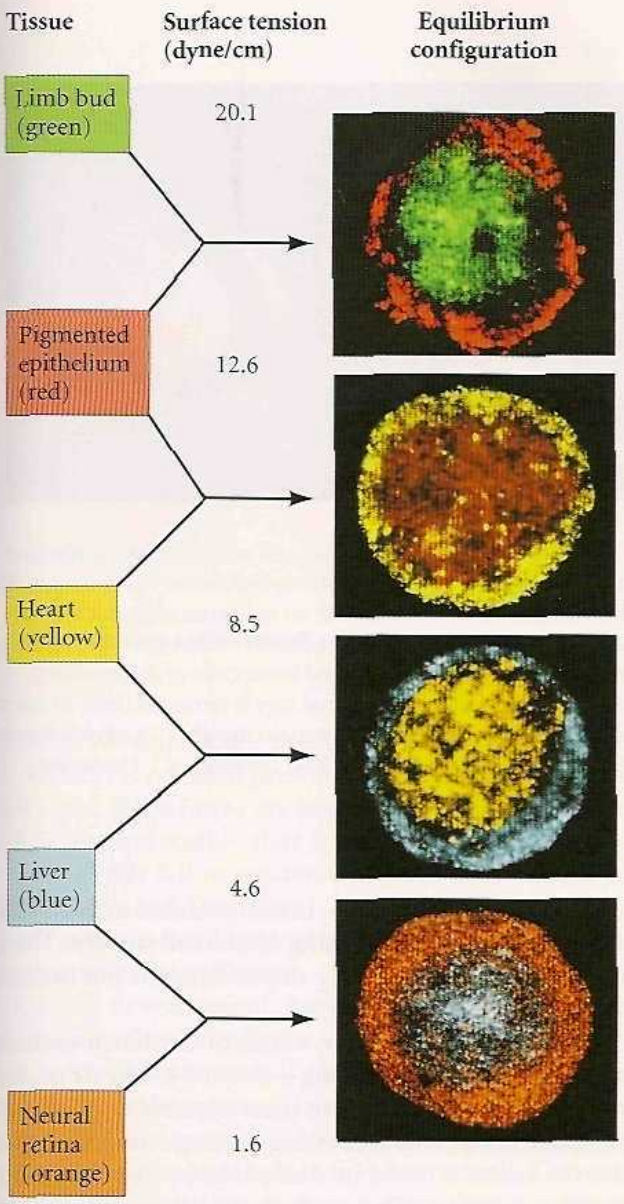


FIGURE 3.4 Hierarchy of cell sorting in order of decreasing surface tensions. The equilibrium configuration reflects the strength of cell cohesion, with the cell types having the greater cell cohesion segregating inside the cells with less cohesion. The images were obtained by sectioning the aggregates and assigning colors to the cell types by computer. The black areas represent cells whose signal was edited out in the program of image optimization. (From Foty et al. 1996, courtesy of M. S. Steinberg and R. A. Foty.)

In 1996, Foty and his colleagues in Steinberg's laboratory demonstrated that this was indeed the case: the cell types that had greater surface cohesion migrated centrally compared with those cells that had less surface tension (Figure 3.4; Foty et al. 1996). In the simplest form of this model, all cells could have the same type of "glue" on the cell surface. The amount of this cell surface product, or the cellular architecture that allows the substance to be differentially distributed across the surface, could cause a difference in the number of stable contacts made between cell types. In a more specific version of this model, the thermodynamic differences could be caused by different types of adhesion molecules (see Moscona 1974). When Holtfreter's studies were revisited using modern techniques, Davis and colleagues (1997) found that the tissue surface tensions of the individual germ layers were precisely those required for the sorting patterns observed both in vitro and in vivo.

See VADE MECUM
The differential adhesion hypothesis

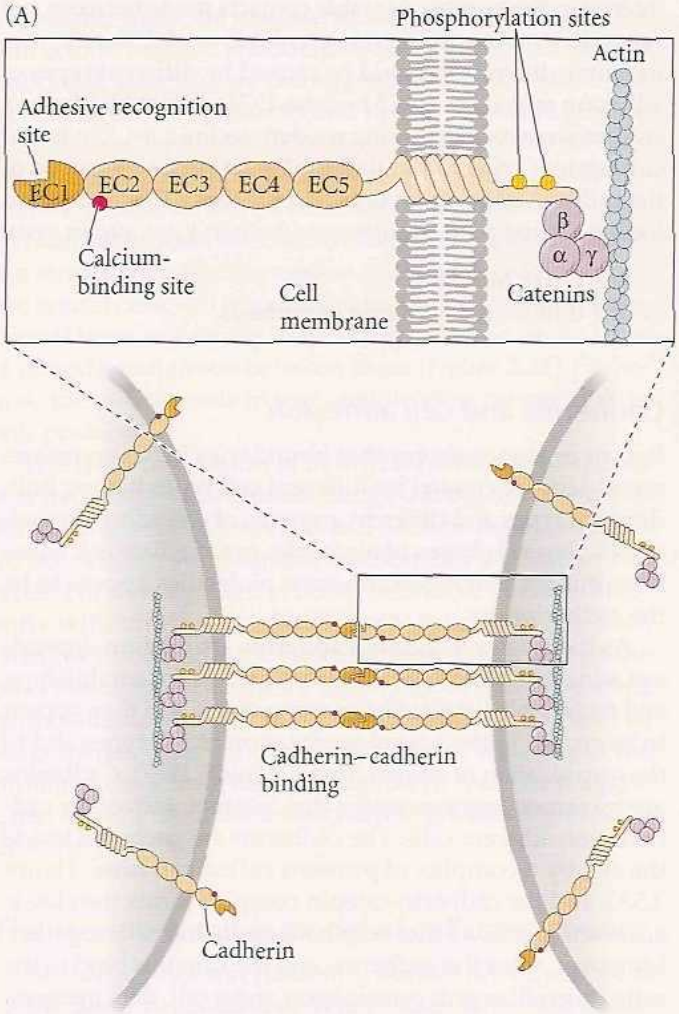
Cadherins and cell adhesion

Recent evidence shows that boundaries between tissues can indeed be created by different cell types having both different types and different amounts of cell adhesion molecules. Several classes of molecules can mediate cell adhesion, but the major cell adhesion molecules appear to be the cadherins.

As their name suggests, **cadherins** are *ca*lcium-*d*ependent ad*he*sion molecules. They are critical for establishing and maintaining intercellular connections, and they appear to be crucial to the spatial segregation of cell types and to the organization of animal form (Takeichi 1987). Cadherins are transmembrane proteins that interact with other cadherins on adjacent cells. The cadherins are anchored inside the cell by a complex of proteins called **catenins** (Figure 3.5A), and the cadherin-catenin complex forms the classic adherens junctions that help hold epithelial cells together. Moreover, since the cadherins and the catenins bind to the actin (microfilament) cytoskeleton of the cell, they integrate the epithelial cells into a mechanical unit. Interfering with cadherin function (by univalent antibodies against cadherin or morpholinos against cadherin mRNA) can prevent the formation of tissues and cause the cells to disaggregate (Figure 3.5B; Takeichi et al. 1979).

Cadherin proteins perform several related functions. First, their external domains serve to adhere cells together. Second, cadherins link to and help assemble the actin cytoskeleton, thereby providing the mechanical forces for forming tubes. Third, cadherins can serve as signaling molecules that change a cell's gene expression.

In vertebrate embryos, several major cadherin types have been identified. **E-cadherin** is expressed on all early mammalian embryonic cells, even at the zygote stage. Later in development, this molecule is restricted to epithelial tissues of embryos and adults. **P-cadherin** is found predominantly on the placenta, where it helps the placenta stick to the uterus (Nose and Takeichi 1986; Kadokawa et al. 1989). **N-cadherin** becomes highly expressed on the cells of the developing central nervous system (Hatta and Takeichi 1986), and it may play roles in mediating neural signals. **R-cadherin** is critical in retina formation (Babb et al. 2005). A class of cadherins called **protocadherins** (Sano et al. 1993) lack the attachment to the actin skeleton through catenins. Expressing similar protocadherins is an impor-



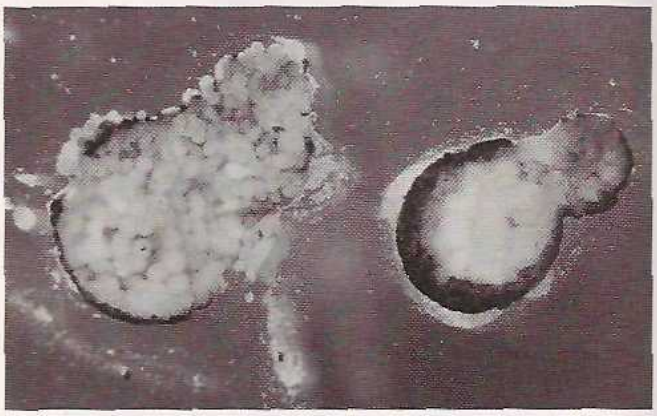


FIGURE 3.5 Cadherin-mediated cell adhesion. (A) Simplified scheme of cadherin linkage to the cytoskeleton via catenins. (B) When an oocyte is injected with an antisense oligonucleotide against a maternally inherited cadherin mRNA (thus preventing the synthesis of the cadherin), the inner cells of the resulting embryo disperse when the animal cap is removed (left). In control embryos (right), the inner cells remain together. (A after Takeichi 1991; B from Heasman et al. 1994, courtesy of J. Heasman.)

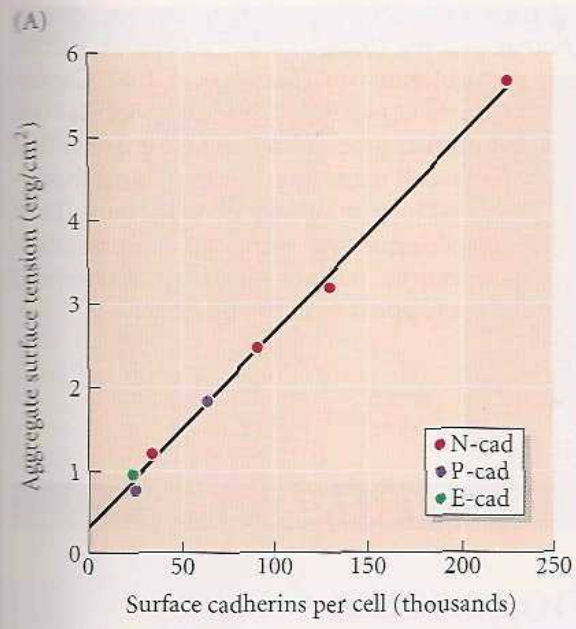
tant means of keeping migrating epithelial cells together; and expressing dissimilar protocadherins is an important way of separating tissues (as when the mesoderm forming the notochord separates from the surrounding mesoderm that will form somites).

Differences in cell surface tension and the tendency of cells to bind together depend on the strength of cadherin interactions (Duguay et al. 2003). This strength can be achieved quantitatively (the more cadherins on the apposing cell surfaces, the tighter the adhesion) or qualitatively (some cadherins will bind to different cadherin types, whereas other cadherins will not bind to different types). The ability to sort cells based on the *amount* of cadherin was first shown when Steinberg and Takeichi (1994) collaborated on an experiment using two cell lines that were identical except that they synthesized different amounts of P-cadherin. When these two groups of cells, each expressing a different amount of cadherin, were mixed, the cells that expressed more cadherin had a higher surface cohesion and migrated internally to the lower-expressing group of cells. Foty and Steinberg (2005) demonstrated that this cadherin-dependent sorting directly correlated with the aggregate surface tension (Figure 3.6). The surface tensions of these aggregates are linearly related to the amount of cadherin they are expressing on the cell surface. The cell sorting hierarchy is strictly dependent on the cadherin interactions between the cells.

Moreover, the energetic value of cadherin-cadherin binding is remarkably strong—about 3400 kcal/mole, or some 200 times stronger than most metabolic protein-protein interactions. This free energy change associated with cadherin function could be dissipated by depolymerizing the actin skeleton. The underlying actin cytoskeleton appears to be crucial in organizing the cadherins in a manner that allows them to form remarkably stable linkages between cells (Foty and Steinberg 2004).

Qualitative interactions are also important. Duguay and colleagues (2003) showed, for instance, that R-cadherin and β-cadherin do *not* bind well to each other, and in these interactions the type of cadherin expressed becomes important. In another example, the expression of N-cadherin is important in separating the precursors of the neural cells from the precursors of the epidermal cells. All early embryonic cells originally contain E-cadherin, but those cells destined to become the neural tube lose E-cadherin and gain N-cadherin. If epidermal cells are experimentally made to express N-cadherin, or if N-cadherin synthesis is blocked in the prospective neural cells, the border between the nervous system and skin fails to form properly (Figure 3.7; Kintner 1993).

The timing of particular developmental events can also depend on cadherin expression. For instance, N-cadherin appears in the mesenchymal cells of the developing chick leg just before these cells condense and form nodules of



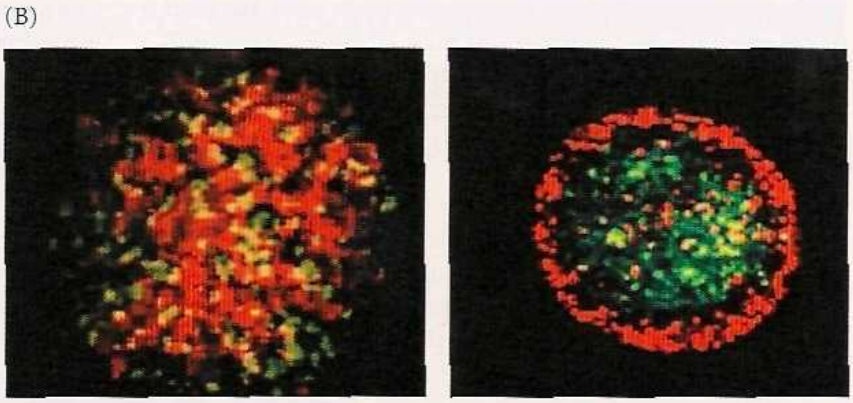


FIGURE 3.6 Importance of the amount of cadherin for correct morphogenesis. (A) Aggregate surface tension correlates with the number of cadherin molecules on the cell membranes. (B) Sorting out of two subclones having different amounts of cadherin on their cell surfaces. The green-stained cells had 2.4 times as many N-cadherin molecules in their membrane as did the other cells. (These cells had no normal cadherin genes being expressed.) At 4 hours of incubation, the cells are randomly distributed, but after 24 hours of incubation, the red cells (with a surface tension of about 2.4 erg/cm^2) have formed an envelope around the more tightly cohering (5.6 erg/cm^2) green cells. (After Foty and Steinberg 2005, photographs courtesy of R. Foty.)

cartilage (which are the precursors of the limb skeleton). N-cadherin is not seen prior to condensation, nor is it seen afterward. If the limbs are injected just prior to condensation with antibodies that block N-cadherin, the mesenchyme cells fail to condense and cartilage fails to form (Oberlander and Tuan 1994). It therefore appears that the signal to begin cartilage formation in the chick limb is the appearance of N-cadherin.

During development, the many cadherins often work with other adhesion systems. For example, one of the most critical times in a mammal's life occurs soon after conception, as the embryo passes from the oviduct and enters the uterus. If development is to continue, the embryo must adhere to and embed itself in the uterine wall. That is why the first differentiation event in mammalian development

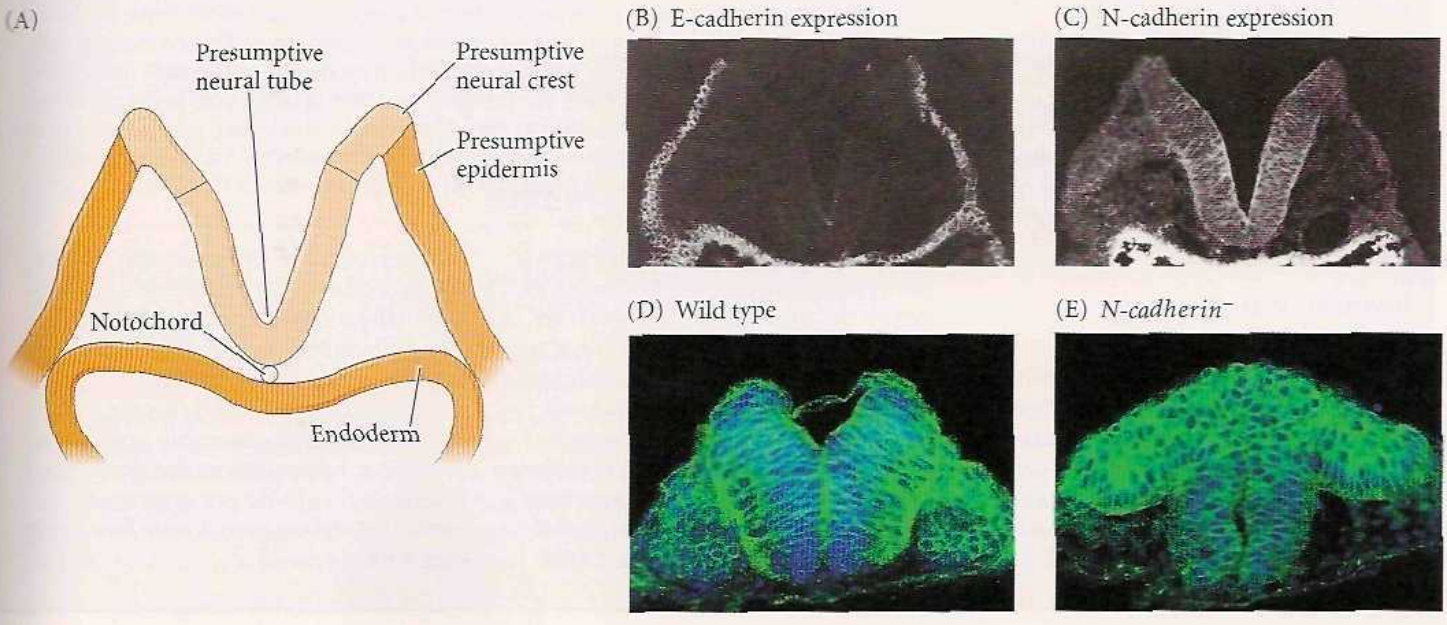


FIGURE 3.7 Importance of the types of cadherin for correct morphogenesis. (A–C) Neural and epidermal tissues in a cross section of a mouse embryo showing the domains of E-cadherin expression (B) and N-cadherin expression (C). N-cadherin is critical for separation of presumptive epidermal and neural tissues during organogenesis. (D,E) The neural tube separates cleanly from surface epidermis in wild-type zebrafish embryos (D) but not in mutant embryos where N-cadherin fails to be made (E). In these images, the cell outlines are stained green with antibodies to β-catenin, while the cell interiors are stained blue. (B,C, photographs by K. Shimamura and H. Matsunami, courtesy of M. Takeichi; D,E from Hong and Brewster 2006, courtesy of R. Brewster.)

distinguishes the **trophoblast** cells (the outer cells that bind to the uterus) from the **inner cell mass** (those cells that will generate the embryo and eventually the mature organism). This differentiation process occurs as the embryo travels from the upper regions of the oviduct on its way to the uterus.

Trophoblast cells are endowed with several adhesion molecules that anchor the embryo to the uterine wall. First, they contain both E- and P-cadherins (Kadokawa et al. 1989), and these two molecule types recognize similar cadherins on the uterine cells. Second, they have receptors (integrin proteins) for the collagen and the heparan sulfate glycoproteins of the uterine wall (Farach et al. 1987; Carson et al. 1988, 1993; Cross et al. 1994). Third, trophoblast cell surfaces have a modified glycosyltransferase enzyme that extends out from the cell membrane and can bind to specific carbohydrate residues on uterine glycoproteins (Dutt et al. 1987). For something as important as the implantation of the mammalian embryo, it is not surprising that several cell adhesion systems appear to work together.

SIDELIGHTS & SPECULATIONS

Shape Change and Epithelial Morphogenesis: "The Force Is Strong in You"

Epithelial cells form sheets and tubes. Their ability to form such structures often depends on cell shape changes that usually involve cadherins and the actin cytoskeleton. The extracellular domains of cadherins bind groups of cells together, while the intracellular domains of the cadherins alter the actin cytoskeleton. The proteins mediating this cadherin-dependent remodeling of the cytoskeleton are usually (1) small GTPases, which convert soluble actin into fibrous actin cables that anchor at the cadherins and (2) non-muscle myosin, which provides the energy for actin contraction. Two examples of cadherin-dependent remodeling of the cytoskeleton are the formation of the neural tube in vertebrates and the internalization of the mesoderm in *Drosophila*. In both cases, the cells (neural ectoderm in frogs, mesoderm in *Drosophila*) are on the outside of the embryo, and it is critical that they migrate to the inside.

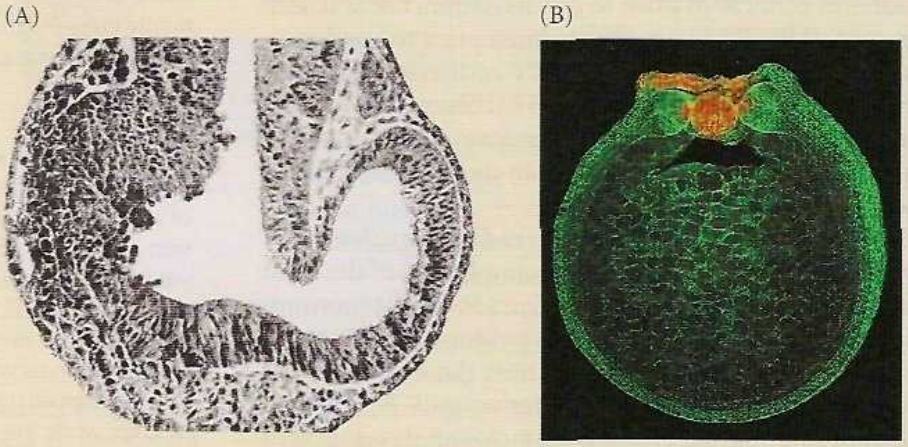


Figure 3.8 Importance of cadherin in cell adhesion and morphogenetic movements. (A) Frog gastrula injected with a nonfunctioning N-cadherin gene on one side. The uninjected side (right) develops normally; on the injected side (left), the epidermis and neural tissue fail to separate. (B) Cross section of a *Xenopus* neurula stained for F-actin (green). The red-staining cells of the neural plate (uppermost cells) and notochord (the mesodermal rod beneath them) have in them a morpholino oligonucleotide that prevents N-cadherin mRNA from being translated. The neural plate cells fail to invaginate into the embryo or to form a neural tube because of the loss of the N-cadherin-based actin assembly in their apical cytoplasm. (A from Kintner et al. 1992, courtesy of C. Kintner; B from Nandadasa et al. 2009, courtesy of C. Wylie.)

Involution of the frog neural tube

In the early frog embryo, each cell's membrane can contain several types of cadherins. Each cell of the gastrula is covered with C-cadherin. However, the presumptive *neural tube* ectoderm cell membranes also contain N-cadherin, concentrated in the apical region of each of these cells; the presumptive *epidermal* cells of the ectoderm contain E-cadherin, which is expressed on the lateral and basal surfaces of the cell. The actin organized in the apical region of the neural cells causes them to change shape and enter the internal region of the embryo as a neural tube. The actin organized on the lateral sides of the epidermal cells enables the movements of the epidermal (skin) cells over the surface of the embryo. If N-cadherin is experimentally removed from a frog gastrula, the cells still adhere (thanks to the C-cadherin that is still present), but the actin (and the activated myosin that binds to it) fail to assemble apically, so there is no neurulation: the presumptive neural cells do not enter the embryo, and no neural tube forms (Figure 3.8).

SIDELIGHTS & SPECULATIONS (Continued)

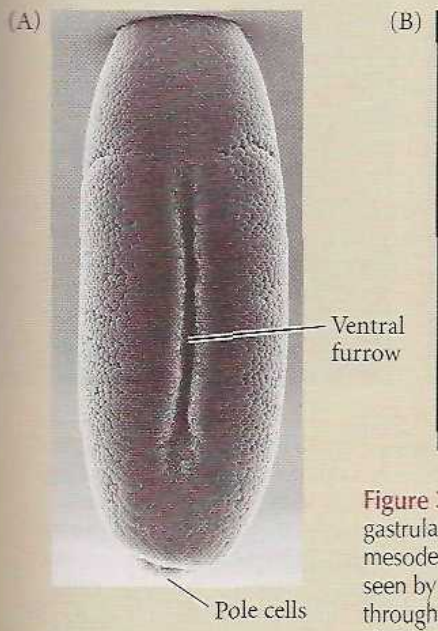


Figure 3.9 Ventral furrow formation during *Drosophila* gastrulation internalizes the cells that will become the mesoderm. (A) Ventral furrow in the *Drosophila* embryo seen by scanning electron microscopy. (B) Cross section through the center of such an embryo, demonstrating cell shape changes and the redistribution of protein products along the dorsal-ventral axis. Apical constrictions of the ventral cells can be observed. (A courtesy of F. R. Turner; B courtesy of V. Kölsch and M. Leptin.)

