2016

Principles Of Differentiation And Morphogenesis

Scott F. Gilbert
Swarthmore College, sgilber1@swarthmore.edu

R. Rice

2 Principles of Differentiation and Morphogenesis

SCOTT F. GILBERT AND RITVA RICE

Developmental biology is the science connecting genetics with anatomy, making sense out of both. The body builds itself from the instructions of the inherited DNA and the cytoplasmic system that interprets the DNA into genes and creates intracellular and cellular networks to generate the observable phenotype. Even ecological factors such as diet and stress may modify the DNA such that different phenotypes can be constructed from the same DNA. During the past two decades, the basic principles of development have become known; although this brief chapter cannot do them justice (see, for example, Gilbert, 2013), they cover the following:

- Mechanisms of differential gene expression
- Combinatorial logic of enhancers and promoters
- Signal-transduction pathways linking cell membrane and nucleus
- Mechanisms by which syndromes occur
- The repertoire of morphogenetic interactions and the molecules causing them
- Environmental agents of phenotype production

MECHANISMS OF DIFFERENTIAL GENE EXPRESSION

With few exceptions (e.g., lymphocytes and blood cells), every cell nucleus in the body contains the complete genome established in the fertilized egg. In molecular terms, the DNAs of all differentiated cells within an organism are identical. This was vividly demonstrated when entire mammalian organisms were generated from the nuclei of adult cells transplanted into enucleated oocytes (Wilmut et al., 1997; Kato et al., 1998; Wakayama et al., 1998). Thus the unused genes in differentiated cells are neither destroyed nor mutated and 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 a cell is specific for that cell.

How, then, is the inherited repertoire of genes differentially expressed during development? It appears that this can be accomplished at the 4 major steps of protein synthesis. Some genes are regulated at different steps in different cells, and certain genes can be regulated at multiple steps in the same cell:

- Differential gene transcription, regulating which of the nuclear genes are transcribed into nuclear (n)RNA
- Selective nRNA processing, regulating which of the transcribed RNAs (or which parts of such an nRNA) enter into the cytoplasm to become messenger (m)RNAs
- Selective mRNA translation, regulating which of the mRNAs in the cytoplasm become translated into proteins
- Differential protein modification, regulating which proteins are allowed to remain or function in the cell

DIFFERENTIAL GENE TRANSCRIPTION

Initiation of Transcription: Promoters and Enhancers

Whether or not a gene is active depends upon RNA polymerase binding to the promoter and the elongation of that RNA transcript. RNA polymerase II, the enzyme responsible for the transcription of protein-encoding genes, does not bind directly to naked DNA. Rather, it binds to a number of proteins, including TFIID, which creates a saddle upon which it sits, and TFIIB, which positions RNA polymerase II on the DNA in such a manner that it can read the codons. Other basal transcription factors, such as TFIIA and TFIIH, help stabilize the polymerase once it is there (Kostrewa et al. 2009).

Where and when a gene is expressed depends on another regulatory unit of the gene, the enhancer. An enhancer is a DNA sequence that can activate or repress the utilization of a promoter, controlling the efficiency and rate of transcription from that particular promoter. Enhancers can activate only *cis*-linked promoters (i.e., promoters on the same chromosome), but they can do so at great distances (some as great as 50 kb away from the promoter). Moreover, enhancers do not need to be upstream of the gene. They can also be at the 3′ end, in the introns, or even on the complementary DNA strand (Maniatis et al., 1987). Like promoters, enhancers function by binding transcription factors, but these transcription factors are often very specific and are not found in every cell. As we will see, the combination of transcription factors activates or represses gene transcription. It is estimated (Rockman and Wray, 2002) that the human species has more polymorphism in its regulatory regions than in the amino acid encoding exon regions of the genome, and mutations in the enhancer and promoter regions are major causes of human congenital anomalies (see VanderMeer and Ahituv, 2011).