Drosophila mesoderm formation

In *Drosophila*, the mesoderm is formed from epithelial cells on the ventral side of the embryo. These cells form a furrow and then migrate inside the embryo (Figure 3.9). To create this furrow, the cube-shaped cells become wedge-shaped, constricting at their apical surfaces. This transition creates a force that pushes the ventral cells inside the embryo. What creates this force? The apical constriction is brought about by the rearrangement of actin microfilaments and myosin II (a "non-skeletal myosin") to the apical end of the cell (Figure 3.10). Actin microfilaments are part of the cytoskeleton and are often found on the periphery of the cell. (Indeed, they are critical for producing the cleavage furrows of cell division.) The instructions for this apical constriction appear to emanate from the *Twist* gene, which

(Continued on next page)

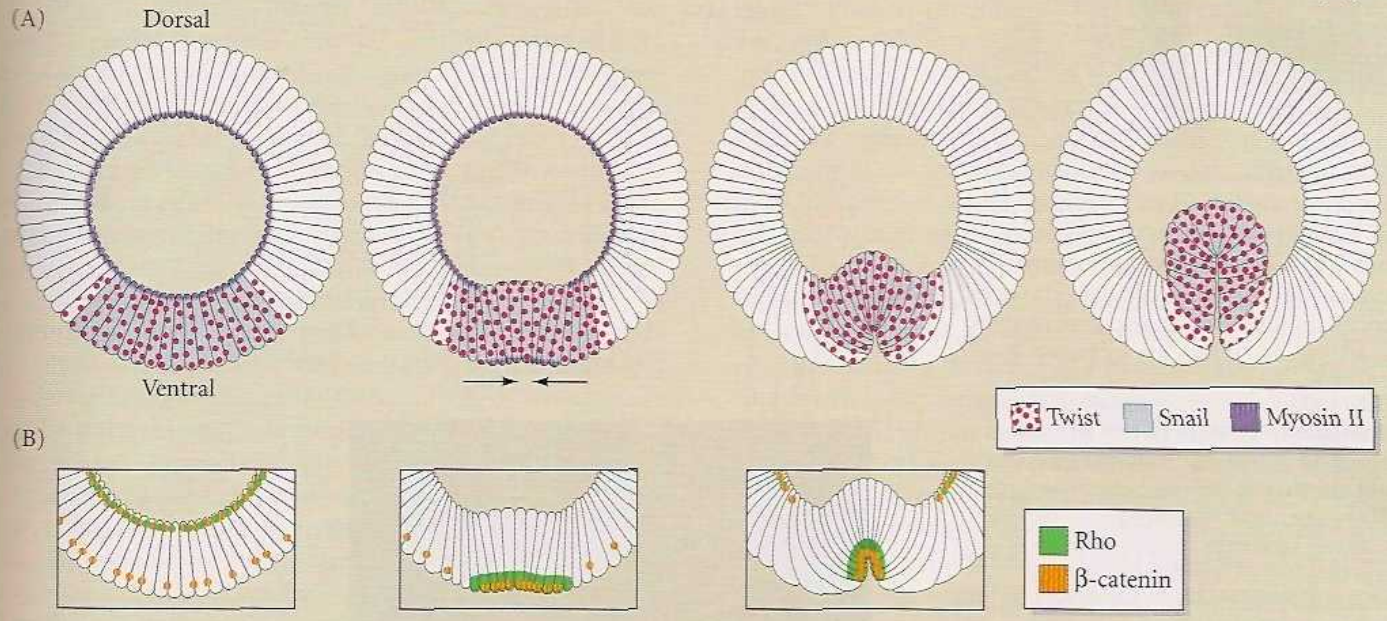


Figure 3.10 Getting mesodermal cells inside the embryo during *Drosophila* gastrulation by regulation of the cytoskeleton. (A) Schematic representation of ventral furrow formation shown as cross sections through the *Drosophila* embryo, progressing in time from left to right. The ventral cells are defined by the expression of transcription factors Twist and Snail. These cells accumulate myosin II at their apical surfaces. When myosin II interacts with actin already present, the cells begin to constrict apically and thus invaginate. (B) Close-ups of the ventral domain. Before the initiation of ventral furrow formation, Rho (green) and β-catenin (orange) both reside along the basal surface (facing the interior of the embryo) of the ventral cells. β-catenin is also found in a subapical region in all cells. Formation of the ventral furrow begins with the relocalization of Rho and β-catenin, which move from the basal surface to accumulate apically, at the opposite end of the cell. (From Kölsch et al. 2007.)

SIDELIGHTS & SPECULATIONS (Continued)

is only expressed in the nuclei of the ventralmost cells (Kölsch et al. 2007). The Twist protein activates other genes, whose protein products cause the actin cytoskeleton to build up on the apical side of the cell. This build-up is accomplished by the binding of a small GTPase (Rho) and β-catenin to E-cadherin on the apical portion of the cell membrane in these most ventral cells. Once stabilized, the actin-myosin complex in the cell's apical cortex constricts like the drawstring of a purse, causing the cells to change shape, buckle inward, and enter the embryo to form the mesoderm.

External signals: Insect trachea

In the above cases, the instructions for folding come from inside the cell. Instructions for cell shape change can also arise outside the cell. For instance, the tracheal (respiratory) system in *Drosophila* embryos develops from epithelial sacs. The approximately 80 cells in each of these sacs become reorganized into primary, secondary, and tertiary branches without any cell division or cell death (Ghabrial and Krasnow 2006). This reorganization is initiated when nearby cells secrete a protein called Branchless, which acts as a chemoattractant.* Branchless binds to a receptor on the cell membranes of the epithelial cells. The cells receiving the most Branchless protein lead the rest, while the followers (connected to each other by cadherins) receive a signal from the leading cells to form the tracheal tube (Figure 3.11). It is the lead cell that will change its shape (by rearranging its actin-myosin cytoskeleton via a small GTPase-mediated process, just like the mesodermal cells) to migrate and to form the secondary branches. During this migration, cadherin proteins are regulated such that the epithelial cells can migrate over one another to form a tube while keeping their integrity as an epithelium (Cela and Llimagas 2006).

But another external force is also at work. The dorsalmost secondary branches of the sac move along a groove that forms between the developing muscles. These tertiary cell migrations cause the trachea to become segmented around the musculature (Franch-Marro and Casanova 2000). In this way, the respiratory tubes are placed close to the larval musculature.

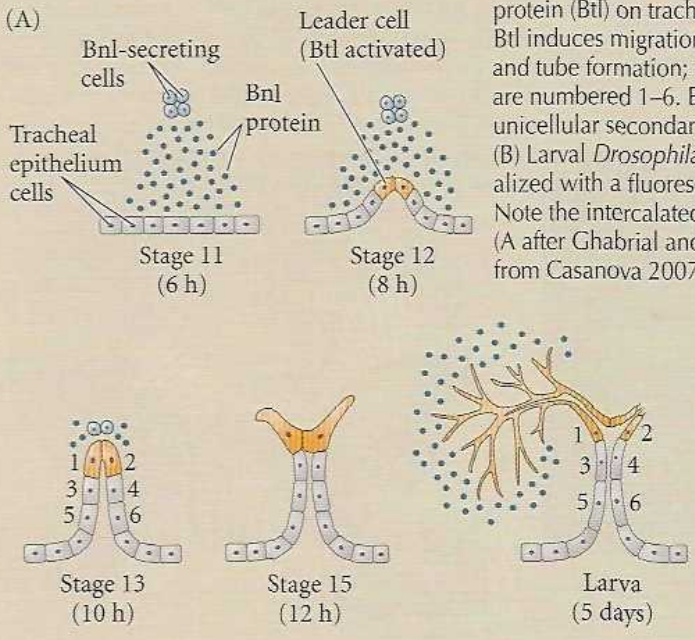


FIGURE 3.11 Tracheal development in *Drosophila*. (A) Diagram of dorsal tracheal branch budding from tracheal epithelium. Nearby cells secrete Branchless protein (Bnl; blue dots), which activates Breathless protein (Btl) on tracheal cells. The activated Btl induces migration of the leader cells and tube formation; the dorsal branch cells are numbered 1–6. Branchless also induces unicellular secondary branches (stage 15). (B) Larval *Drosophila* tracheal system visualized with a fluorescent red antibody. Note the intercalated branching pattern. (A after Ghabrial and Krasnow 2006; B from Casanova 2007.)

*Chemoattractants are usually diffusible molecules that attract a cell to migrate along an increasing concentration gradient toward the cells secreting the factor. There are also *chemorepulsive* factors that send the migrating cells in an opposite direction. Generally speaking, *chemotactic factors*—soluble factors that cause cells to move in a particular direction—are assumed to be chemoattractive unless otherwise described.

Cell Migration

Cell migration is a common feature of both epithelial and mesenchymal cells (Kurosaka and Kashina 2008). The cells of the embryo move extensively during gastrulation to form the three germ layers; the neural tube folds into the vertebrate embryo; the mesoderm folds into the fly embryo; and the precursors of the germ cells, blood cells, and pigment cells undergo individual and extensive migrations. In epithelia, the motive force for migration is usual-

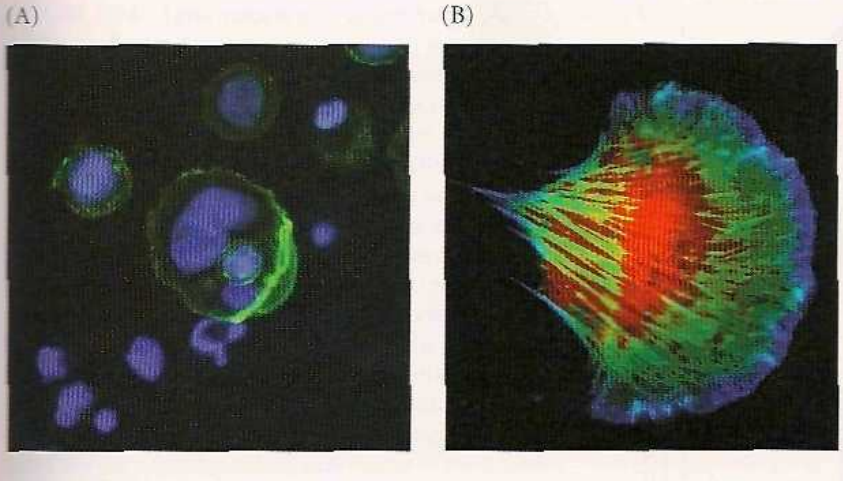


FIGURE 3.12 Cell migration. (A) Polarization of a migrating cell. Cell interiors are stained blue. The actin (stained green with antibodies to β-catenin) becomes redistributed by interstitial flow to the leading edge of the cell. (B) In the lamellipodium of a migrating mesenchymal cell, the ratio of filamentous actin to globular actin is visualized by different colors, blue being the highest filamentous actin level, red being the lowest. (A from Shields et al. 2007, courtesy of M. A. Swartz; B from Cramer et al. 2002, courtesy of L. Cramer.)

ly provided by the cells at the edge of the sheet, and the rest of the cells follow passively. In mesenchymal cell migration, individual cells become polarized and migrate through the extracellular milieu. In both cases, there is a wide-scale reorganization of the actin cytoskeleton (Figure 3.12A). The first stage of migration is **polarization**, wherein a cell defines its front and its back. Polarization can be directed by diffusing signals (such as a chemotactic protein) or by signals from the extracellular matrix. These signals reorganize the cytoskeleton so that the cell has a front and a back, and so that the front part of the cell becomes structurally different from the back of the cell (Rodionov and Borisy 1997; Malikov et al. 2005).

The second stage of migration is the protrusion of the cell's leading edge. The mechanical force for this is the polymerization of the actin microfilaments at the cell membrane, creating long parallel bundles (forming filopodia) or broad sheets (forming lamellipodia; Figure 3.12B.) The membrane-bound Rho G-proteins activate the WASP-N proteins to nucleate actin and connect it to cadherins and the cell membrane (Co et al. 2007).

The third stage of migration involves the **adhesion** of the cell to its extracellular substrate. The moving cell needs something to push on, and attaches to the surrounding matrix. The key molecules in this process (as we will detail later in this chapter) are *integrin* proteins. Integrins span the cell membrane, connecting the extracellular matrix outside the cell to the actin cytoskeleton on the inside of the cell. These connections of actin to integrin form **focal adhesions** on the cell membrane where the membrane contacts the extracellular matrix. Myosin and its regulators provide the motive force along these actin microfilaments, and they are linked with the lamellipodial actin at the sites of adhesion (Giannone et al. 2007).

The fourth stage of cell migration concerns the release of adhesions in the rear, allowing the cell to migrate in the forward direction. It is probable that stretch-sensitive calcium channels are opened and that the released calcium ions activate proteases that destroy the focal adhesion sites.

Cell Signaling

Induction and competence

From the earliest stages of development through the adult, cell differentiation and behavior (such as adhesion, migration, and cell division) are regulated by signals from one cell being received by another cell. Indeed, these interactions (which are often reciprocal) are what allow organs to be constructed. The development of the vertebrate eye is a classic example used to describe the modus operandi of tissue organization via intercellular interactions.

In the vertebrate eye, light is transmitted through the transparent corneal tissue and focused by the lens tissue (the diameter of which is controlled by muscle tissue), eventually impinging on the tissue of the neural retina. The precise arrangement of tissues in the eye cannot be disturbed without impairing its function. Such coordination in the construction of organs is accomplished by one group of cells changing the behavior of an adjacent set of cells, thereby causing them to change their shape, mitotic rate, or cell fate. This kind of interaction at close range between two or more cells or tissues of different histories and properties is called **induction**.

There are at least two components to every inductive interaction. The first component is the **inducer**: the tissue that produces a signal (or signals) that changes the cellular behavior of the other tissue. Often, this signal is a secreted protein called a *paracrine factor*. **Paracrine factors** are proteins made by a cell or a group of cells that alter the behavior or differentiation of adjacent cells. In contrast to endocrine factors (i.e., hormones), which travel through the blood and exert their effects on cells and tissues far away, paracrine factors are secreted into the extracellular space and influence their close neighbors. (The Branchless protein, mentioned in the Sidelights & Speculations on p. 78, is such a factor). The second component, the **responder**, is the tissue being induced. Cells of the responding tissue must have both a receptor protein for the inducing

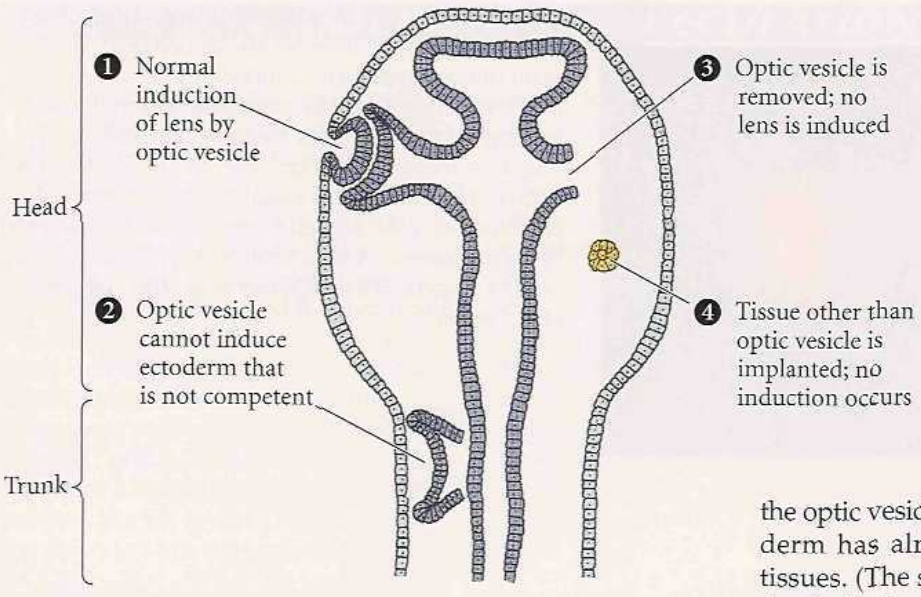


FIGURE 3.13 Ectodermal competence and the ability to respond to the optic vesicle inducer in *Xenopus*. The optic vesicle is able to induce lens formation in the anterior portion of the ectoderm (1) but not in the presumptive trunk and abdomen (2). If the optic vesicle is removed (3), the surface ectoderm forms either an abnormal lens or no lens at all. (4) Most other tissues are not able to substitute for the optic vesicle.

factor (the receptor for Branchless is the Breathless protein) and the *ability* to respond to the signal. The ability to respond to a specific inductive signal is called **competence** (Waddington 1940).

Even if receptor proteins are present, not every tissue type is competent to respond to an inducer's signal. For instance, if the optic vesicle (the presumptive retina) of a *Xenopus laevis* embryo is placed in an ectopic location underneath the head ectoderm (i.e., in a different part of the head from where the frog's optic vesicle normally occurs), it will induce that ectoderm to form lens tissue. Only the optic vesicle appears to be able to do this; therefore, it is an inducer. However, if the optic vesicle is placed beneath ectoderm in the flank or abdomen of the same organism, that ectoderm will not be able to form lens tissue. Only head ectoderm is *competent* to respond to the signals from the optic vesicle by producing a lens (Figure 3.13; Saha et al. 1989; Grainger 1992).

Often, one induction will give a tissue the competence to respond to another inducer. Studies on amphibians suggest that the first inducers of the lens may be the foregut endoderm and heart-forming mesoderm that underlie the lens-forming ectoderm during the early and mid gastrula stages (Jacobson 1963, 1966). The anterior neural plate may produce the next signals, including a signal that promotes the synthesis of Pax6 transcription factor in the anterior ectoderm (Figure 3.14; Zygar et al. 1998). Pax6 is important in providing the competence to respond to the inducers from the optic cup (Fujiwara et al. 1994). Thus, although the optic vesicle appears to be *the* inducer, the anterior ectoderm has already been induced by at least two other tissues. (The situation is like that of the player who kicks the "winning goal" of a soccer match.) The optic vesicle appears to secrete two paracrine factors, one of which may be BMP4 (Furuta and Hogan 1998), a protein that is received by the lens cells and induces the production of the Sox2 transcription factors. The other paracrine factor is Fgf8, a signal that induces the appearance of the L-Maf transcription factor (Ogino and Yasuda 1998; Vogel-Höpker et al. 2000). As we saw in Chapter 2, the combination of Pax6, Sox2, and L-Maf in the ectoderm is needed for the production of the lens and the activation of lens-specific genes such as δ-crystallin.

Cascades of induction: Reciprocal and sequential inductive events

Another feature of induction is the *reciprocal nature* of many inductive interactions. To continue the above example, once the lens has formed, it induces other tissues. One of these responding tissues is the optic vesicle itself; thus the inducer becomes the induced. Under the influence of factors secreted by the lens, the optic vesicle becomes the optic cup and the wall of the optic cup differentiates into two layers, the pigmented retina and the neural retina (Figure 3.15; Cvekl and Piatigorsky 1996; Strickler et al. 2007). Such interactions are called **reciprocal inductions**.

At the same time, the lens is inducing the ectoderm above it to become the cornea. Like the lens-forming ectoderm, the cornea-forming ectoderm has achieved a particular competence to respond to inductive signals, in this case the signals from the lens (Meier 1977; Thut et al. 2001). Under the influence of the lens, the corneal ectoderm cells become columnar and secrete multiple layers of collagen. Mesenchymal cells from the neural crest use this collagen matrix to enter the area and secrete a set of proteins (including the enzyme hyaluronidase) that further differentiate the cornea. A third signal, the hormone thyroxine, dehydrates the tissue and makes it transparent (Hay 1980; Bard 1990). Thus, there are sequential inductive events, and multiple causes for each induction.

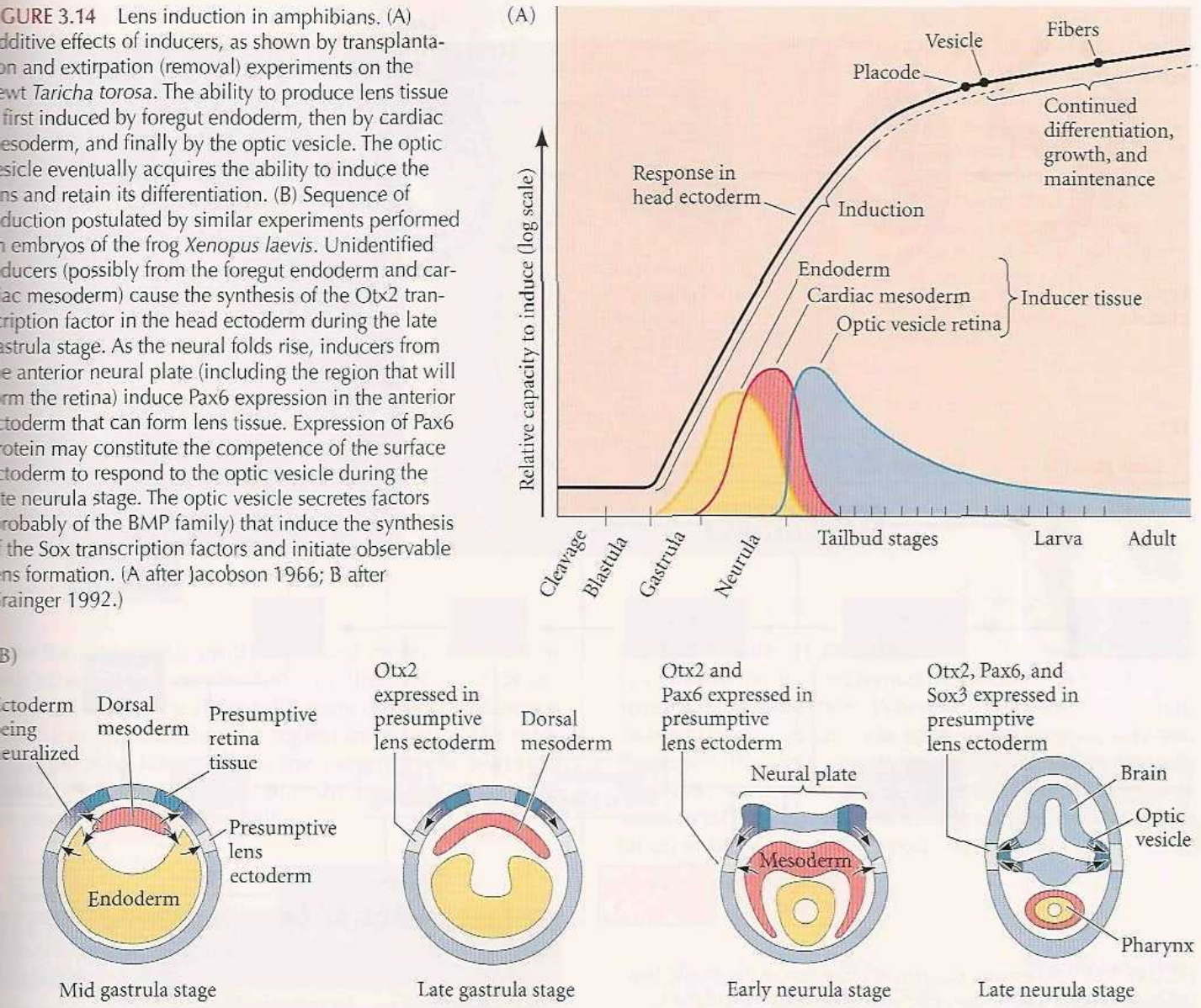


FIGURE 3.14 Lens induction in amphibians. (A) Additive effects of inducers, as shown by transplantation and extirpation (removal) experiments on the newt *Taricha torosa*. The ability to produce lens tissue is first induced by foregut endoderm, then by cardiac mesoderm, and finally by the optic vesicle. The optic vesicle eventually acquires the ability to induce the lens and retain its differentiation. (B) Sequence of induction postulated by similar experiments performed on embryos of the frog *Xenopus laevis*. Unidentified inducers (possibly from the foregut endoderm and cardiac mesoderm) cause the synthesis of the Otx2 transcription factor in the head ectoderm during the late gastrula stage. As the neural folds rise, inducers from the anterior neural plate (including the region that will form the retina) induce Pax6 expression in the anterior ectoderm that can form lens tissue. Expression of Pax6 protein may constitute the competence of the surface ectoderm to respond to the optic vesicle during the late neurula stage. The optic vesicle secretes factors (probably of the BMP family) that induce the synthesis of the Sox transcription factors and initiate observable lens formation. (A after Jacobson 1966; B after Grainger 1992.)

Another principle can be seen in such reciprocal inductions: a structure does not need to be fully differentiated in order to have a function. As we will detail in Chapter 9, the optic vesicle induces the lens placode before it becomes the retina; the lens placode (the prospective lens) reciprocates by inducing the optic vesicle before the lens forms its characteristic fibers. Thus, before a tissue has its "adult" functions, it has critically important transient functions in building the organs of the embryo.

Instructive and permissive interactions

Howard Holtzer (1968) distinguished two major modes of inductive interaction. In **instructive interaction**, a signal from the inducing cell is necessary for initiating new gene expression in the responding cell. Without the inducing cell, the responding cell is not capable of differentiating in that particular way. For example, when the optic vesicle is experimentally placed under a new region of the head ectoderm and causes that region of the ectoderm to form a lens, that is an instructive interaction.

The second type of inductive interaction is **permissive interaction**. Here, the responding tissue has already been specified, and needs only an environment that allows the expression of these traits. For instance, many tissues need a solid substrate containing *fibronectin* or *laminin* in order to develop. The fibronectin or laminin does not alter the type of cell that is produced, but it enables what has already been determined to be expressed.*

*It is easy to distinguish permissive and instructive interactions by an analogy with a more familiar situation. This textbook is made possible by both permissive and instructive interactions. A reviewer can convince me to change the material in the chapters. This is an instructive interaction, as the information expressed in the book is changed from what it would have been. However, the information in the book could not be expressed at all without permissive interactions with the publisher and printer.

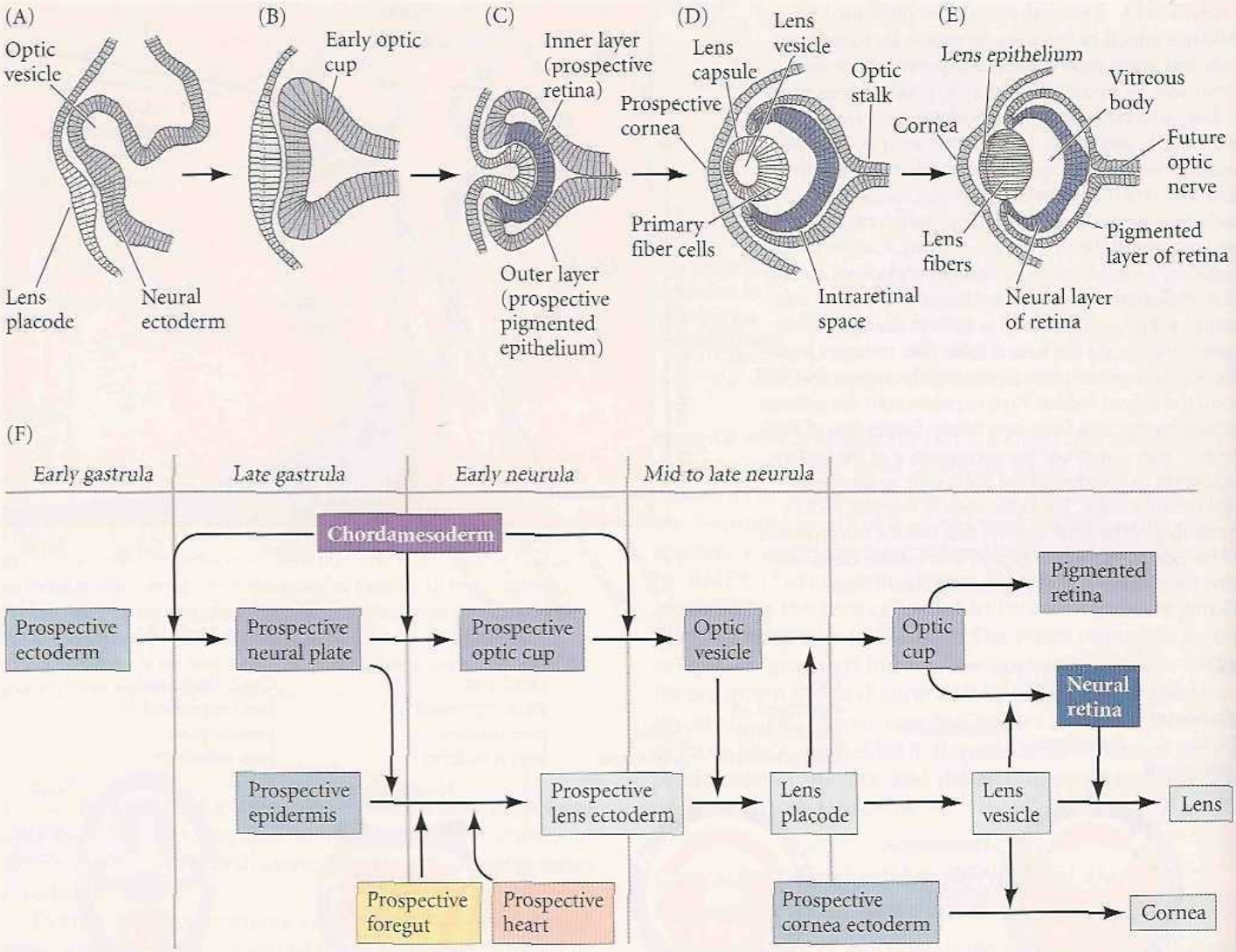


FIGURE 3.13 Schematic diagram of induction of the mouse lens. (A) At embryonic day 9, the optic vesicle extends toward the surface ectoderm from the forebrain. The lens placode (the prospective lens) appears as a local thickening of the surface ectoderm near the optic vesicle. (B) By the middle of day 9, the lens placode has enlarged and the optic vesicle has formed an optic cup. (C) By the middle of day 10, the central portion of the lens-forming ectoderm invaginates while the two layers of the retina become distinguished. (D) By the middle of day 11, the lens vesicle has formed. (E) By day 13, the lens consists of anterior cuboidal epithelial cells and elongating posterior fiber cells. The cornea develops in front of the lens. (F) Summary of some of the inductive interactions during eye development. (A–E after Cvekl and Piatigorsky 1996.)

Epithelial-mesenchymal interactions

Some of the best-studied cases of induction involve the interactions of sheets of epithelial cells with adjacent mesenchymal cells. All organs consist of an epithelium and an associated mesenchyme, so these **epithelial-mesenchymal interactions** are among the most important phenomena in nature. Some examples are listed in Table 3.1.

REGIONAL SPECIFICITY OF INDUCTION Using the induction of cutaneous (skin) structures as our examples, we will look at the properties of epithelial-mesenchymal interactions. The first of these properties is the regional specificity of induction. Skin is composed of two main tissues: an outer epidermis (an epithelial tissue derived from ectoderm), and a dermis (a mesenchymal tissue derived from mesoderm). The chick epidermis secretes proteins that signal the underlying dermal cells to form condensations, and the condensed dermal mesenchyme responds by secreting factors that cause the epidermis to form regionally specific cutaneous structures (Figure 3.16; Nohno et al. 1995; Ting-Berreth and Chuong 1996). These structures can be the broad feathers of the wing, the narrow feathers of the thigh, or the scales and claws of the feet.

As Figure 3.17 demonstrates, the dermal mesenchyme is responsible for the regional specificity of induction in the competent epidermal epithelium. Researchers can sep-

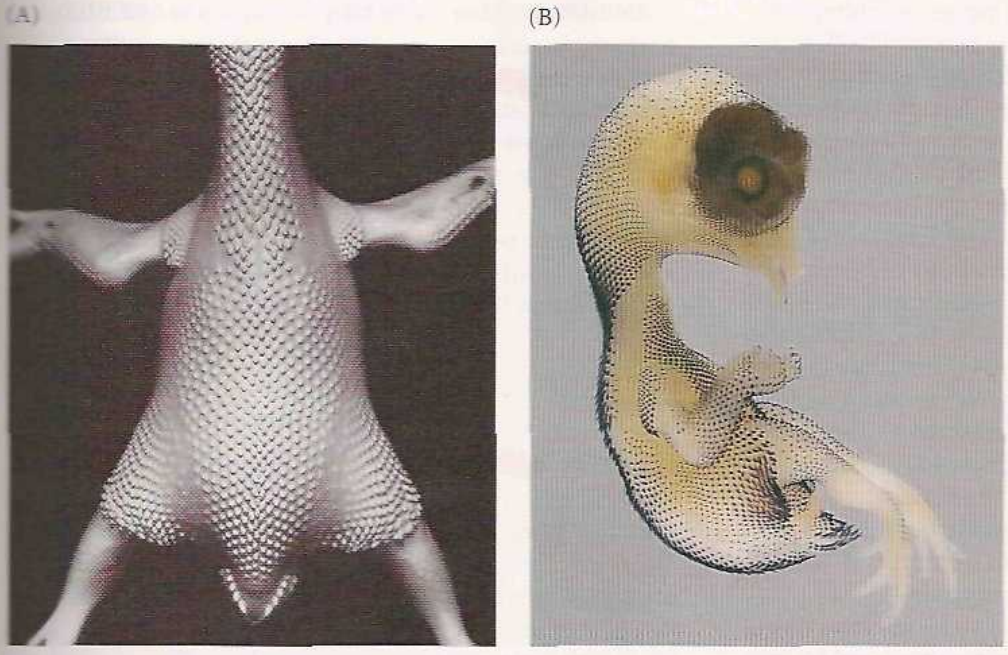


FIGURE 3.16 Feather induction in the chick. (A) Feather tracts on the dorsum of a 9-day chick embryo. Note that each feather primordium is located between the primordia of adjacent rows. (B) In situ hybridization of a 10-day chick embryo shows Sonic hedgehog expression (dark spots) in the ectoderm of the developing feathers and scales. (A courtesy of P. Sengal; B courtesy of W.-S. Kim and J. F. Fallon.)

arate the embryonic epithelium and mesenchyme from each other and recombine them in different ways (Saunders et al. 1957). The same epithelium develops cutaneous structures according to the region from which the mesenchyme was taken. Here, the mesenchyme plays an instructive role, calling into play different sets of genes in the responding epithelial cells.

TABLE 3.1 Some epithelial-mesenchymal interactions

Organ	Mesenchymal component	Epithelial component
Cutaneous structures (hair, feathers, sweat glands, mammary glands)	Epidermis (ectoderm)	Dermis (mesoderm)
Limb	Epidermis (ectoderm)	Mesenchyme (mesoderm)
Gut organs (liver, pancreas, salivary glands)	Epithelium (endoderm)	Mesenchyme (mesoderm)
Foregut and respiratory associated organs (lungs, thymus, thyroid)	Epithelium (endoderm)	Mesenchyme (mesoderm)
Kidney	Ureteric bud epithelium (mesoderm)	Mesenchyme (mesoderm)
Tooth	Jaw epithelium (ectoderm)	Mesenchyme (neural crest)

GENETIC SPECIFICITY OF INDUCTION The second property of epithelial-mesenchymal interactions is the genetic specificity of induction. Whereas the mesenchyme may instruct the epithelium as to what sets of genes to activate, the responding epithelium can comply with these instructions only so far as its genome permits. This property was discovered through experiments involving the transplantation of tissues from one species to another.

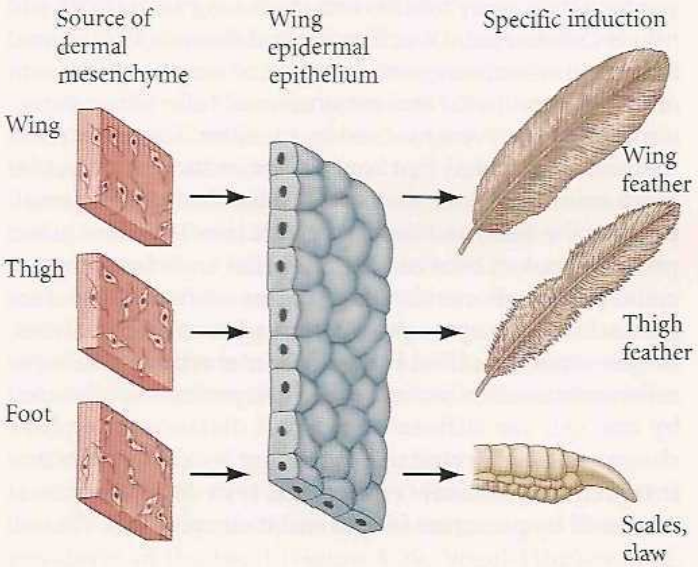


FIGURE 3.17 Regional specificity of induction in the chick. When cells from different regions of the dermis (mesenchyme) are recombined with the epidermis (epithelium), the type of cutaneous structure made by the epidermal epithelium is determined by the original source of the mesenchyme. (After Saunders 1980.)

In one of the most dramatic examples of interspecific induction, Hans Spemann and Oscar Schotté (1932) transplanted flank ectoderm from an early *frog* gastrula to the region of a *newt* gastrula destined to become parts of the mouth. Similarly, they placed presumptive flank ectodermal tissue from a *newt* gastrula into the presumptive oral regions of *frog* embryos. The structures of the mouth region differ greatly between salamander and frog larvae. The salamander larva has club-shaped balancers beneath its mouth, whereas the frog tadpole produces mucus-secreting glands and suckers (Figure 3.18). The frog tadpole also has a horny jaw without teeth, whereas the salamander has a set of calcareous teeth in its jaw. The larvae resulting from the transplants were chimeras. The salamander larvae had froglike mouths, and the frog tadpoles had salamander teeth and balancers. In other words, the mesenchymal cells instructed the ectoderm to make a mouth, but the ectoderm responded by making the only kind of mouth it "knew" how to make, no matter how inappropriate.*

Thus the instructions sent by the mesenchymal tissue can cross species barriers. Salamanders respond to frog inducers, and chick tissue responds to mammalian inducers. The response of the epithelium, however, is species-specific. So, whereas organ-type specificity (e.g., feather or claw) is usually controlled by the mesenchyme, species specificity is usually controlled by the responding epithelium. As we will see in Chapter 19, major evolutionary changes in the phenotype can be brought about by changing the response to a particular inducer.

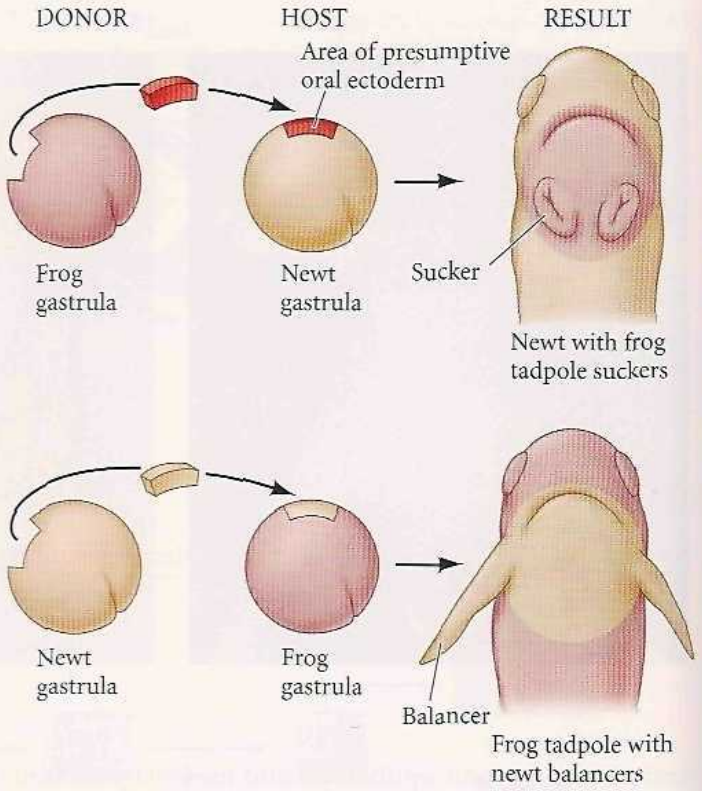


FIGURE 3.18 Genetic specificity of induction in amphibians. Reciprocal transplantation between the presumptive oral ectoderm regions of salamander and frog gastrulae leads to newts with tadpole suckers and tadpoles with newt balancers. (After Hamburgh 1970.)

Paracrine Factors: The Inducer Molecules

How are the signals between inducer and responder transmitted? While studying the mechanisms of induction that produce the kidney tubules and teeth, Grobstein (1956) and others (Saxén et al. 1976; Slavkin and Bringas 1976) found that some inductive events could occur despite a filter separating the epithelial and mesenchymal cells. Other inductions, however, were blocked by the filter. The researchers therefore concluded that some of the inductive molecules were soluble factors that could pass through the small pores of the filter, and that other inductive events required physical contact between the epithelial and mesenchymal cells. When cell membrane proteins on one cell surface interact with receptor proteins on adjacent cell surfaces, these events are called **juxtacrine interactions** (since the cell membranes are *juxtaposed*). When proteins synthesized by one cell can diffuse over small distances to induce changes in neighboring cells, the event is called a **paracrine interaction**. A we saw earlier, this type of interaction is mediated by paracrine factors and their receptors. We will consider paracrine interactions first, returning to juxtacrine interactions later in the chapter.

Whereas **endocrine factors*** (hormones) travel through the blood to exert their effects, paracrine factors are secreted into the immediate spaces around the cell producing them. These proteins are the "inducing factors" of the classic experimental embryologists. There is considerable debate as to the distances at which paracrine factors can operate. The proteins Nodal and activin, for instance, can diffuse over many cell diameters and induce different sets of genes at different concentrations (Gurdon et al. 1994, 1995). The Wnt, Vg1, and BMP4 proteins, however, probably work only on their adjacent neighbors (Jones et al. 1996; Reilly and Melton 1996). These factors may induce the expression of other short-range factors from these neighbors, and a cascade of paracrine inductions can be initiated.

*Spemann is reported to have put it this way: "The ectoderm says to the inducer, 'you tell me to make a mouth; all right, I'll do so, but I can't make your kind of mouth; I can make my own and I'll do that.'" (Quoted in Harrison 1933.)

*Some endocrine factors are active in development; these include the hormones estrogen, testosterone, and thyroxine. They affect many tissues simultaneously and often coordinate development throughout the body (as in metamorphosis or the morphogenesis of sexual phenotypes). These hormones work directly, binding to a dormant transcription factor ("receptor"), thereby activating the transcription factor and allowing it to enter the nucleus and bind to DNA.

In addition to endocrine, paracrine, and juxtacrine interactions, there are also **autocrine** interactions. Autocrine interactions occur when the same cells that secrete paracrine factors also respond to them. In other words, the cell synthesizes a molecule for which it has its own receptor. Although autocrine regulation is not common, it is seen in placental cytotrophoblast cells; these cells synthesize and secrete platelet-derived growth factor, whose receptor is on the cytotrophoblast cell membrane (Goustin et al. 1985). The result is the explosive proliferation of that tissue.

Signal transduction cascades: The response to inducers

The induction of numerous organs is actually effected by a relatively small set of paracrine factors. The embryo inherits a rather compact genetic "tool kit" and uses many of the same proteins to construct the heart, kidneys, teeth, eyes, and other organs. Moreover, the same proteins are used throughout the animal kingdom: the factors active in creating the *Drosophila* eye or heart are very similar to those used in generating mammalian organs. Many of the paracrine factors can be grouped into one of four major families on the basis of their structure:

1. The fibroblast growth factor (FGF) family
2. The Hedgehog family
3. The Wnt family
4. The TGF-β superfamily, encompassing the TGF-β family, the activin family, the bone morphogenetic proteins (BMPs), the Nodal proteins, the Vg1 family, and several other related proteins.

Paracrine factors function by binding to a receptor that initiates a series of enzymatic reactions within the cell. These enzymatic reactions have as their end point either the regulation of transcription factors (such that different genes are expressed in the cells reacting to these paracrine factors) or the regulation of the cytoskeleton (such that the cells responding to the paracrine factors alter their shape or are permitted to migrate). These pathways of responses to the paracrine factor often have several end points and are called **signal transduction cascades**.

The major signal transduction pathways all appear to be variations on a common and rather elegant theme, exemplified in Figure 3.19. Each receptor spans the cell membrane and has an extracellular region, a transmembrane region, and a cytoplasmic region. When a ligand (here, the paracrine factor) binds to its receptor's extracellular domain, that ligand induces a conformational change in the receptor's structure. This shape change is transmitted through the membrane and alters the shape of the receptor's cytoplasmic domains. This conformational change in cytoplasmic domains gives those domains enzymatic activity—usually a kinase activity that can use ATP to phosphorylate specific tyrosine residues of particular proteins. Thus, this type of receptor is often called a **receptor tyrosine kinase** (**RTK**). The active receptor can now

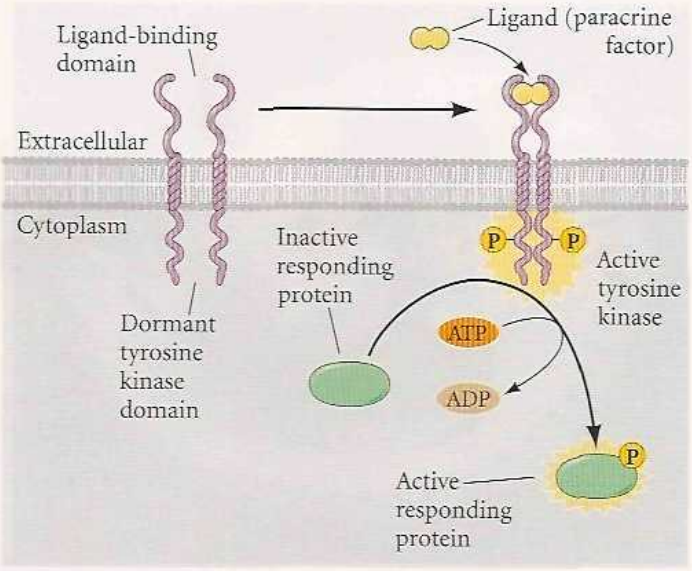


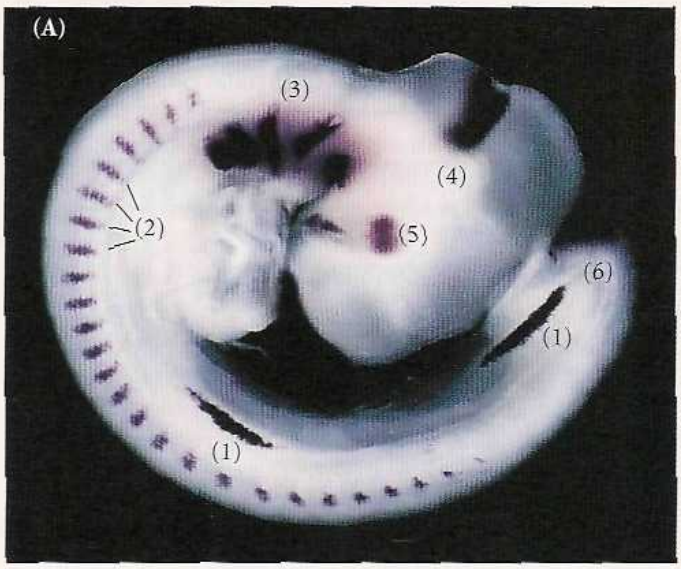
FIGURE 3.19 Structure and function of a receptor tyrosine kinase. The binding of a paracrine factor (such as Fgf8) by the extracellular portion of the receptor protein activates the dormant tyrosine kinase, whose enzyme activity phosphorylates specific tyrosine residues of certain proteins.

catalyze reactions that phosphorylate other proteins, and this phosphorylation in turn activates their latent activities. Eventually, the cascade of phosphorylation activates a dormant transcription factor or a set of cytoskeletal proteins.

Fibroblast growth factors and the RTK pathway

The **fibroblast growth factor** (**FGF**) family of paracrine factors comprises nearly two dozen structurally related members, and the FGF genes can generate hundreds of protein isoforms by varying their RNA splicing or initiation codons in different tissues (Lappi 1995). Fgf1 protein is also known as acidic FGF and appears to be important during regeneration (Yang et al. 2005); Fgf2 is sometimes called basic FGF and is very important in blood vessel formation; and Fgf7 sometimes goes by the name of keratinocyte growth factor and is critical in skin development. Although FGFs can often substitute for one another, the expression patterns of the FGFs and their receptors give them separate functions. In *Drosophila*, Breathless is an FGF protein.

One member of this family, Fgf8, is especially important during limb development and lens induction. Fgf8 is usually made by the optic vesicle that contacts the outer ectoderm of the head (Figure 3.20; Vogel-Höpker et al. 2000). After contact with the outer ectoderm occurs, *fgf8* gene expression becomes concentrated in the region of the presumptive neural retina—the tissue directly apposed to the presumptive lens. Moreover, if Fgf8-containing beads are placed adjacent to head ectoderm, this ectopic Fgf8 will



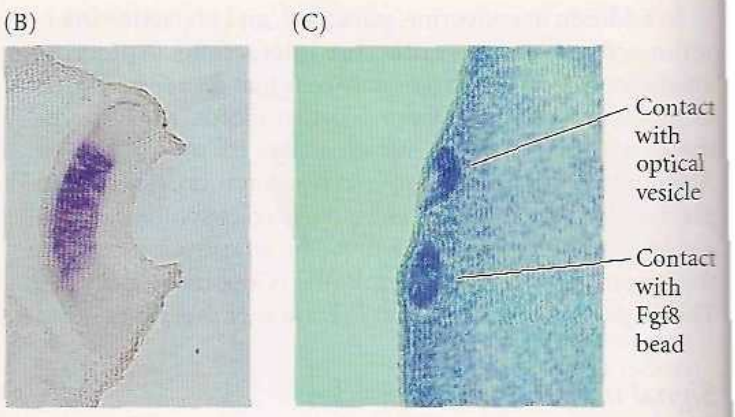


FIGURE 3.20 Fgf8 in the developing chick. (A) *Fgf8* gene expression pattern in the 3-day chick embryo, shown by in situ hybridization. Fgf8 protein (dark areas) is seen in the distalmost limb bud ectoderm (1); in the somitic mesoderm (the segmented blocks of cells along the anterior-posterior axis; 2); in the branchial arches of the neck (3); at the boundary between the midbrain and hindbrain (4); in the developing eye (5); and in the tail (6). (B,C) Fgf8 function in the developing eye. (B) In situ hybridization of *fgf8* in the optic vesicle. The *fgf8* mRNA (purple) is localized to the presumptive neural retina of the optic cup and is in direct contact with the outer ectoderm cells that will become the lens. (C) Ectopic expression of L-Maf in competent ectoderm can be induced by the optic vesicle (above) and by an Fgf8-containing bead (below). (A courtesy of E. Laufer, C.-Y. Yeo, and C. Tabin; B,C courtesy of A. Vogel-Höpker.)

induce this ectoderm to produce ectopic lenses and to express the lens-associated transcription factor L-Maf (see Figure 3.20B). FGFs often work by activating a set of receptor tyrosine kinases called the **fibroblast growth factor receptors (FGFRs)**. The Branchless protein is an FGF receptor in *Drosophila*.

When an FGFR binds an FGF (and only when it binds an FGF), the dormant kinase is activated and phosphorylates certain proteins (including other FGFRs) within the responding cell. These proteins, once activated, can perform new functions. The **RTK pathway** was one of the first signal transduction pathways to unite various areas of developmental biology (Figure 3.21). Researchers studying *Drosophila* eyes, nematode vulvae, and human cancers found that they were all studying the same genes. The pathway begins at the cell surface, where an RTK binds its specific ligand. Ligands that bind to RTKs include the fibroblast growth factors, epidermal growth factors, platelet-derived growth factors, and stem cell factor. Each RTK can bind only one or a small set of these ligands. The RTK spans the cell membrane, and when it binds its ligand, it undergoes a conformational change that enables it to dimerize with another RTK. This conformational change activates the latent kinase activity of each RTK, and these receptors phosphorylate each other on particular tyrosine residues (see Figure 3.19). Thus, the binding of the ligand to the receptor causes the autophosphorylation of the cytoplasmic domain of the receptor.

The phosphorylated tyrosine on the receptor is then recognized by an adaptor protein. The adaptor protein serves as a bridge that links the phosphorylated RTK to a powerful intracellular signaling system. While binding to the phosphorylated RTK through one of its cytoplasmic domains, the adaptor protein also activates a **G protein**, such as **Ras**. Normally, the G protein is in an inactive, GDP-bound state. The activated receptor stimulates the adaptor protein to activate the **guanine nucleotide releasing factor (GNRP)**. This protein exchanges a phosphate from a GTP to transform the bound GDP into GTP. The GTP-bound G protein is an active form that transmits the signal to the next molecule. After the signal is delivered, the GTP on the G protein is hydrolyzed back into GDP. This catalysis is greatly stimulated by the complexing of the Ras protein with the **GTPase-activating protein (GAP)**. In this way, the G protein is returned to its inactive state, where it can await further signaling. Without the GAP protein, Ras protein cannot catalyze GTP well, and so remains in its active configuration (Cales et al. 1988; McCormick 1989). Mutations in the *RAS* gene account for a large proportion of cancerous human tumors (Shih and Weinberg 1982), and the mutations of *RAS* that make it oncogenic all inhibit the binding of the GAP protein.

The active Ras G protein associates with a kinase called Raf. The G protein recruits the inactive Raf protein to the cell membrane, where it becomes active (Leevers et al. 1994; Stokoe et al. 1994). The Raf protein is a kinase that activates the MEK protein by phosphorylating it. MEK is itself a kinase, which activates the ERK protein by phosphorylation. In turn, ERK is a kinase that enters the nucleus and phosphorylates certain transcription factors.

The RTK pathway is critical in numerous developmental processes. Moreover, it can be activated by paracrine factors other than those of the FGF family. (And FGF fam-

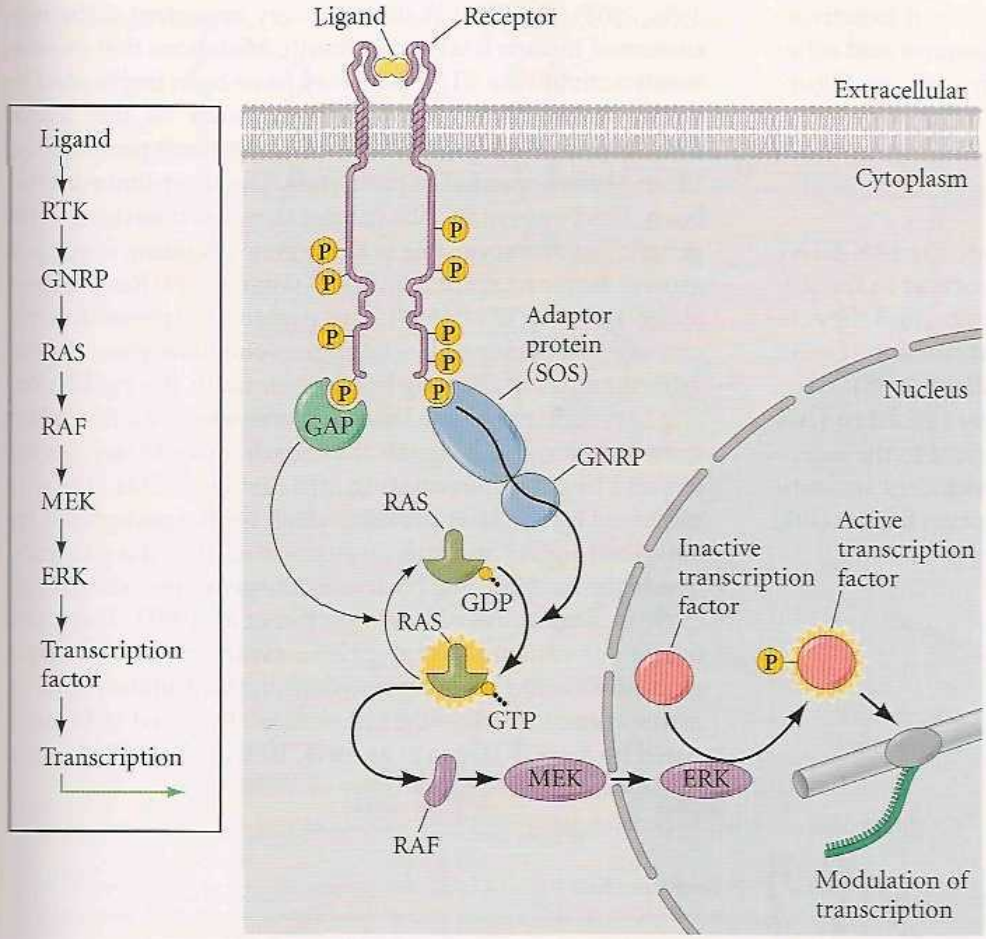


FIGURE 3.21 The widely used RTK signal transduction pathway. The receptor tyrosine kinase is dimerized by the ligand, which causes the autophosphorylation of the receptor. The adaptor protein recognizes the phosphorylated tyrosines on the RTK and activates an intermediate protein, GNRP, which activates the Ras G protein by allowing the phosphorylation of the GDP-bound Ras. At the same time, the GAP protein stimulates the hydrolysis of this phosphate bond, returning Ras to its inactive state. The active Ras activates the Raf protein kinase C (PKC), which in turn phosphorylates a series of kinases. Eventually, the activated kinase ERK alters gene expression in the nucleus of the responding cell by phosphorylating certain transcription factors (which can then enter the nucleus to change the types of genes transcribed) and certain translation factors (which alter the level of protein synthesis). In many cases, this pathway is reinforced by the release of calcium ions. A simplified version of the pathway is shown on the left.

ily proteins can activate other signaling pathways, depending on the receptor.) In the migrating neural crest cells of humans and mice, the pathway is important in activating the microphthalmia transcription factor (MITF) to produce the pigment cells (Figure 3.22). MITF, whose mechanism of action was described in Chapter 2, is transcribed in the pigment-forming melanoblast cells that migrate from the neural crest into the skin and in the melanin-forming cells of the pigmented retina. But we have not yet discussed what proteins signal this transcription factor to become active. The clue lies in two mouse mutants whose phenotypes resemble those of mice homozygous for *microphthalmia* mutations. Like *Mitf* mutant mice, homozygous *White* mice and homozygous *Steel* mice are white because their pigment cells have failed to migrate. Could it be that all three genes (*Mitf*, *Steel*, and *White*) are on the same developmental pathway?

In 1990, several laboratories demonstrated that the *Steel* gene encodes a paracrine protein called **stem cell factor** (see Witte 1990). Stem cell factor binds to and activates the **Kit** receptor tyrosine kinase encoded by the *White* gene (Spritz et al. 1992; Wu et al. 2000; also see Chapter 1). The binding of stem cell factor to the Kit RTK dimerizes the Kit protein, causing it to become phosphorylated. The phosphorylated Kit activates the pathway whereby phosphorylated ERK is able to phosphorylate the MITF transcription factor (Hsu et al. 1997; Hemesath et al. 1998). Only the phosphorylated form of MITF is able to bind the p300/CBP

histone acetyltransferase protein that enables it to activate transcription of the genes encoding tyrosinase and other proteins of the melanin-formation pathway (see Figure 3.22; Price et al. 1998).

The JAK-STAT pathway

Fibroblast growth factors can also activate the JAK-STAT cascade. This pathway is extremely important in the differentiation of blood cells, the growth of limbs, and the activation of the casein gene during milk production (Figure 3.23; Briscoe et al. 1994; Groner and Gouilleux 1995). Here, the ligand is bound by receptors that are linked to **JAK** (*Ja*nus *k*inase) proteins. The binding of ligand to the receptor phosphorylates the **STAT** (signal *t*ransducers and *a*ctivators of *t*ranscription) family of transcription factors (Ihle 1996, 2001). The STAT pathway is very important in the regulation of human fetal bone growth. Mutations that prematurely activate the STAT pathway have been implicated in some severe forms of dwarfism, such as the lethal **thanatophoric dysplasia**, wherein the growth plates of the rib and limb bones fail to proliferate. The short-limbed newborn dies because its ribs cannot support breathing. The genetic lesion responsible is in the gene encoding fibroblast growth factor receptor 3,or FgfR3 (Figure 3.24; Rousseau et al. 1994; Shiang et al. 1994). This protein is expressed in the cartilage precursor cells—known as **chondrocytes**—in the growth plates of the long bones. Normally, the FgfR3 protein (a receptor tyrosine kinase) is activated by a fibroblast growth factor, and it signals the chondrocytes to stop dividing and begin differentiating into cartilage. This signal is mediated by the Stat1 protein, which is phosphorylated by activated FgfR3 and then translocated into the nucleus. Inside the nucleus, Stat1 activates the genes encoding a cell cycle inhibitor, the p21 protein (Su et al. 1997). Thus, the mutations causing thanatophoric dwarfism result from a gain-of-function phenotype, wherein the mutant FgfR3 is active constitutively—that is, without the need to be activated by an FGF (Deng et al. 1996; Webster and Donoghue

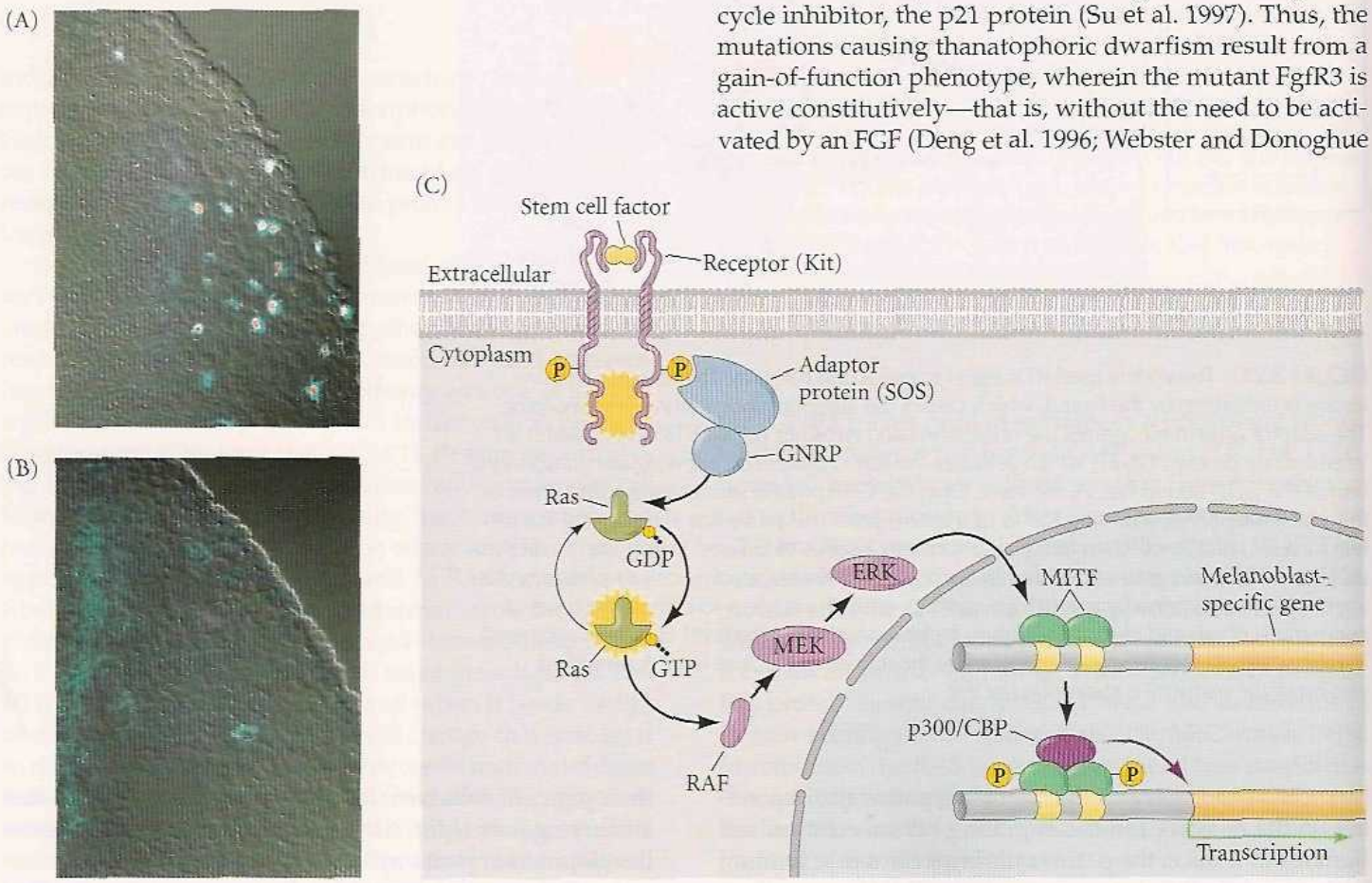


FIGURE 3.22 Activation of MITF transcription factor through the binding of stem cell factor by the Kit RTK protein. The information received at the cell membrane is sent to the nucleus by the RTK signal transduction pathway. (A,B) Demonstration that Kit protein and MITF are present in the same cells. Antibodies to these proteins stain the Kit protein (red) and MITF (green). The overlap is yellow or yellow-green. Both proteins are present in the migrating melanocyte precursor cells (melanoblasts). (A) Migrating melanoblasts can be seen in a wild-type mouse embryo at day 10.5. (B) No melanoblasts are visible in a *microphthalmia* mutant embryo of the same age. The lack of melanoblasts in the mutant is due to the relative absence of MITF. (C) Signal transduction pathway leading from the cell membrane to the nucleus. When the receptor domain of the Kit RTK protein binds the stem cell factor, Kit dimerizes and becomes phosphorylated. This phosphorylation is used to activate the Ras G protein, which activates the chain of kinases that will phosphorylate the MITF protein. Once phosphorylated, MITF can bind the cofactor p300/CBP, acetylate the nucleosome histones, and initiate transcription of the genes needed for melanocyte development. (A,B from Nakayama et al. 1998, courtesy of H. Arnheiter; C after Price et al. 1998.)

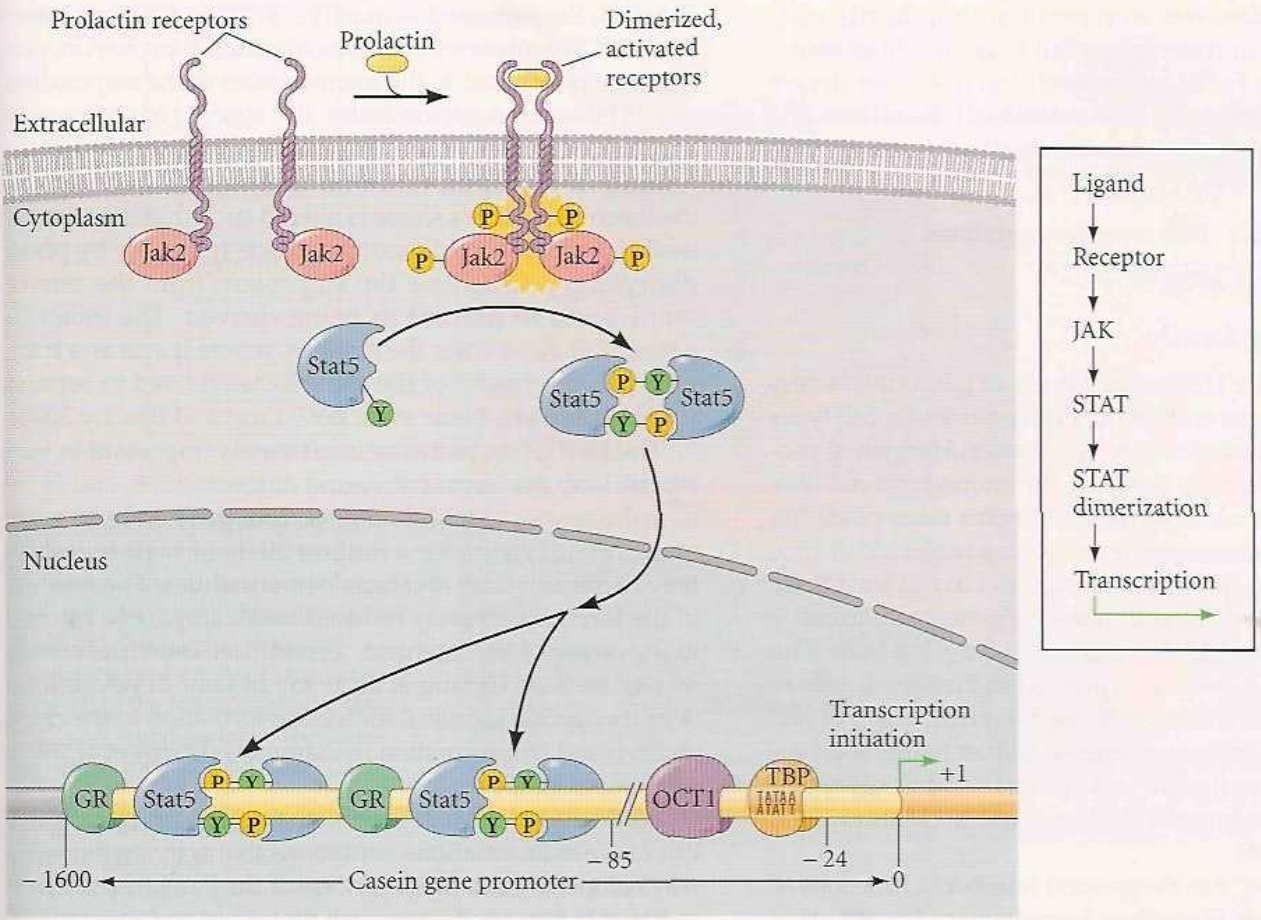


FIGURE 3.23 A STAT pathway: the casein gene activation pathway activated by prolactin. The casein gene is activated during the last (lactogenic) phase of mammary gland development, and its signal is the secretion of the hormone prolactin from the anterior pituitary gland. Prolactin causes the dimerization of prolactin receptors in the mammary duct epithelial cells. A particular JAK protein (Jak2) is "hitched" to the cytoplasmic domain of these receptors. When the receptors bind prolactin and dimerize, the JAK proteins phosphorylate each other and the dimerized receptors, activating the dormant kinase activity of the receptors. The activated receptors add a phosphate group to a tyrosine residue (Y) of a particular STAT protein—in this case, Stat5. This allows Stat5 to dimerize, be translocated into the nucleus, and bind to particular regions of DNA. In combination with other transcription factors (which presumably have been waiting for its arrival), the Stat5 protein activates transcription of the casein gene. GR is the glucocorticoid receptor, OCT1 is a general transcription factor, and TBP is the TATA-binding protein (see Chapter 2) responsible for binding RNA polymerase. A simplified diagram is shown to the right. (For details, see Groner and Gouilleux 1995.)

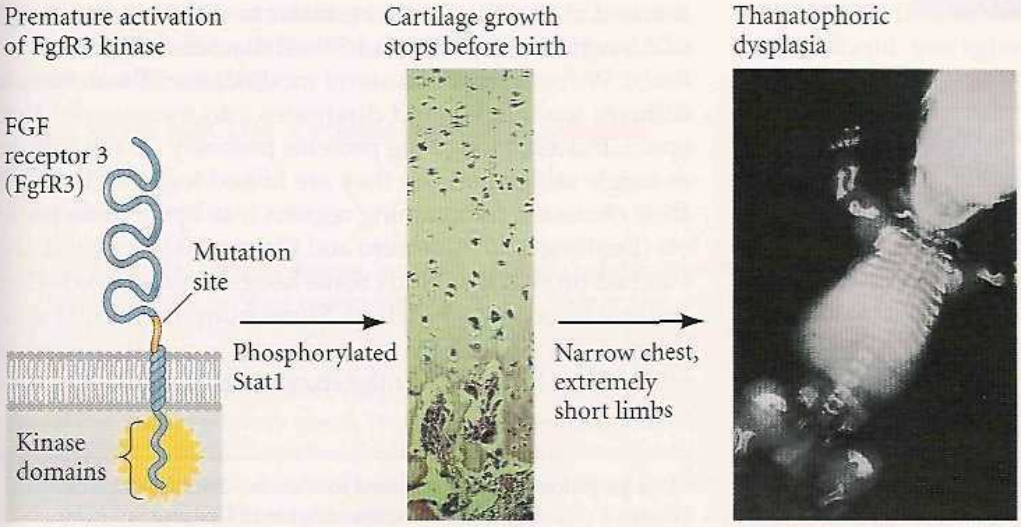


FIGURE 3.24 A mutation in the gene for FgfR3 causes the premature constitutive activation of the STAT pathway and the production of phosphorylated Stat1 protein. This transcription factor activates genes that cause the premature termination of chondrocyte cell division. The result is thanatophoric dysplasia, a condition of failed bone growth that results in the death of the newborn infant because the thoracic cage cannot expand to allow breathing. (After Gilbert-Barness and Opitz 1996.)

1996). The chondrocytes stop proliferating shortly after they are formed, and the bones fail to grow. Other mutations that activate FgfR3 prematurely but to a lesser degree produce **achondroplasic (short-limbed) dwarfism**, the most prevalent of the human dominant syndromes (Legeai-Mallet et al. 2004).

See WEBSITE 3.1 FGF receptor mutations

The Hedgehog family

The proteins of the Hedgehog family of paracrine factors are often used by the embryo to induce particular cell types and to create boundaries between tissues. Hedgehog proteins are processed such that only the amino-terminal two-thirds of the molecule is secreted; once this takes place, the protein must become complexed with a molecule of cholesterol in order to function. Vertebrates have at least three homologues of the *Drosophila hedgehog* gene: *sonic hedgehog* (*shh*), *desert hedgehog* (*dhh*), and *indian hedgehog* (*ihh*). The Desert hedgehog protein is expressed in the Sertoli cells of the testes, and mice homozygous for a null allele of *dhh* exhibit defective spermatogenesis. Indian hedgehog protein is expressed in the gut and cartilage and is important in postnatal bone growth (Bitgood and McMahon 1995; Bitgood et al. 1996).

Sonic hedgehog* has the greatest number of functions of the three vertebrate Hedgehog homologues. Among other important functions, this paracrine factor is responsible for assuring that motor neurons come only from the ventral portion of the neural tube (see Chapter 10), that a portion of each somite forms the vertebrae (see Chapter 12), that the feathers of the chick form in their proper places (see Figure 3.16), and that our pinkies are always our most posterior digits (see Chapter 14). Sonic hedgehog often works with other paracrine factors, such as Wnt and FGF proteins.

THE HEDGEHOG PATHWAY Proteins of the Hedgehog family function by binding to a receptor called Patched. The Patched protein, however, is not a signal transducer. Rather, it is *bound to* a signal transducer, the Smoothened protein. The Patched protein prevents Smoothened from functioning. In the absence of Hedgehog binding to Patched, Smoothened is inactive, and the Cubitus interruptus (Ci) protein (or the homologous Gli protein in vertebrates) is tethered to the microtubules of the responding cell. While on the microtubules, it is cleaved in such a way that a portion of it enters the nucleus and acts as a transcriptional repressor. When Hedgehog binds to Patched, the Patched protein's shape is altered such that it no longer inhibits Smoothened. Smoothened acts (probably by phosphorylation) to release the Ci protein from the microtubules and to prevent its being cleaved. The intact Ci protein can now enter the nucleus, where it acts as a transcriptional *activator* of the same genes it used to repress (Figure 3.25; Aza-Blanc et al. 1997; Lum and Beachy 2004).

The Hedgehog pathway is extremely important in vertebrate limb development, neural differentiation, and facial morphogenesis (McMahon et al. 2003). When mice were made homozygous for a mutant allele of *sonic hedgehog*, they had major limb and facial abnormalities. The midline of the face was severely reduced and a single eye formed in the center of the forehead, a condition known as *cyclopia*† (Figure 3.26; Chiang et al. 1996). In later development, Sonic hedgehog is critical for feather formation in the chick embryo and hair formation in mammals (Harris et al. 2002; Michino et al. 2003).

While mutations that inactivate the Hedgehog pathway can cause malformations, mutations that activate the pathway ectopically can cause cancers. If the Patched protein is mutated in somatic tissues such that it can no longer inhibit Smoothened, it can cause tumors of the basal cell layer of the epidermis (basal cell carcinomas). Heritable mutations of the *patched* gene cause basal cell nevus syndrome, a rare autosomal dominant condition characterized by both developmental anomalies (fused fingers; rib and facial abnormalities) and multiple malignant tumors such as basal cell carcinoma (Hahn et al. 1996; Johnson et al. 1996).

One remarkable feature of the Hedgehog signal transduction pathway is the importance of cholesterol. First, cholesterol is critical for the catalytic cleavage of Sonic hedgehog protein. Only the amino-terminal portion of the protein is functional and secreted. The cholesterol also binds to the active N-terminus of the Sonic hedgehog protein and allows this paracrine factor to diffuse over a range of a few hundred μm (about 30 cell diameters in the mouse limb). Without this cholesterol modification, the molecule diffuses too quickly and dissipates into the surrounding space. Indeed, Hedgehog proteins probably do not diffuse as single molecules, but they are linked together through their cholesterol-containing regions into lipoprotein packets (Breitling 2007; Guerrero and Chiang 2007). Second, the Patched protein that binds Sonic hedgehog also needs cholesterol in order to function. Some human cyclopia syndromes are caused by mutations in genes that encode either Sonic hedgehog or the enzymes that synthesize cho-

*Yes, it is named after the Sega Genesis character. The original *hedgehog* gene was found in *Drosophila*, in which genes are named after their mutant phenotypes—the loss-of-function *hedgehog* mutation causes the fly embryo to be covered with pointy denticles on its cuticle, so it looks like a hedgehog. The vertebrate Hedgehog genes were discovered by searching vertebrate gene libraries (chick, rat, zebrafish) with probes that would find sequences similar to that of the fruit fly *hedgehog* gene. Riddle and colleagues (1993) discovered three genes homologous to *Drosophila hedgehog*. Two were named after existing species of hedgehog; the third was named after the animated character. Two other Hedgehog genes, found only in fish, are named *echidna hedgehog* (after the spiny Australian marsupial mammal) and *Tiggywinkle hedgehog* (after Beatrix Potter's fictional hedgehog).

†This pathology, which is named for the one-eyed Cyclops of Homer's *Odyssey*, will be discussed again in Chapter 9.

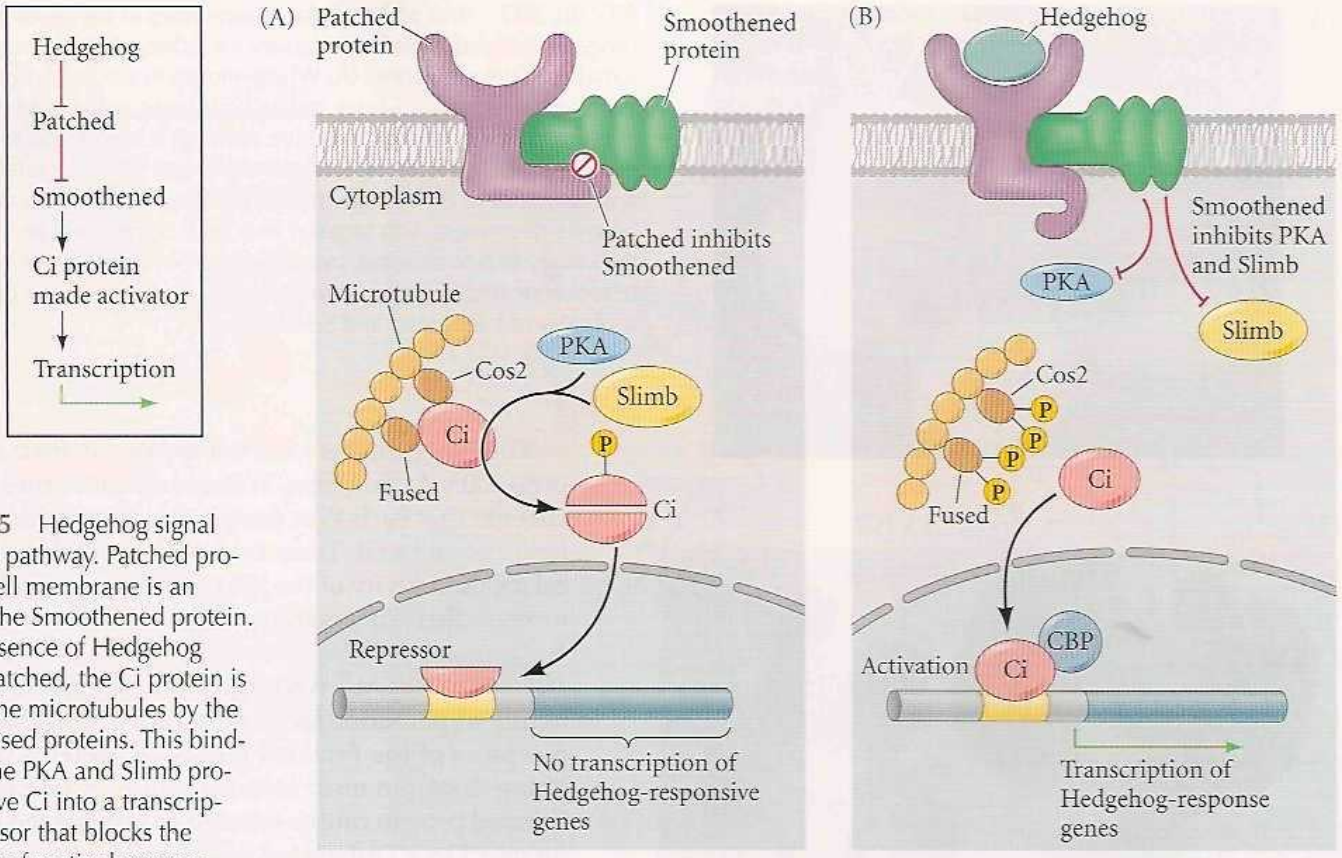


FIGURE 3.25 Hedgehog signal transduction pathway. Patched protein in the cell membrane is an inhibitor of the Smoothened protein. (A) In the absence of Hedgehog binding to Patched, the Ci protein is tethered to the microtubules by the Cos2 and Fused proteins. This binding allows the PKA and Slimb proteins to cleave Ci into a transcriptional repressor that blocks the transcription of particular genes. (B) When Hedgehog binds to Patched, its conformation changes, releasing the inhibition of the Smoothened protein. Smoothened then releases Ci from the microtubules (probably by adding more phosphates to the Cos2 and Fused proteins) and inactivates the cleavage proteins PKA and Slimb. The Ci protein enters the nucleus, binds a CBP protein, and acts as a transcriptional activator of particular genes. (After Johnson and Scott 1998.)

lesterol (Kelley et al. 1996; Roessler et al. 1996). Moreover, certain chemicals that induce cyclopia do so by interfering with the cholesterol biosynthetic enzymes (Beachy et al. 1997; Cooper et al. 1998). Two teratogens known to cause cyclopia in vertebrates are jervine and cyclopamine. Both substances are found in the plant *Veratrum californicum*, and both block the synthesis of cholesterol (see Figure 3.26; Keeler and Binns 1968).

See VADE MECUM Cyclopia induced in zebrafish

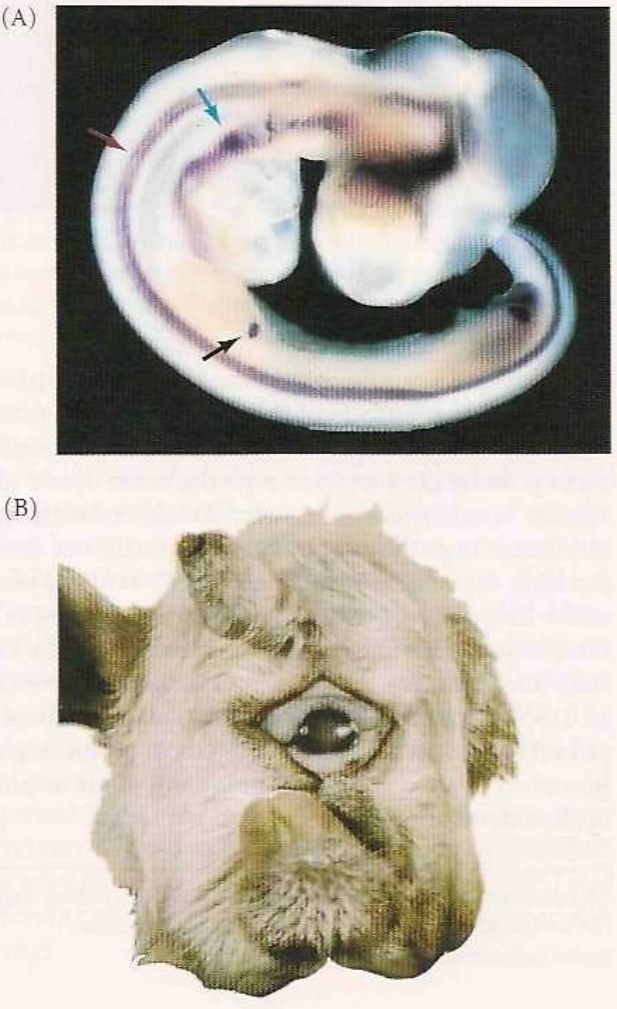
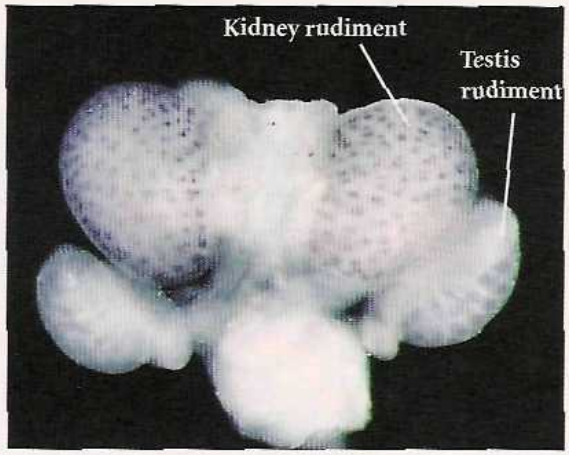


FIGURE 3.26 (A) Sonic hedgehog is shown by in situ hybridization to be expressed in the nervous system (red arrow), gut (blue arrow), and limb bud (black arrow) of a 3-day chick embryo. (B) Head of a cyclopic lamb born of a ewe that ate *Veratrum californicum* early in pregnancy. The cerebral hemispheres fused, resulting in the formation of a single, central eye and no pituitary gland. The jervine alkaloid made by this plant inhibits cholesterol synthesis, which is needed for Hedgehog production and reception. (A courtesy of C. Tabin; B courtesy of L. James and USDA Poisonous Plant Laboratory.)



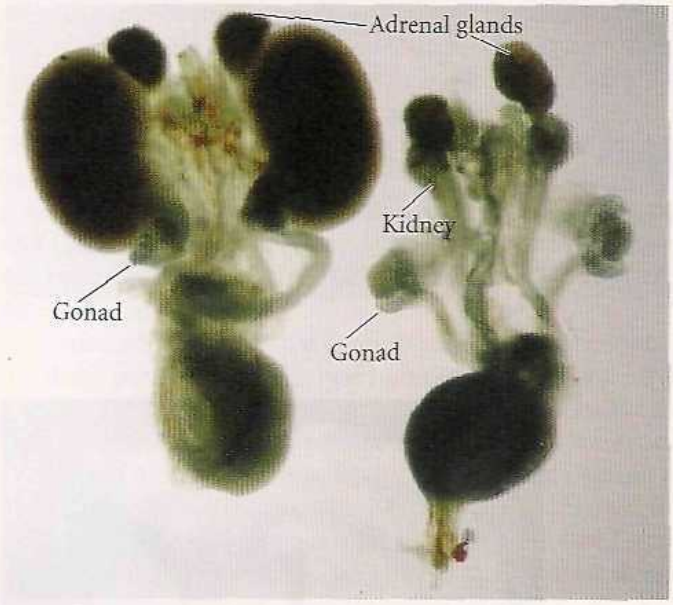


FIGURE 3.27 Wnt proteins play several roles in the development of the urogenital organs. Wnt4 is necessary for kidney development and for female sex determination. (A) Whole-mount in situ hybridization of Wnt4 expression in a 14-day mouse embryonic male urogenital rudiment. Expression (dark purple-blue staining) is seen in the mesenchyme that condenses to form the kidney's nephrons. (B) Urogenital rudiment of a wild-type newborn female mouse. (C) Urogenital rudiment of a newborn female mouse with targeted knockout of the *Wnt4* gene shows that the kidney fails to develop. In addition, the ovary starts synthesizing testosterone and becomes surrounded by a modified male duct system. (Courtesy of J. Perasaari and S. Vainio.)

The Wnt family

The Wnts are a family of cysteine-rich glycoproteins. There are at least 15 members of this gene family in vertebrates.* Their name is a fusion of the name of the *Drosophila* segment polarity gene *wingless* with the name of one of its vertebrate homologues, *integrated*. While Sonic hedgehog is important in patterning the ventral portion of the somites (causing the cells to become cartilage), Wnt1 appears to be active in inducing the dorsal cells of the somites to become muscle and is involved in the specification of the midbrain cells (see Chapter 11; McMahon and Bradley 1990; Stern et al. 1995). Wnt proteins also are critical in establishing the polarity of insect and vertebrate limbs, promoting the proliferation of stem cells, and in several steps of urogenital system development (Figure 3.27).

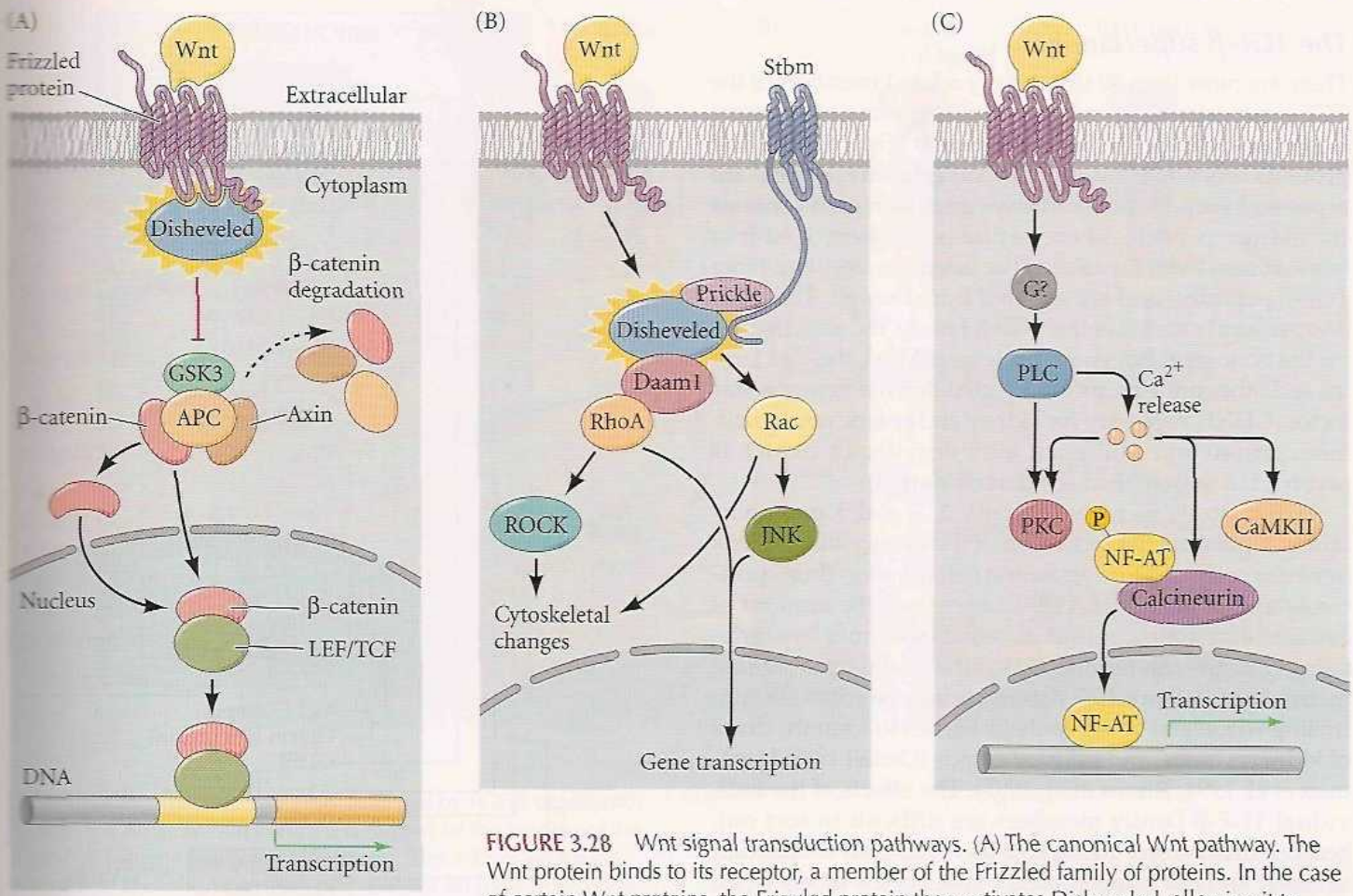
*A summary of all the Wnt proteins and Wnt signaling components can be found at http://www.stanford.edu/~rnusse/wntwindow.htm

The Wnt proteins were not isolated in their active form until 2003. At that time, Willert and colleagues (2003) discovered that each Wnt protein has a lipid molecule covalently bound to it. These hydrophobic molecules are critical for the activity of the Wnt proteins and probably act to increase their concentration in the cell membrane.

THE "CANONICAL" WNT PATHWAY Members of the Wnt family of paracrine factors interact with transmembrane receptors of the Frizzled family of proteins (Logan and Nusse 2004). In most instances, the binding of Wnt by a Frizzled protein causes Frizzled to activate the Disheveled protein. Once Disheveled is activated, it inhibits the activity of the glycogen synthase kinase 3 (GSK3) enzyme. GSK3, if it were active, would prevent the dissociation of the β-catenin protein from the APC protein, which targets β-catenin for degradation. However, when the Wnt signal is present and GSK3 is inhibited, β-catenin can dissociate from the APC protein and enter the nucleus. Once inside the nucleus, β-catenin binds to a Lef/Tcf transcription factor that is already on the DNA, repressing the genes it has bound. The binding of β-catenin to the Lef/Tcf protein converts the repressor into a transcriptional activator, thereby activating the Wnt-responsive genes (Figure 3.28A; Behrens et al. 1996; Cadigan and Nusse 1997).

This model is undoubtedly an oversimplification, because different cells use this pathway in different ways (see McEwen and Peifer 2001). Moreover, its components can have more than one function in the cell. In addition to being part of the Wnt signal transduction cascade, GSK3 is also an enzyme that regulates glycogen metabolism. The β-catenin protein was recognized as being part of the cell adhesion complex on the cell surface before it was also found to be a transcription factor. The APC protein also functions as a tumor suppressor. The transformation of normal adult colon epithelial cells into colon cancer is thought to occur when the *APC* gene is mutated and can no longer keep β-catenin out of the nucleus (Korinek et al. 1997; He et al. 1998). Once in the nucleus, β-catenin can bind with another transcription factor and activate genes for cell division.

One overriding principle is readily evident in both the Wnt pathway and the Hedgehog pathway: *activation is often accomplished by inhibiting an inhibitor*. Thus, in the Wnt



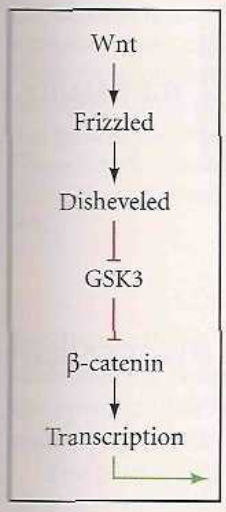


FIGURE 3.28 Wnt signal transduction pathways. (A) The canonical Wnt pathway. The Wnt protein binds to its receptor, a member of the Frizzled family of proteins. In the case of certain Wnt proteins, the Frizzled protein then activates Disheveled, allowing it to become an inhibitor of glycogen synthase kinase 3 (GSK3). GSK3, if it were active, would prevent the dissociation of β-catenin from the APC protein. So by inhibiting GSK3, the Wnt signal frees β-catenin to associate with an LEF or TCF protein and become an active transcription factor. (B) In a pathway that regulates cell morphology, division, and movement, certain Wnt proteins activate Frizzled in a way that causes Frizzled to activate the Disheveled protein, which has been tethered to the plasma membrane (through the Prickle protein). Here, Disheveled activates Rac and RhoA proteins, which coordinate the cytoskeleton and which can also regulate gene expression. (C) In a third pathway, certain Wnt proteins activate Frizzled receptors in a way that releases calcium ions and can cause Ca^{2+}-dependent gene expression.

pathway, the GSK3 protein is an inhibitor that is itself repressed by the Wnt signal.

THE "NONCANONICAL" WNT PATHWAYS The pathway described above is often called the "canonical" Wnt pathway because it was the first one to be discovered. However, in addition to sending signals to the nucleus, Wnt can also affect the actin and microtubular cytoskeleton. Here, Wnt activates alternative, "noncanonical," pathways. For instance, when Wnt activates Disheveled, the Disheveled protein can interact with a Rho GTPase. This GTPase can activate the kinases that phosphorylate cytoskeletal proteins and thereby alter cell shape, cell polarity (where the upper and lower portions of the cell differ), and motility (Figure 3.28B; Shulman et al. 1998; Winter et al. 2001). A third Wnt pathway diverges earlier than Disheveled. Here, the Frizzled receptor protein activates a phospholipase (PLC) that synthesizes a compound that releases calcium ions from the endoplasmic reticulum (Figure 3.28C). The released calcium can activate enzymes, transcription factors, and translation factors.

It is probable that the Frizzled proteins (of which there are many) can be used to couple different signal transduction cascades to the Wnt signal (see Chen et al. 2005) and that different cells have evolved to use Wnt factors in different ways.

The TGF-β superfamily

There are more than 30 structurally related members of the TGF-β superfamily,* and they regulate some of the most important interactions in development (Figure 3.29). The proteins encoded by TGF-β superfamily genes are processed such that the carboxy-terminal region contains the mature peptide. These peptides are dimerized into homodimers (with themselves) or heterodimers (with other TGF-β peptides) and are secreted from the cell. The TGF-β superfamily includes the TGF-β family, the activin family, the bone morphogenetic proteins (BMPs), the Vg1 family, and other proteins, including glial-derived neurotrophic factor (GDNF; necessary for kidney and enteric neuron differentiation) and Müllerian inhibitory factor (which is involved in mammalian sex determination).

TGF-β family members TGF-β1, 2, 3, and 5 are important in regulating the formation of the extracellular matrix between cells and for regulating cell division (both positively and negatively). TGF-β1 increases the amount of extracellular matrix epithelial cells make (both by stimulating collagen and fibronectin synthesis and by inhibiting matrix degradation). TGF-β proteins may be critical in controlling where and when epithelia branch to form the ducts of kidneys, lungs, and salivary glands (Daniel 1989; Hardman et al. 1994; Ritvos et al. 1995). The effects of the individual TGF-β family members are difficult to sort out, because members of the TGF-β family appear to function similarly and can compensate for losses of the others when expressed together.

The members of the BMP family can be distinguished from other members of the TGF-β superfamily by having seven (rather than nine) conserved cysteines in the mature polypeptide. Because they were originally discovered by their ability to induce bone formation, they were given the name **bone morphogenetic proteins**. But it turns out that bone formation is only one of their many functions; the BMPs are extremely multifunctional. †They have been found to regulate cell division, apoptosis (programmed cell death), cell migration, and differentiation (Hogan 1996). They include proteins such as BMP4 (which in some tissues causes bone formation, in other tissues causes cell death, and in other instances specifies the epidermis) and BMP7 (which is important in neural tube polarity, kidney development, and sperm formation). As it (rather oddly) turns out, however, BMP1 is not a member of the BMP family at all; it is a protease.

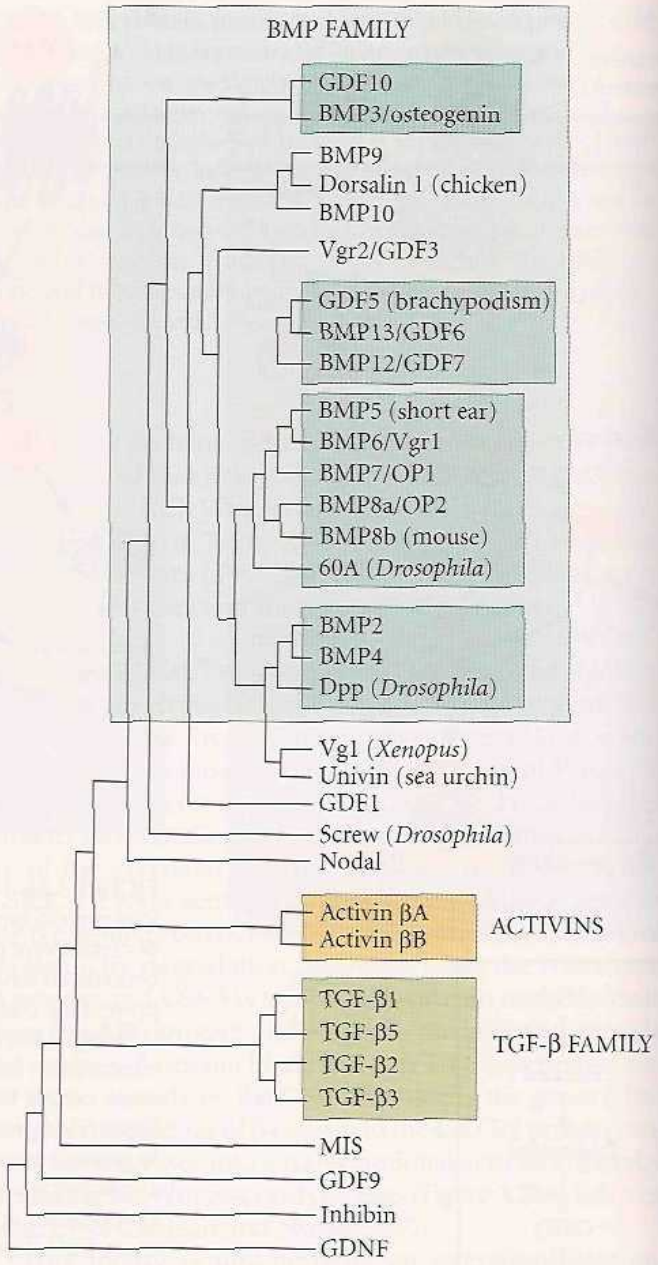


FIGURE 3.29 Relationships among members of the TGF-β superfamily. (After Hogan 1996.)

The *Drosophila* Decapentaplegic (Dpp) protein is homologous to vertebrate BMP4, and human BMP4 can replace Dpp and thus "rescue" *dpp*-deficient flies (Padgett et al. 1993). BMPs are thought to work by diffusion from the cells

*TGF stands for "*t*ransforming *g*rowth *f*actor." The designation "superfamily" is often given when each of the different classes of molecules constitutes a "family." The members of a superfamily all have similar structures but are not as close as the molecules within a family are to one another.

†One of the many reasons why humans don't seem to need an enormous genome is that the gene products—proteins—involved in our construction and development often have many functions. Many of the proteins we are familiar with in adults (such as hemoglobin, keratins, insulin, and the like) *do* have only one function, which led to the erroneous conclusion that this is the norm. Indeed, the "one-function-per-entity" concept is a longstanding one in science, having been credited to Aristotle. Philosopher John Thorp has called this *monotelism* (Greek, "one end") "Aristotle's worst idea."

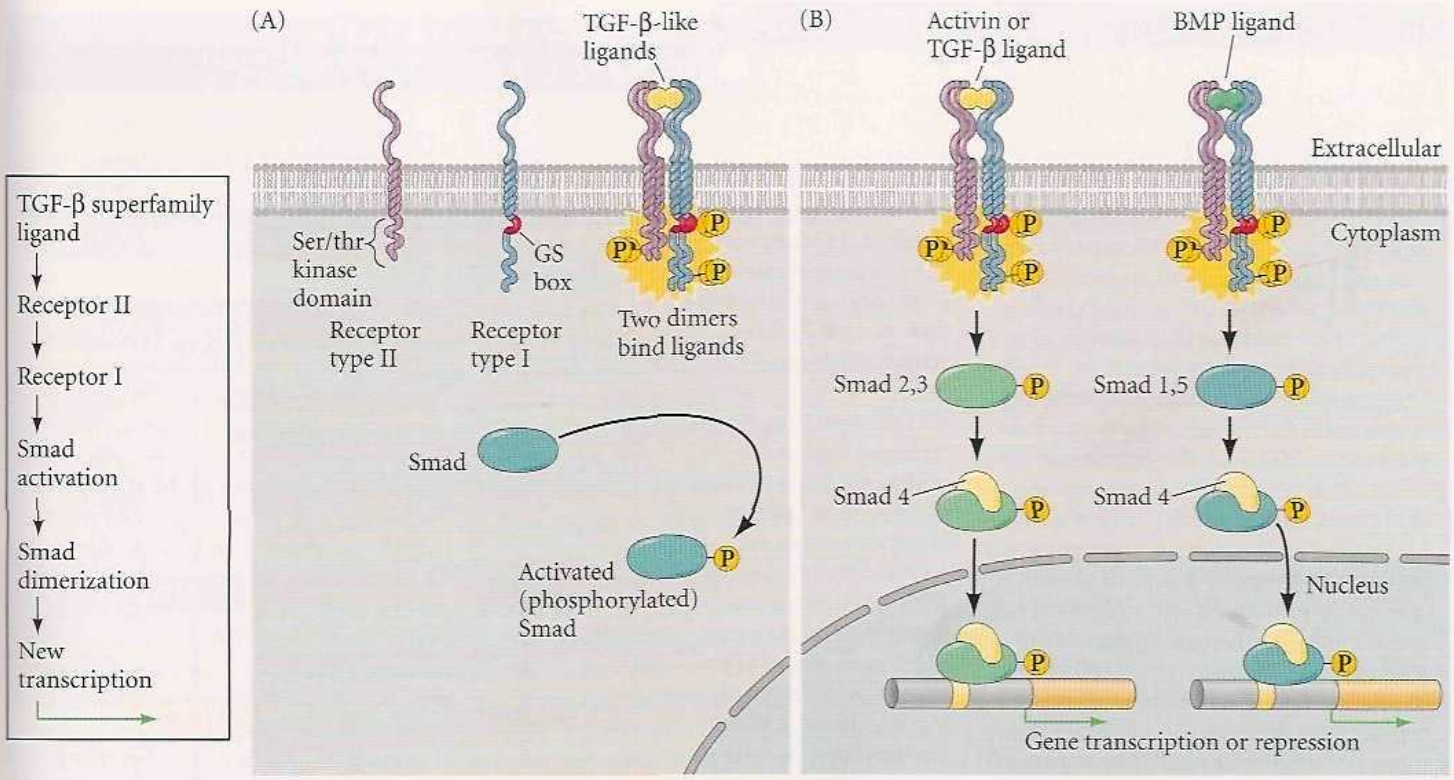


FIGURE 3.30 The Smad pathway activated by TGF-β superfamily ligands. (A) An activation complex is formed by the binding of the ligand by the type I and type II receptors. This allows the type II receptor to phosphorylate the type I receptor on particular serine or threonine residues (of the "GS box"). The phosphorylated type I receptor protein can now phosphorylate the Smad proteins. (B) Those receptors that bind TGF-β family proteins or members of the activin family phosphorylate Smads 2 and 3. Those receptors that bind to BMP family proteins phosphorylate Smads 1 and 5. These Smads can complex with Smad4 to form active transcription factors. A simplified version of the pathway is shown at the left.

producing them. Their range is determined by the amino acids in their N-terminal region, which determine whether the specific BMP will be bound by proteoglycans, thereby restricting its diffusion (Ohkawara et al. 2002).

The Nodal and activin proteins are also members of the TGF-β superfamily. These proteins are extremely important in specifying the different regions of the mesoderm and for distinguishing the left and right sides of the vertebrate body axis.

THE SMAD PATHWAY Members of the TGF-β superfamily activate members of the Smad family of transcription factors (Heldin et al. 1997; Shi and Massague 2003). The TGF-β ligand binds to a type II TGF-β receptor, which allows that receptor to bind to a type I TGF-β receptor. Once the two receptors are in close contact, the type II receptor phosphorylates a serine or threonine on the type I receptor, thereby activating it. The activated type I receptor can now phosphorylate the Smad* proteins (Figure 3.30A). Smads 1 and 5 are activated by the BMP family of TGF-β factors, while the receptors binding activin, Nodal, and the TGF-β family phosphorylate Smads 2 and 3. These phosphorylated Smads bind to Smad4 and form the transcription factor complex that will enter the nucleus (Figure 3.30B).

Other paracrine factors

Although most paracrine factors are members of one of the four families described above, some of these proteins have few or no close relatives. Epidermal growth factor, hepatocyte growth factor, neurotrophins, and stem cell factor are not included among these families, but each of these factors plays important roles during development. In addition, there are numerous paracrine factors involved almost exclusively with developing blood cells: erythropoietin, the cytokines, and the interleukins. These factors will be discussed when we detail blood cell formation in Chapter 12.

*Researchers named the Smad proteins by merging the names of the first identified members of this family: the *C. elegans* SMA protein and the *Drosophila* Mad protein.

SIDELIGHTS & SPECULATIONS

Cell Death Pathways

"To be, or not to be: that is the question." While we all are poised at life-or-death decisions, this existential dichotomy is exceptionally stark for embryonic cells. **Programmed cell death**, or **apoptosis**,* is a normal part of development (see Baehrecke 2002). In the nematode *C. elegans*, in which we can count the number of cells, exactly 131 cells die according to the normal developmental pattern. All the cells of this nematode are "programmed" to die unless they are actively told not to undergo apoptosis. In humans, as many as 10^{11} cells die in each adult each day and are replaced by other cells. (Indeed, the mass of cells we lose each year through normal cell death is close to our entire body weight!) Within the uterus, we were constantly making and destroying cells, and we generated about three times as many neurons as we eventually ended up with when we were born. Lewis Thomas (1992) has aptly noted,

> By the time I was born, more of me had died than survived. It was no wonder I cannot remember; during that time I went through brain after brain for nine months, finally contriving the one model that could be human, equipped for language.

Apoptosis is necessary not only for the proper spacing and orientation of neurons, but also for generating the middle ear space, the vaginal opening, and the spaces between our fingers and toes (Saunders and Fallon 1966; Roberts and Miller 1998; Rodriguez et al. 1997). Apoptosis prunes unneeded structures (frog tails, male mammary tissue), controls the number of cells in particular tissues (neurons in vertebrates and flies), and sculpts complex organs (palate, retina, digits, and heart).

Different tissues use different signals for apoptosis. One of the signals often used in vertebrates is bone morphogenetic protein 4 (BMP4). Some tissues, such as connective tissue, respond to BMP4 by differentiating into bone. Others, such as the frog gastrula ectoderm, respond to BMP4 by differentiating into skin. Still others, such as neural crest cells and tooth primordia, respond by degrading their DNA and dying. In the developing tooth, for instance, numerous growth and differentiation factors are secreted by the enamel knot. After the cusp has grown, the enamel knot synthesizes BMP4 and shuts itself down by apoptosis (see Chapter 10; Vaahtokari et al. 1996).

In other tissues, the cells are "programmed" to die, and will remain alive only if some growth or differentiation factor is present to "rescue" them. This happens during the development of mammalian red blood cells. The red blood cell precursors in the mouse liver need the hormone erythropoietin in order to survive. If they do not receive it, they undergo apoptosis. The erythropoietin receptor works through the JAK-STAT pathway, activating the Stat5 transcription factor. In this way, the amount of erythropoietin present can determine how many red blood cells enter the circulation.

One of the pathways for apoptosis was largely delineated through genetic studies of *C. elegans*. Indeed, the importance of this pathway was recognized by awarding a Nobel Prize to Sydney Brenner, Bob Horvitz, and Jonathan Sulston in 2002. It was found that the proteins encoded by the *ced-3* and *ced-4* genes were essential for

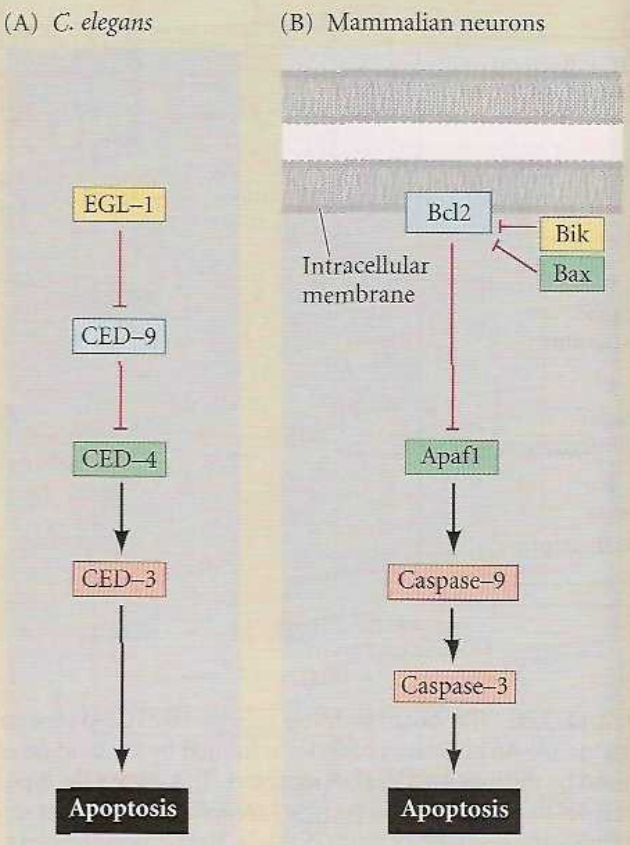


Figure 3.31 Apoptosis pathways in nematodes and mammals. (A) In *C. elegans*, the CED-4 protein is a protease-activating factor that can activate the CED-3 protease. The CED-3 protease initiates the cell destruction events. CED-9 can inhibit CED-4 (and CED-9 can be inhibited upstream by EGL-1). (B) In mammals, a similar pathway exists, and appears to function in a similar manner. In this hypothetical scheme for the regulation of apoptosis in mammalian neurons, Bcl-X_L (a member of the Bcl2 family) binds Apaf1 and prevents it from activating the precursor of caspase-9. The signal for apoptosis allows another protein (here, Bik) to inhibit the binding of Apaf1 to Bcl-X_L. Apaf1 is now able to bind to the caspase-9 precursor and cleave it. Caspase-9 dimerizes and activates caspase-3, which initiates apoptosis. The same colors are used to represent homologous proteins. (After Adams and Cory 1998.)

*The term *apoptosis* (both *p*s are pronounced) comes from the Greek word for the natural process of leaves falling from trees or petals falling from flowers. Apoptosis is an active process that can be subject to evolutionary selection. A second type of cell death, *necrosis*, is a pathological death caused by external factors such as inflammation or toxic injury.

SIDELIGHTS & SPECULATIONS (Continued)

apoptosis, and that in the cells that did not undergo apoptosis, those genes were turned off by the product of the *ced-9* gene (Figure 3.31A; Hengartner et al. 1992). The CED-4 protein is a protease-activating factor that activates the gene for CED-3, a protease that initiates the destruction of the cell. The CED-9 protein can bind to and inactivate CED-4. Mutations that inactivate the gene for CED-9 cause numerous cells that would normally survive to activate their *ced-3* and *ced-4* genes and die, leading to the death of the entire embryo. Conversely, gain-of-function mutations in the *ced-9* gene cause its protein to be made in cells that would normally die, resulting in those cells' survival. Thus, the *ced-9* gene appears to be a binary switch that regulates the choice between life and death on the cellular level. It is possible that every cell in the nematode embryo is poised to die, with those cells that survive being rescued by the activation of the *ced-9* gene.

The CED-3 and CED-4 proteins are at the center of the apoptosis pathway that is common to all animals studied. The trigger for apoptosis can be a developmental cue such as a particular molecule (e.g., BMP4 or glucocorticoids) or the loss of adhesion to a matrix. Either type of cue can activate CED-3 or CED-4 proteins or inactivate CED-9 molecules. In mammals, the homologues of the CED-9 protein are members of the Bcl2 family (which includes Bcl2, Bcl-X, and similar proteins; Figure 3.31B). The functional similarities are so strong that if an active human *BCL2* gene is placed in *C. elegans* embryos, it prevents normally occurring cell death (Vaux et al. 1992).

The mammalian homologue of CED-4 is Apaf1 (apoptotic protease activating factor 1), and it participates in the cytochrome c-dependent activation of the mammalian CED-3 homologues, the proteases caspase-9 and caspase-3 (see Figure 3.31B; Shaham and Horvitz 1996; Cecconi et al. 1998; Yoshida et al. 1998). Activation of the caspase proteins results in autodigestion—caspases are strong proteases that digest the cell from within, cleaving cellular proteins and fragmenting the DNA.

While apoptosis-deficient nematodes deficient for CED-4 are viable (despite having 15% more cells than wild-type worms), mice with loss-of-function mutations for either *caspase-3* or *caspase-9* die around birth from massive cell overgrowth in the nervous system (Figure 3.32; Kuida et al. 1996, 1998; Jacobson et al. 1997). Mice homozygous for targeted deletions of *Apaf1* have similarly severe craniofacial abnormalities, brain overgrowth, and webbing between their toes.

There are instances where cell death is the normal state unless some ligand "rescues" the cells. In the chick neural tube, Patched protein (a Hedgehog receptor) will activate caspases. The binding of Sonic hedgehog (from the notochord and ventral neural tube cells) suppresses Patched, and the caspases are not activated to start apoptosis (Thibert et al. 2003). Such "dependence receptors" probably prevent neural cells from proliferating outside the proper tissue, and the loss of such receptors is associated with cancers (Porter and Dhakshinamoorty 2004). Moreover, we will soon see that certain epithelial cells must be attached to the extracellular matrix in order to function. If the cell is removed from the matrix, the apoptosis pathway is activated and the cell dies (Jan et al. 2004). This, too, is probably a mechanism that prevents cancers once cells have lost their adhesion to extracellular matrix proteins.

See WEBSITE 3.2
The uses of apoptosis

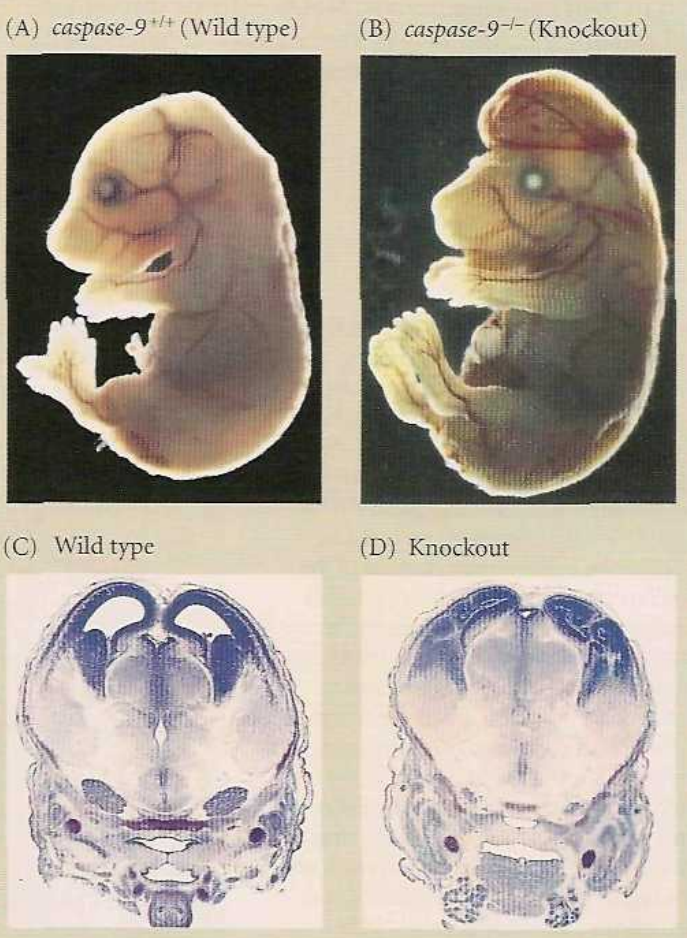


Figure 3.32 Disruption of normal brain development by blocking apoptosis. In mice in which the genes for caspase-9 have been knocked out, normal neural apoptosis fails to occur, and the overproliferation of brain neurons is obvious. (A) 6-day embryonic wild-type mouse. (B) A *caspase-9* knockout mouse of the same age. The enlarged brain protrudes above the face, and the limbs are still webbed. (C,D) This effect is confirmed by cross sections through the forebrain at day 13.5. The knockout exhibits thickened ventricle walls and the near-obliteration of the ventricles. (From Kuida et al. 1998.)

Juxtacrine Signaling

In juxtacrine interactions, proteins from the inducing cell interact with receptor proteins of adjacent responding cells *without diffusing from the cell producing it*. Two of the most widely used families of juxtacrine factors are the *Notch proteins* (which bind to a family of ligands exemplified by the Delta protein) and the **eph receptors** and their **ephrin ligands**. When the ephrin on one cell binds with the eph receptor on an adjacent cell, signals are sent to each of the two cells (Davy et al. 2004; Davy and Soriano 2005). These signals are often those of either attraction or repulsion, and ephrins are often seen where cells are being told where to migrate or where boundaries are forming. We will see the ephrins and the eph receptors functioning in the formation of blood vessels, neurons, and somites. For the moment, we will look at the Notch proteins and their ligands.

The Notch pathway: Juxtaposed ligands and receptors

While most known regulators of induction are diffusible proteins, some inducing proteins remain bound to the inducing cell surface. In one such pathway, cells expressing the **Delta**, **Jagged**, or **Serrate** proteins in their cell membranes activate neighboring cells that contain **Notch** protein in their cell membranes. Notch extends through the cell membrane, and its external surface contacts Delta, Jagged, or Serrate proteins extending out from an adjacent cell. When complexed to one of these ligands, Notch undergoes a conformational change that enables a part of its cytoplasmic domain to be cut off by the presenilin-1 protease. The cleaved portion enters the nucleus and binds to a dormant transcription factor of the CSL family. When bound to the Notch protein, the CSL transcription factors activate their target genes (Figure 3.33; Lecourtois and Schweisguth 1998; Schroeder et al. 1998; Struhl and Adachi 1998). This activation is thought to involve the recruitment of histone acetyltransferases (Wallberg et al. 2002). Thus, Notch can be considered as a transcription factor tethered to the cell membrane. When the attachment is broken, Notch (or a piece of it) can detach from the cell membrane and enter the nucleus (Kopan 2002).

Notch proteins are involved in the formation of numerous vertebrate organs—kidney, pancreas, and heart—and they are extremely important receptors in the nervous system. In both the vertebrate and *Drosophila* nervous systems, the binding of Delta to Notch tells the receiving cell not to become neural (Chitnis et al. 1995; Wang et al. 1998). In the vertebrate eye, the interactions between Notch and its ligands seem to regulate which cells become optic neurons and which become glial cells (Dorsky et al. 1997; Wang et al. 1998). Notch proteins are also important in the patterning of the nematode vulva. The vulval precursor cell closest to the anchor cell becomes the central vulva cell, and this cell is able to prevent its neighbors from becoming central vulval cells by signaling to them through its Notch homologue, the LIN-12 receptor (Berset et al. 2001).

See WEBSITE 3.3 Notch mutations

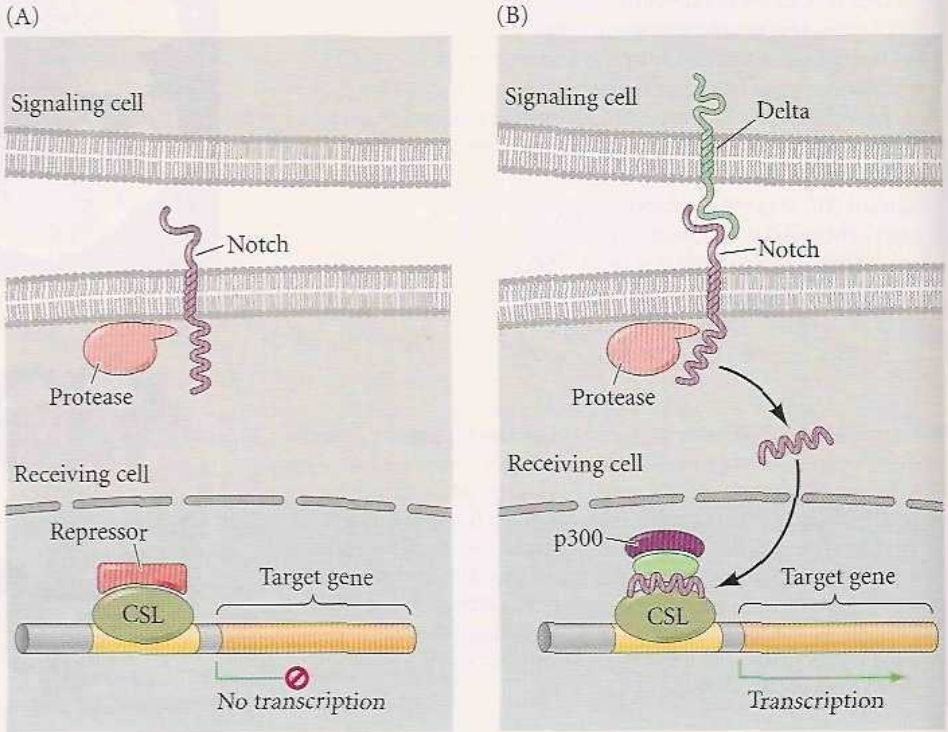


FIGURE 3.33 Mechanism of Notch activity. (A) Prior to Notch signaling, a CSL transcription factor (such as Suppressor of hairless or CBF1) is on the enhancer of Notch-regulated genes. The CSL binds repressors of transcription. (B) Model for the activation of Notch. A ligand (Delta, Jagged, or Serrate protein) on one cell binds to the extracellular domain of the Notch protein on an adjacent cell. This binding causes a shape change in the intracellular domain of Notch, which activates a protease. The protease cleaves Notch and allows the intracellular region of the Notch protein to enter the nucleus and bind the CSL transcription factor. This intercellular region of Notch displaces the repressor proteins and binds activators of transcription, including the histone acetyltransferase p300. The activated CSL can then transcribe its target genes. (After Koziol-Dube, Pers. Comm.)

SIDELIGHTS & SPECULATIONS

Juxtacrine Signaling and Cell Patterning

Induction does indeed occur on the cell-to-cell level, and one of the best examples is the formation of the vulva in the nematode worm *Caenorhabditis elegans*. Remarkably, the signal transduction pathways involved turn out to be the same as those used in the formation of retinal receptors in *Drosophila*; only the targeted transcription factors are different. In both cases, an epidermal growth factor-like inducer activates the RTK pathway.

Vulval induction in *C. elegans*

Most *C. elegans* individuals are hermaphrodites. In their early development, they are male and the gonad produces sperm, which is stored for later use. As they grow older, they develop ovaries. The eggs "roll" through the region of sperm storage, are fertilized inside the nematode, and then pass out of the body through the vulva (see Figure 5.43).

The formation of the vulva in *C. elegans* represents a case in which one inductive signal generates a variety of cell types. This organ forms during the larval stage from six cells called the **vulval precursor cells** (**VPCs**). The cell connecting the overlying gonad to the vulval precursor cells is called the **anchor cell** (Figure 3.34). The anchor cell secretes the LIN-3 protein, a paracrine factor (similar to mammalian epidermal growth factor, or EGF) that activates the RTK pathway (Hill and Sternberg 1992). If the anchor cell is destroyed (or if the *lin-3* gene is mutated), the VPCs will not form a vulva, but instead become part of the hypodermis (skin) (Kimble 1981).

The six VPCs influenced by the anchor cell form an **equivalence group**. Each member of this group is competent to become induced by the anchor cell and can assume any of three fates, depending on its proximity to the anchor cell. The cell directly beneath the anchor cell divides to form the central vulval cells. The two cells flanking that central cell divide to become the lateral vulval cells, while the three cells farther away from the anchor cell generate hypodermal cells. If the anchor cell is destroyed, all six cells of the equivalence group

(Continued on next page)

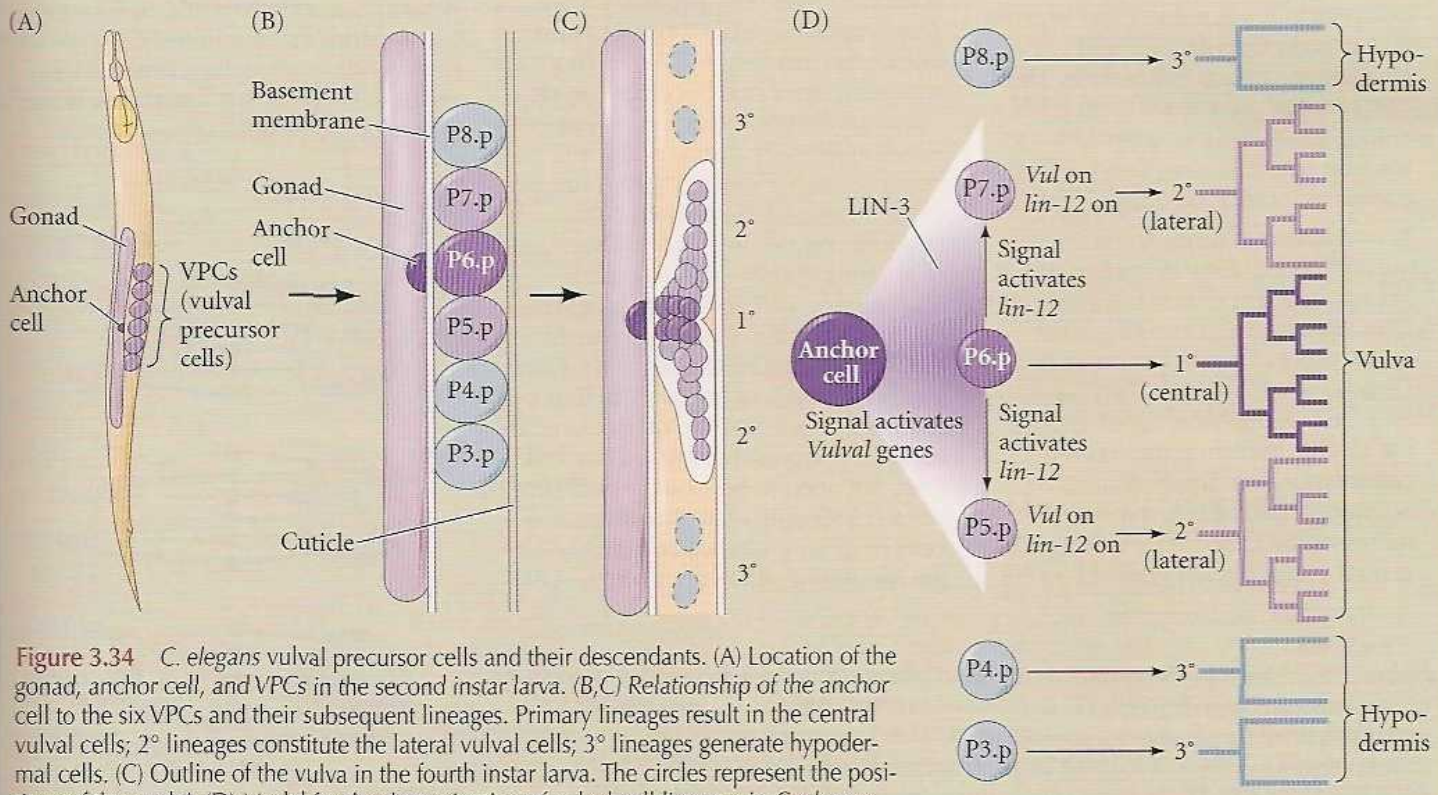


Figure 3.34 *C. elegans* vulval precursor cells and their descendants. (A) Location of the gonad, anchor cell, and VPCs in the second instar larva. (B,C) Relationship of the anchor cell to the six VPCs and their subsequent lineages. Primary lineages result in the central vulval cells; 2° lineages constitute the lateral vulval cells; 3° lineages generate hypodermal cells. (C) Outline of the vulva in the fourth instar larva. The circles represent the positions of the nuclei. (D) Model for the determination of vulval cell lineages in *C. elegans*. The LIN-3 signal from the anchor cell causes the determination of the P6.p cell to generate the central vulval lineage (dark purple). Lower concentrations of LIN-3 cause the P5.p and P7.p cells to form the lateral vulval lineages. The P6.p (central lineage) cell also secretes a short-range juxtacrine signal that induces the neighboring cells to activate the LIN-12 (Notch) protein. This signal prevents the P5.p and P7.p cells from generating the primary, central vulval cell lineage. (After Katz and Sternberg 1996.)

SIDELIGHTS & SPECULATIONS (Continued)

divide once and contribute to the hypodermal tissue. If the three central VPCs are destroyed, the three outer cells, which normally form hypodermis, generate vulval cells instead.

The LIN-3 protein is received by the LET-23 receptor tyrosine kinase on the VPCs, and the signal is transferred to the nucleus through the RTK pathway. The target of the kinase cascade is the LIN-31 protein (Tan et al. 1998). When this protein is phosphorylated in the nucleus, it loses its inhibitory protein partner and is able to function as a transcription factor, promoting vulval cell fates. Two mechanisms coordinate the formation of the vulva through this induction, as shown in Figure 3.34 (Katz and Sternberg 1996; Félix 2007):

1. The LIN-3 protein forms a concentration gradient. Here, the VPC closest to the anchor cell (i.e., the P6.p cell) receives the highest concentration of LIN-3 protein and generates the central vulval cells. The two VPCs adjacent to it (P5.p and P7.p) receive a lower amount of LIN-3 and become the lateral vulval cells. The VPCs farther away from the anchor cell do not receive enough LIN-3 to have an effect, so they become hypodermis (Katz et al. 1995).

2. In addition to forming the central vulval lineage, the VPC closest to the anchor cell also signals laterally to the two adjacent (P5.p and P7.p) cells and instructs them not to generate the central vulval lineages. The P5.p and P7.p cells receive the signal through the LIN-12 (Notch) proteins on their cell membranes. The Notch signal activates a microRNA, *mir-61*, which represses the gene that would specify central vulval fate, as well as promoting those genes that are involved in forming the lateral vulval cells (Sternberg 1988; Yoo et al. 2005). The lateral cells do not instruct the peripheral VPCs to do anything, so they become hypodermis (Koga and Ohshima 1995; Simske and Kim 1995).

Cell-cell interactions and chance in the determination of cell types

The development of the vulva in *C. elegans* offers several examples of induction on the cellular level. We have already discussed the reception of the EGF-like LIN-3 signal by the cells of the equivalence group that forms the vulva. But before this induction occurs, there is an earlier interaction that forms the anchor cell. The formation of the anchor cell is mediated by *lin-12*, the *C. elegans* homologue of the *Notch* gene. In wild-type *C. elegans* hermaphrodites, two adjacent cells, Z1.ppp and Z4.aaa, have the potential to become the anchor cell. They interact in a manner that causes one of them to become the anchor cell while the other one becomes the precursor of the uterine tissue. In loss-of-function *lin-12* mutants, both cells become anchor cells, whereas in gain-of-function mutations, both cells become uterine precursors (Greenwald et al. 1983). Studies using genetic mosaics and cell ablations have shown that this decision is made in the second larval stage, and that the *lin-12* gene needs to function only in that cell destined to become the uterine precursor cell. The presumptive anchor cell does not need it. Seydoux and Greenwald (1989) speculate that these two cells originally synthesize both the signal for uterine differentiation (the LAG-2 protein, homologous to Delta) and the receptor for this molecule (the LIN-12 protein, homologous to Notch; Wilkinson et al. 1994).

During a particular time in larval development, the cell that, by chance, is secreting more LAG-2 causes its neighbor to cease its production of this differentiation signal and to increase its production of LIN-12 protein. The cell secreting LAG-2 becomes the gonadal anchor cell, while the cell receiving the signal through its LIN-12 protein becomes the ventral uterine precursor cell (Figure 3.35). Thus, the two cells are thought to determine each other prior to their respective differentiation events. When the LIN-12 protein is used again during vulva formation, it is activated by the primary vulval lineage to stop the lateral vulval cells from forming the central vulval phenotype (see Figure 3.34). Thus, the anchor cell/ventral uterine precursor decision illustrates two important aspects of determination in two originally equivalent cells. First, the initial difference between the two cells is created by chance. Second, this initial difference is reinforced by feedback.

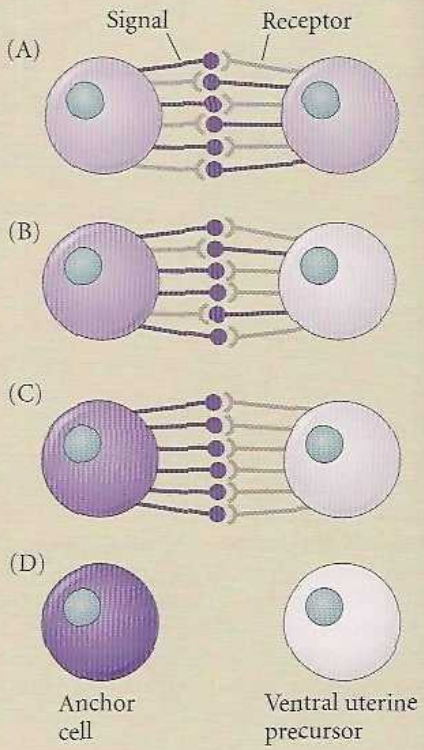


Figure 3.35 Model for the generation of two cell types (anchor cell and ventral uterine precursor) from two equivalent cells (Z1.ppp and Z4.aaa) in *C. elegans*. (A) The cells start off as equivalent, producing fluctuating amounts of signal and receptor (inverted arrow). The *lag-2* gene is thought to encode the signal; the *lin-12* gene is thought to encode the receptor. Reception of the signal turns down LAG-2 (Delta) production and upregulates LIN-12 (Notch). (B) A stochastic (chance) event causes one cell to produce more LAG-2 than the other cell at some particular critical time. This stimulates more LIN-12 production in the neighboring cell. (C) This difference is amplified, since the cell producing more LIN-12 produces less LAG-2. Eventually, just one cell is delivering the LAG-2 signal, and the other cell is receiving it. (D) The signaling cell becomes the anchor cell; the receiving cell becomes the ventral uterine precursor. (After Greenwald and Rubin 1992.)

Maintaining the Differentiated State

Development obviously means more than initiating gene expression. For a cell to become committed to a particular phenotype, gene expression must be maintained. Evolution has resulted in four major pathways for maintaining differentiation once it has been initiated (Figure 3.36):

1. The transcription factor whose gene is activated by a signal transduction cascade can bind to the enhancer of its own gene. In this way, once the transcription factor is made, its synthesis becomes independent of the signal that induced it originally. The MyoD transcription factor in muscle cells is produced in this manner.
2. A cell can stabilize its differentiation by synthesizing proteins that act on chromatin to keep the gene accessible. Such proteins include the Trithorax family discussed in Chapter 2.
3. A cell can maintain its differentiation in an autocrine fashion. If differentiation is dependent on a particular signaling molecule, the cell can make both that signaling molecule and that molecule's receptor. This produces a "community effect" (Grobstein 1955; Saxén and Wartiovaara 1966; Gurdon 1988), where the capacity to express a developmental potential exists only when a critical cell density of induced cells is present. In other words, once a group of cells has been induced, autocrine factors can sustain that induction and complete their differentiation.* In *Xenopus* muscle development, this community effect is mediated through FGF signaling. Standley and colleagues (2001) have shown (1) that FGF signaling can simulate the community effect in isolated muscle precursor cells; (2) that the muscle precursor cells have the receptors for FGFs at the critical time; and

*Community effect is also extremely important in bacterial development. Here it is called "quorum sensing," and it is critical in permitting emergent phenotypes such as light production, biofilm formation, invasiveness, and virulence. These phenotypes are expressed only in groups of bacteria and not in individuals. Each bacterium makes a small amount of a diffusible autocrine inducer that will induce the phenotype only at relatively high concentrations (see Zhu et al. 2002; Podbielski and Kreikemeyer 2004).

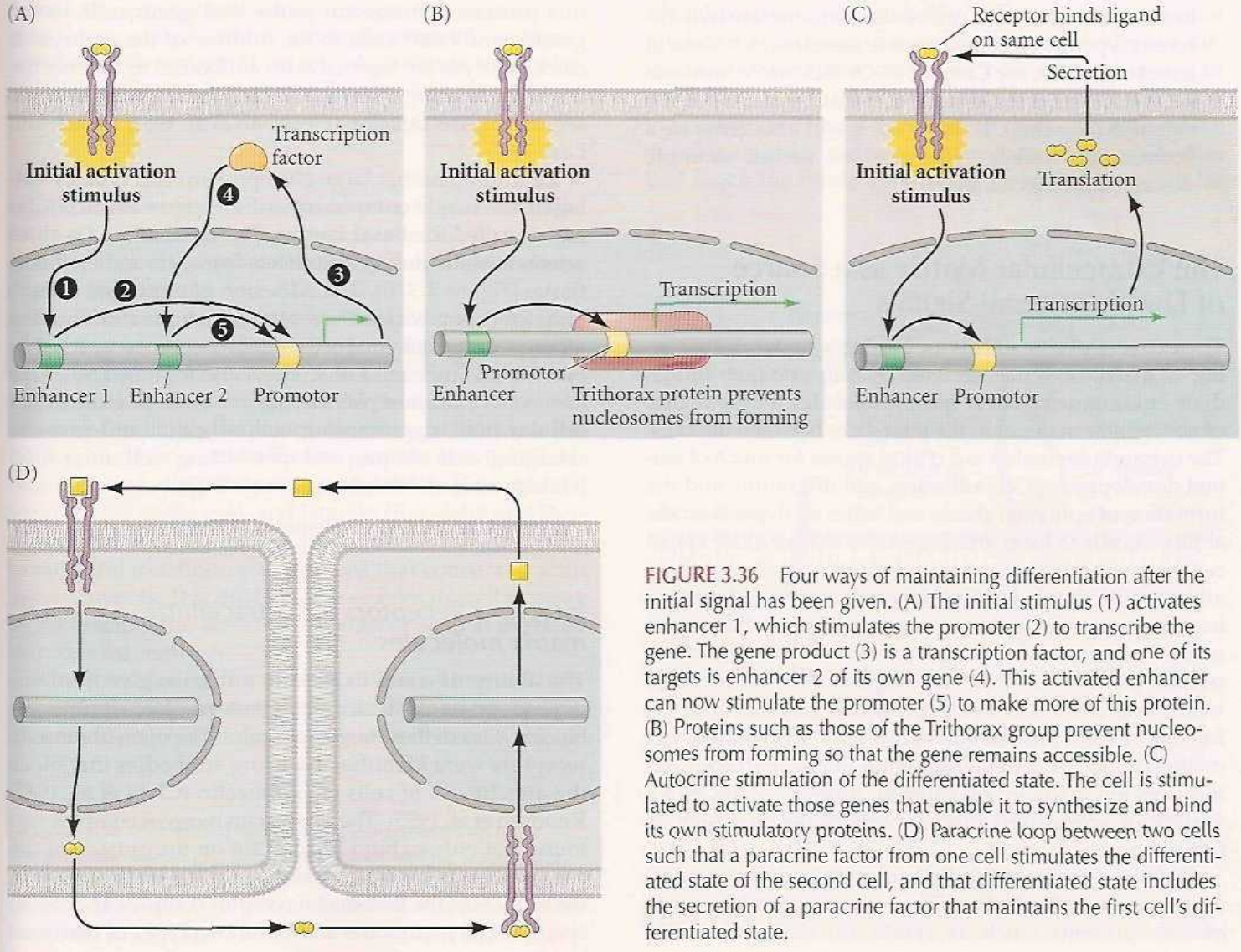


FIGURE 3.36 Four ways of maintaining differentiation after the initial signal has been given. (A) The initial stimulus (1) activates enhancer 1, which stimulates the promoter (2) to transcribe the gene. The gene product (3) is a transcription factor, and one of its targets is enhancer 2 of its own gene (4). This activated enhancer can now stimulate the promoter (5) to make more of this protein. (B) Proteins such as those of the Trithorax group prevent nucleosomes from forming so that the gene remains accessible. (C) Autocrine stimulation of the differentiated state. The cell is stimulated to activate those genes that enable it to synthesize and bind its own stimulatory proteins. (D) Paracrine loop between two cells such that a paracrine factor from one cell stimulates the differentiated state of the second cell, and that differentiated state includes the secretion of a paracrine factor that maintains the first cell's differentiated state.

(3) that the muscle precursor cells express FGFs at this time. It thus appears that part of the developmental program for *Xenopus* muscle cells is to make an FGF protein to which those same cells can also respond (i.e., an autocrine factor). This protein has to be present in sufficiently high density for the continuation of the processes leading to muscle development.

In the ectoderm of sea urchins, the identity of a group of cells that become the ectoderm around the mouth is also coordinated in this manner. Here, a Nodal paracrine factor from one cell causes the phosphorylation of Smad2 transcription factor in the neighboring cell receiving the signal. This activated transcription factor causes the neighboring cell to also make and secrete a Nodal signal. This keeps all the cells in the field actively secreting Nodal (Bolouri and Davidson 2009).

4. A cell may interact with its neighboring cells such that each one stimulates the differentiation of the other, and part of each neighbor's differentiated phenotype is the production of a paracrine factor that stimulates the other's phenotype. This type of I-scratch-your-back-you-scratch-mine strategy is found in the neighboring cells of the developing vertebrate limb and insect segments. This technique can be used to generate two different types of cells next to one another (as it does in insect segments; see Chapter 6). Or this mechanism can be used to generate a single population of cells (if the signal is the same). If this is so, it will also generate a "community effect," like the sea urchin example described for mechanism 3.

The Extracellular Matrix as a Source of Developmental Signals

The **extracellular matrix** is an insoluble network consisting of macromolecules secreted by cells into their immediate environment. These macromolecules form a region of noncellular material in the interstices between the cells. The extracellular matrix is a critical region for much of animal development. Cell adhesion, cell migration, and the formation of epithelial sheets and tubes all depend on the ability of cells to form attachments to extracellular matrices. In some cases, as in the formation of epithelia, these attachments have to be extremely strong. In other instances, as when cells migrate, attachments have to be made, broken, and made again. In some cases, the extracellular matrix merely serves as a permissive substrate to which cells can adhere, or upon which they can migrate. In other cases, it provides the directions for cell movement or the signal for a developmental event. Extracellular matrices are made up of collagen, proteoglycans, and a variety of specialized glycoprotein molecules such as fibronectin and laminin.

Proteoglycans play critically important roles in the delivery of the paracrine factors. These large molecules consist of core proteins (such as syndecan) with covalently attached glycosaminoglycan polysaccharide side chains. Two of the most widespread proteoglycans are heparan sulfate and chondroiton sulfate proteoglycans. Heparan sulfate proteoglycans can bind many members of the TGF-β, Wnt, and FGF families, and they appear to be essential for presenting the paracrine factor in high concentrations to their receptors. In *Drosophila*, *C. elegans*, and mice, mutations that prevent proteoglycan protein or carbohydrate synthesis block normal cell migration, morphogenesis, and differentiation (Garcia-Garcia and Anderson 2003; Hwang et al. 2003; Kirn-Safran et al. 2004).

The large glycoproteins are responsible for organizing the matrix and the cells into an ordered structure. **Fibronectin** is a very large (460 kDa) glycoprotein dimer synthesized by numerous cell types. One function of fibronectin is to serve as a general adhesive molecule, linking cells to one another and to other substrates such as collagen and proteoglycans. Fibronectin has several distinct binding sites, and their interaction with the appropriate molecules results in the proper alignment of cells with their extracellular matrix (Figure 3.37A). Fibronectin also has an important role in cell migration, since the "roads" over which certain migrating cells travel are paved with this protein. Fibronectin paths lead germ cells to the gonads and heart cells to the midline of the embryo. If chick embryos are injected with antibodies to fibronectin, the heart-forming cells fail to reach the midline, and two separate hearts develop (Heasman et al. 1981; Linask and Lash 1988).

Laminin (another large glycoprotein) and **type IV collagen** are major components of a type of extracellular matrix called the **basal lamina**. The basal lamina is characteristic of the closely knit sheets that surround epithelial tissue (Figure 3.37B). The adhesion of epithelial cells to laminin (upon which they sit) is much greater than the affinity of mesenchymal cells for fibronectin (to which they must bind and release if they are to migrate). Like fibronectin, laminin plays a role in assembling the extracellular matrix, promoting cell adhesion and growth, changing cell shape, and permitting cell migration (Hakamori et al. 1984; Morris et al. 2003).

See VADE MECUM Elements of the ECM

Integrins: Receptors for extracellular matrix molecules

The ability of a cell to bind to adhesive glycoproteins depends on its expressing membrane receptors for the cell-binding sites of these large molecules. The main fibronectin receptors were identified by using antibodies that block the attachment of cells to fibronectin (Chen et al. 1985; Knudsen et al. 1985). The fibronectin receptor complex was found not only to bind fibronectin on the outside of the cell, but also to bind cytoskeletal proteins on the inside of the cell. Thus, the fibronectin receptor complex appears to span the cell membrane and unite two types of matrices.

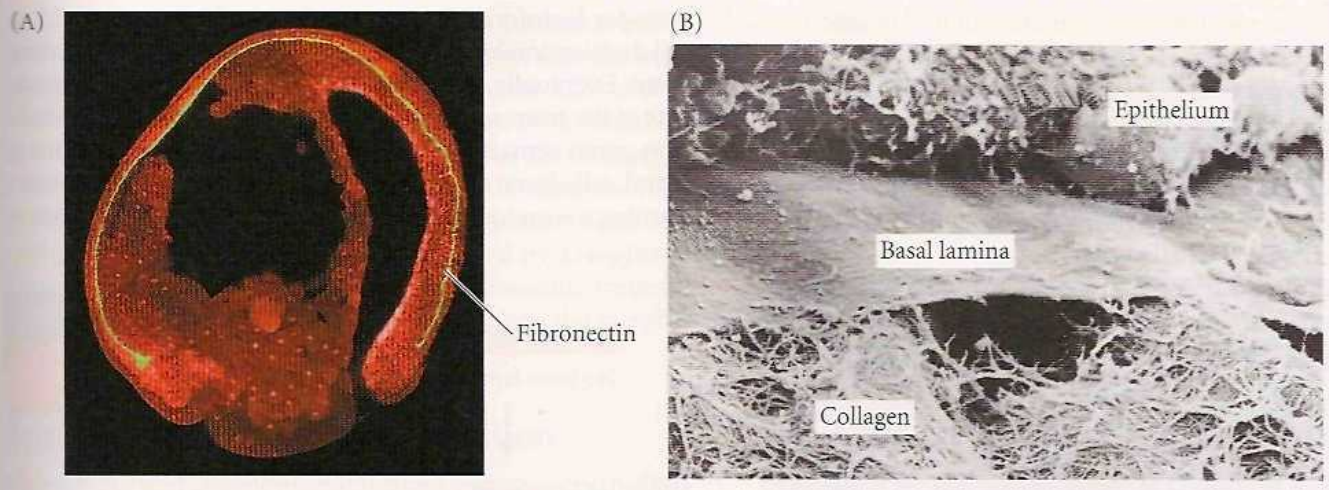


FIGURE 3.37 Extracellular matrices in the developing embryo. (A) Fluorescent antibodies to fibronectin show fibronectin deposition as a green band in the *Xenopus* embryo during gastrulation. The fibronectin will orient the movements of the mesoderm cells. (B) Fibronectin links together migrating cells, collagen, heparin sulfate proteoglycans, and other extracellular matrix proteins. This scanning electron micrograph shows the extracellular matrix at the junction of the epithelial cells (above) and mesenchymal cells (below). The epithelial cells synthesize a tight, laminin-based basal lamina, while the mesenchymal cells secrete a loose reticular lamina made primarily of collagen. (A courtesy of M. Marsden and D. W. DeSimone; B courtesy of R. L. Trelsted.)

On the outside of the cell, it binds to the fibronectin of the extracellular matrix; on the inside of the cell, it serves as an anchorage site for the actin microfilaments that move the cell (Figure 3.38).

Horwitz and co-workers (1986; Tamkun et al. 1986) have called this family of receptor proteins **integrins** because they *integrate* the extracellular and intracellular scaffolds, allowing them to work together. On the extracellular side, integrins bind to the sequence arginine-glycine-aspartate (RGD), found in several adhesive proteins in extracellular matrices, including fibronectin, vitronectin (found in the basal lamina of the eye), and laminin (Ruoslahti and Pierschbacher 1987). On the cytoplasmic side, integrins bind to talin and α-actinin, two proteins that connect to actin microfilaments. This dual binding enables the cell to move by contracting the actin microfilaments against the fixed extracellular matrix.

Bissell and her colleagues (1982; Martins-Green and Bissell 1995) have shown that the extracellular matrix is capable of inducing specific gene expression in developing tissues, especially those of the liver, testis, and mammary gland. In these tissues, the induction of specific transcription factors depends on cell-substrate binding (Figure 3.39; Liu et al. 1991; Streuli et al. 1991; Notenboom et al. 1996). Often, the presence of bound integrin prevents the activation of genes that specify apoptosis (Montgomery et al. 1994; Frisch and Ruoslahti 1997). The chondrocytes that produce the cartilage of our vertebrae and limbs can survive and differentiate only if they are surrounded by an extracellular matrix and are joined to that matrix through their integrins (Hirsch et al. 1997). If chondrocytes from the developing chick sternum are incubated with antibodies that block the binding of integrins to the extracellular

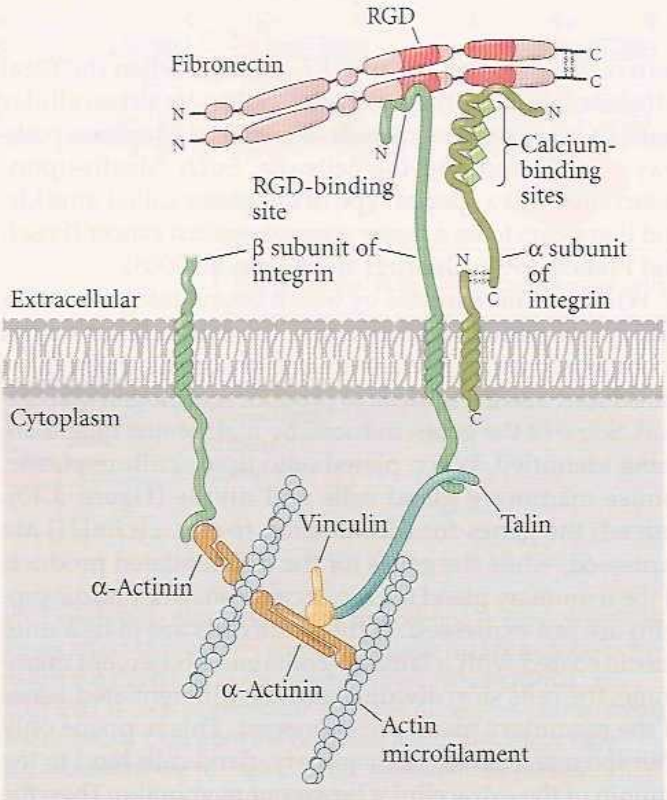


FIGURE 3.38 Simplified diagram of the fibronectin receptor complex. The integrins of the complex are membrane-spanning receptor proteins that bind fibronectin on the outside of the cell while binding cytoskeletal proteins on the inside of the cell. (After Luna and Hitt 1992.)

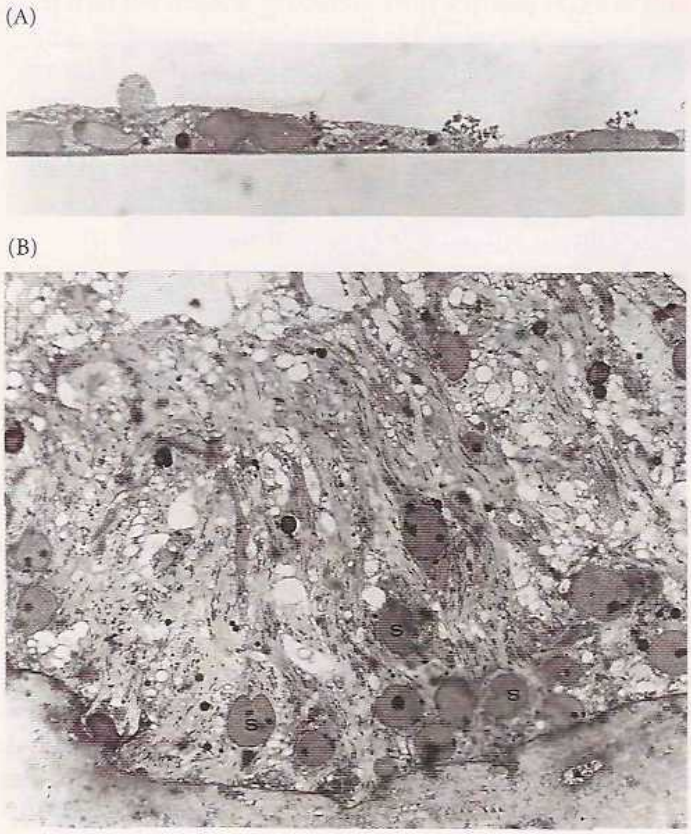


FIGURE 3.39 Role of the extracellular matrix in cell differentiation. Light micrographs of rat testis Sertoli cells grown for 2 weeks on tissue culture plastic dishes (A) and on dishes coated with basal lamina (B). The two photographs were taken at the same magnification, 1200×. (From Hadley et al. 1985, courtesy of M. Dym.)

matrix, they shrivel up and die. Indeed, when the focal adhesions linking the epithelial cell to its extracellular matrix are broken, the caspase-dependent apoptosis pathway is activated and the cells die. Such "death-upon-detachment" is a special type of apoptosis called **anoikis**, and it appears to be a major weapon against cancer (Frisch and Francis 1994; Chiarugi and Giannoni 2008).

While the mechanisms by which bound integrins inhibit apoptosis remain controversial, the extracellular matrix is obviously an important source of signals that can be transduced into the nucleus to produce specific gene expression. Some of the genes induced by matrix attachment are being identified. When plated onto tissue culture plastic, mouse mammary gland cells will divide (Figure 3.40). Indeed, the genes for cell division (*c-myc*, *cyclinD1*) are expressed, while the genes for the differentiated products of the mammary gland (casein, lactoferrin, whey acidic protein) are not expressed. If the same cells are plated onto plastic coated with a laminin-containing basement membrane, the cells stop dividing and the differentiated genes of the mammary gland are expressed. This happens only after the integrins of the mammary gland cells bind to the laminin of the extracellular basement membrane. Then the gene for lactoferrin is expressed, as is the gene for p21, a cell division inhibitor. The *c-myc* and *cyclinD1* genes become silent. Eventually, all the genes for the developmental products of the mammary gland are expressed, and the cell division genes remain turned off. By this time, the mammary gland cells have enveloped themselves in a basal lamina, forming a secretory epithelium reminiscent of the mamma-

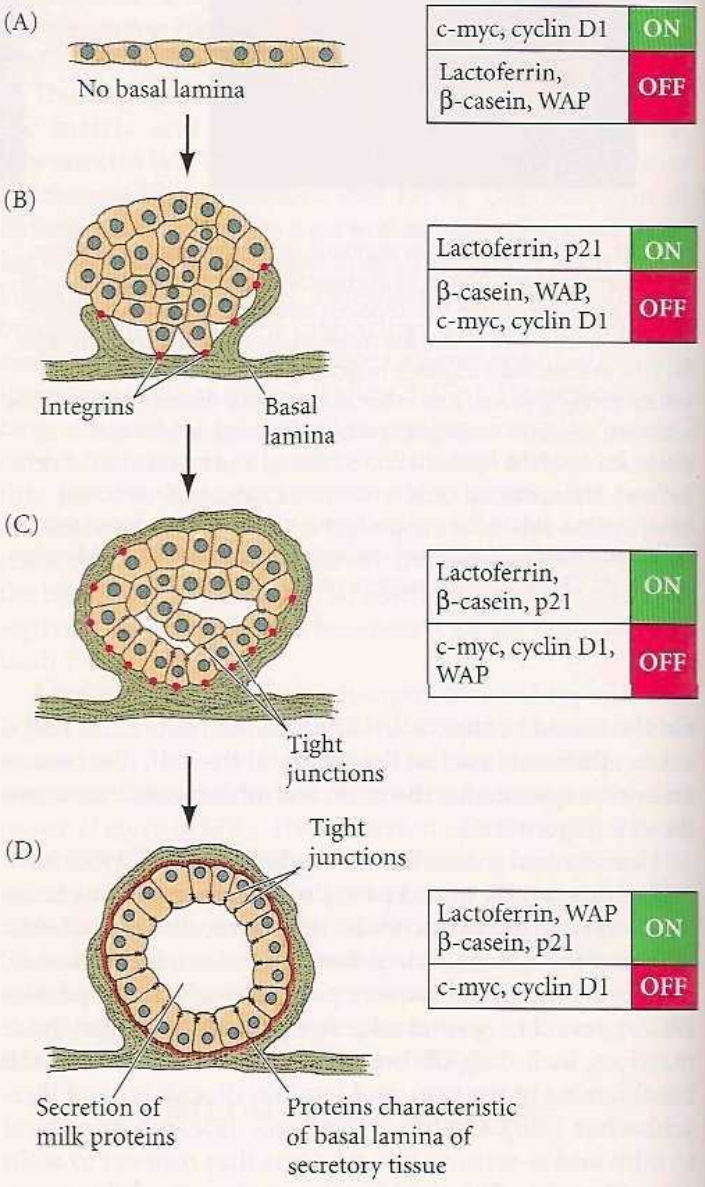


FIGURE 3.40 Basement membrane-directed gene expression in mammary gland tissue. (A) Mouse mammary gland tissue divides when placed on tissue culture plastic. The genes encoding cell division proteins are on, and the genes capable of synthesizing the differentiated products of the mammary gland—lactoferrin, casein, and whey acidic protein (WAP)—are off. (B) When these cells are placed on a basement membrane that contains laminin (basal lamina), the genes for cell division proteins are turned off, while the genes encoding inhibitors of cell division (such as p21) and the gene for lactoferrin are turned on. (C,D) The mammary gland cells wrap the basal lamina around them, forming a secretory epithelium. The genes for casein and WAP are sequentially activated. (After Bissell et al. 2003.)

ry gland tissue. The binding of integrins to laminin is essential for the transcription of the casein gene, and the integrins act in concert with prolactin (see Figure 3.23) to activate that gene's expression (Roskelley et al. 1994; Muschler et al. 1999). Several studies have shown that the binding of integrins to an extracellular matrix can stimulate the RTK pathway. When an integrin on the cell membrane of one cell binds to fibronectin or collagen secreted by a neighboring cell, the integrin can activate the RTK cascade through an adaptor protein-like complex that connects the integrin to the Ras G protein (Wary et al. 1998).

Epithelial-Mesenchymal Transition

One important developmental phenomenon, **epithelial-mesenchymal transition**, or **EMT**, integrates all the processes we have discussed in this chapter. EMT is an orderly series of events whereby epithelial cells are transformed into mesenchymal cells. In this transition, a polarized stationary epithelial cell, which normally interacts with basement membrane through its basal surface, becomes a migratory mesenchymal cell that can invade tissues and form organs in new places (Figure 3.41A). EMT is usually initiated when paracrine factors from neighboring cells activate gene expression in the target cells, instructing the target cells to downregulate their cadherins, release their attachment to laminin and other basement membrane components, rearrange their actin cytoskeleton, and secrete new extracellular matrix molecules characteristic of mesenchymal cells.

Epithelial-mesenchymal transition is critical during development (Figure 3.41B,C). Examples of developmental processes in which this transition is active include (1) the formation of neural crest cells from the dorsalmost region of the neural tube; (2) the formation of mesoderm in chick embryos, wherein cells that had been part of an epithelial layer become mesodermal and migrate into the embryo; and (3) the formation of vertebrae precursor cells from the somites, wherein these cells detach from the somite and migrate around the developing spinal cord. EMT is also important in adults, in whom it is needed for wound healing. However, the most critical adult form of EMT is seen in cancer metastasis, wherein cells that have been part of a solid tumor mass leave that tumor to invade other tissues and form secondary tumors elsewhere in the body. It appears that in metastasis, the processes that generated the cellular transition in the embryo have been reactivated, allowing cancer cells to migrate and become invasive. Cadherins are downregulated, the actin cytoskeleton is reorganized, and the cells secrete mesenchymal extracellular matrix while undergoing cell division (Acloque et al. 2009; Kalluri and Weinberg 2009).

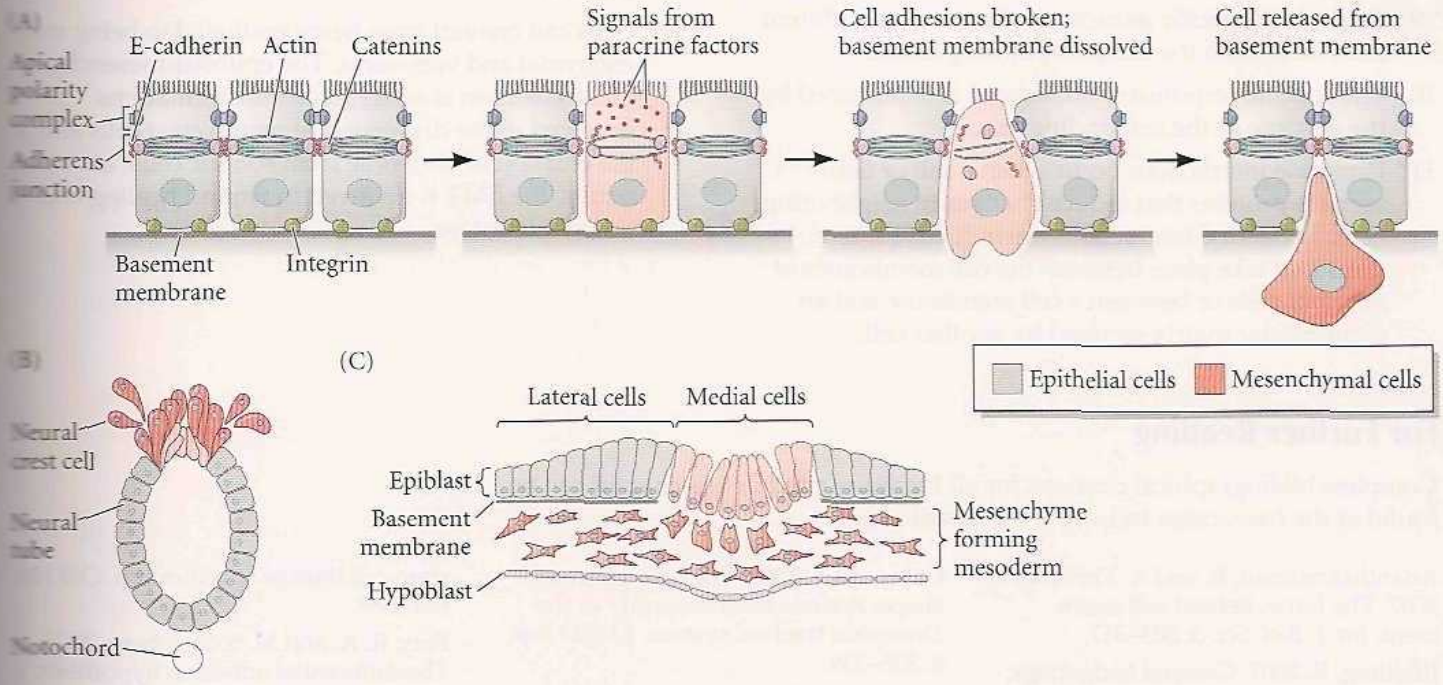


FIGURE 3.41 *Epithelial-mesenchymal transition, or EMT.* (A) Normal epithelial cells are attached to one another through adherens junctions containing cadherin, catenins, and actin rings. They are attached to the basement membrane through integrins. Paracrine factors can repress the expression of genes that encode these cellular components, causing the cell to lose polarity, lose attachment to the basement membrane, and lose cohesion with other epithelial cells. Cytoskeletal remodeling occurs, as well as the secretion of proteases and extracellular matrix molecules that enable the migration of the newly formed mesenchymal cell. (B,C) EMT is seen in vertebrate embryos during the normal formation of neural crest from the dorsal region of the neural tube (B), and during the formation of the mesoderm by mesenchymal cells delaminating from the epiblast (C).

Snapshot Summary: *Cell-Cell Communication*

1. The sorting out of one cell type from another results from differences in the cell membrane.
2. The membrane structures responsible for cell sorting out are often cadherin proteins that change the surface tension properties of the cells.
3. Cadherin proteins can cause cells to sort out by both quantitative differences (different amounts of cadherin) or qualitative differences (different types of cadherin). Cadherins appear to be critical during certain morphological changes.
4. Migration occurs through changes in the actin cytoskeleton. These changes can be directed by internal instructions (from the nucleus) or by external instructions (from the extracellular matrix or chemoattractant molecules).
5. Inductive interactions involve inducing and responding tissues.
6. The ability to respond to inductive signals depends on the competence of the responding cells.
7. Reciprocal induction occurs when the two interacting tissues are both inducers and are competent to respond to each other's signals.
8. Cascades of inductive events are responsible for organ formation.
9. Regionally specific inductions can generate different structures from the same responding tissue.
10. The specific response to an inducer is determined by the genome of the responding tissue.
11. Paracrine interactions occur when a cell or tissue secretes proteins that induce changes in neighboring cells. Juxtacrine interactions are inductive interactions that take place between the cell membranes of adjacent cells or between a cell membrane and an extracellular matrix secreted by another cell.
12. Paracrine factors are proteins secreted by inducing cells. These factors bind to cell membrane receptors in competent responding cells.
13. Competent cells respond to paracrine factors through signal transduction pathways. Competence is the ability to bind and to respond to inducers, and it is often the result of a prior induction.
14. Signal transduction pathways begin with a paracrine or juxtacrine factor causing a conformational change in its cell membrane receptor. The new shape results in enzymatic activity in the cytoplasmic domain of the receptor protein. This activity allows the receptor to phosphorylate other cytoplasmic proteins. Eventually, a cascade of such reactions activates a transcription factor (or set of factors) that activates or represses specific gene activity.
15. Programmed cell death is one possible response to inductive stimuli. Apoptosis is a critical part of life.
16. The maintenance of the differentiated state can be accomplished by positive feedback loops involving transcription factors, autocrine factors, or paracrine factors.
17. The extracellular matrix is a source of signals for the differentiating cells and plays critical roles in cell migration.
18. Cells can convert from being epithelial to being mesenchymal and vice-versa. The epithlial-mesenchymal transitiion is an series of transformations involved in the dispersion of neural crest cells and the creation of vertebrae from somitic cells. In adults, the EMT is involved in wound healing and cancer metastasis.

For Further Reading

Complete bibliographical citations for all literature cited in this chapter can be found at the free-access website **www.devbio.com**

Ananthakrishnan, R. and A. Ehrlicher. 2007. The forces behind cell movement. *Int. J. Biol. Sci.* 3: 303–317.

Breitling, R. 2007. Greased hedgehogs: New links between hedgehog signaling and cholesterol metabolism. *BioEssays* 29: 1085–1094

Cadigan, K. M. and R. Nusse. 1997. Wnt signaling: A common theme in animal development. *Genes Dev.* 24: 3286–3306.

Casanova, J. 2007. The emergence of shape: notions from the study of the *Drosophila* tracheal system. *EMBO Rep.* 4: 335–339.

Cooper, M. K., J. A. Porter, K. E. Young and P. A. Beachy. 1998. Teratogen-mediated inhibition of target tissue response to hedgehog signaling. *Science* 280: 1603–1607.

Engler, A., S. Sen, H. Sweeney and D. Discher. 2006. Matrix elasticity directs stem cell lineage specification. *Cell* 126: 677–689.

Foty, R. A. and M. S. Steinberg. 2005. The differential adhesion hypothesis: a direct evaluation. *Dev. Biol.* 278: 255–263.

Heldin, C.-H., K. Miyazono and P. ten Dijke. 1997. TGF-β signaling from cell membrane to nucleus through SMAD proteins. *Nature* 390: 465–471.

Lum, L. and P. A. Beachy. 2004. The hedgehog response network: sensors, switches, and routers. *Science* 304: 1755–1759.

Rousseau, F. and 7 others. 1994. Mutations in the gene encoding fibroblast growth factor receptor-3 in achondroplasia. *Nature* 371: 252–254.

Steinberg, M. S. and S. F. Gilbert. 2004. Townes and Holtfreter (1955): Directed movements and selective adhesion of embryonic amphibian cells. *J. Exp. Zool. A: Comp Exp. Biol.* 301: 701–706.

Go Online

WEBSITE 3.1 FGF receptor mutations. Mutations of human FGF receptors have been associated with several skeletal malformation syndromes, including syndromes wherein skull cartilage, rib cartilage, or limb cartilage fail to grow or differentiate.

WEBSITE 3.2 The uses of apoptosis. Apoptosis is used for numerous processes throughout development. This website explores the role of apoptosis in such phenomena as *Drosophila* germ cell development and the eyes of blind cave fish.

WEBSITE 3.3 Notch mutations. Mutations in the genes that encode Notch proteins can cause nervous system abnormalities in humans. Humans have more than one Notch gene and more than one ligand. Their interactions may be critical in neural development. Moreover, the association of Notch with the presenilin protease suggests that disruption of Notch functioning might lead to Alzheimer disease.

Vade Mecum

The differential adhesion hypothesis. These movies show the pioneering work of Townes and Holtfreter and Malcolm Steinberg. These experiments demonstrated the phenomenon of cell sorting and how cell surface adhesion molecules can direct sorting behaviors.

Induced cyclopia in zebrafish. As seen in the segment on zebrafish development, alcohol can act as a teratogen and induce cyclopia in these embryos.

Elements of the ECM. Movies on *Vade Mecum* review the molecular components of the extracellular matrix, how cells are influenced by them, and the work of Elizabeth Hay, who was among the first scientists to show the importance of the ECM to tissue differentiation.