Enhancer-bound transcription factors work in two basic (and non-exclusionary) ways. First, transcription factors recruit enzymes that modify the histone proteins that form the nucleosomes surrounding the DNA. Enzymes such as histone acetyltransferases chemically modify the histone proteins in ways that usually cause the dispersion of the nucleosomes in the area. Histone methyltransferases, however, often increase the stability of the nucleosomes and thereby repression of the gene. Second, they can bind to a large multimeric complex called the Mediator, whose nearly 30 protein subunits bind it to RNA polymerase II and relay developmental signals (Malik and Roeder 2010.) This forms the preinitiation complex at the promoter. The enhancer, with its bound transcription factors, is thought to loop around to contact the transcription factors at the promoter site. Most genes have several enhancers, enabling them to be expressed in multiple tissues (Figure 2). Within each enhancer, however, there are usually several transcription factor–binding sites, and the enhancer does not function unless many sites are occupied simultaneously.

Moreover, enhancers must be told where to stop. Since enhancers can work at relatively long distances, it is possible for them to activate several nearby promoters. To stop this spreading of the enhancer's power, there are insulator sequences in the DNA (Zhou et al., 1995; Bell et al., 2001). Insulators bind proteins that prevent the enhancer from activating an adjacent promoter. They are often between the enhancer and a particular promoter. For instance, the chick β-globin enhancers are located in the locus control region (LCR), which is limited by insulators on both sides. On one side is an insulator that prevents the LCR enhancers from activating odorant receptor genes (which are active in the nasal neurons) and on the other side is an insulator preventing the LCR from activating the folate receptor gene.

Transcription Factors

Transcription factors are proteins that bind to enhancer or promoter regions and interact to activate or repress the transcription of a particular gene. Most transcription factors can bind to specific DNA sequences. These proteins can be grouped into families based on similarities in structure (Table 2–1). The transcription factors within such a family share a framework structure in their DNA-binding sites, and

Table 2–1. Partial List of Transcription Factor Families and Functions

Family	Representative Transcription Factors	Some Functions
HOMEODOMAIN		
HOX	HOXA-1, HOXB-2, etc.	Axis formation
POU	PIT1, Unc-86, Oct-2	Pituitary development, neural fate
LIM	Lim-1	Head development
PAX	PAX1, -2, -3, etc.	Neural specification, eye development
Forkhead/winged helix	FOXC1	Eye and skeletal development
Basic helix-loop-helix	MYOD, achaete	Muscle and nerve specification
Basic leucine zipper	C/EBP, AP1	Liver differentiation, fat cell specification
ZINC FINGER		
Standard	WT1, Krüppel	Kidney, gonad development
Hormone receptors	Estrogen receptor	Secondary sex determination
Sry-Sox	Sry, SOXD, Sox2	Bone, primary sex determination

slight differences in the amino acids at the binding site can alter the sequence of the DNA to which the factor binds.

Transcription factors have three major domains. The first is a DNA-binding domain, which recognizes a particular DNA sequence. The second is a *trans*-activating domain, which activates or suppresses the transcription of the gene whose promoter or enhancer it has bound. Usually, the *trans*-activating domain enables the transcription factor to interact with proteins involved in binding RNA polymerase (e.g., TFIIB or TFIIE; see Sauer et al., 1995) or with nucleosome-remodeling enzymes that regulate the access of RNA polymerase to the promoter. In addition, there may be a protein–protein interaction domain, which allows the transcription factor's activity to be modulated by other transcription factors.

Transcription factors can also be grouped together based on their function. Constitutively active nuclear transcription factors are found in all cells at all times, and many are essential to the initiation of transcription. For instance, the basal transcription factors belong to this group. Regulatory transcription factors require activation to function. Some of these proteins are specific to particular cell types or specific temporal stages of development. Others are present ubiquitously but cannot function without being activated (often by the paracrine factor signaling cascades discussed further on) (Brivanlou and Darnell, 2002). It is the combinatorial use of transcription factors that is the driving force behind gene-specific transcription and eventually differentiation and morphogenesis.

Numerous diseases are caused by deficiencies of transcription factors. The first identified, transcription factoropathy, was probably the androgen insensitivity syndrome. Here the testosterone receptor, a dormant transcription factor, is absent or deficient and therefore will not bind to the DNA activating male-specific genes, even in the presence of its activator, testosterone (Meyer et al., 1975). One of the first human genetic diseases to be understood from the binding of the ligand to the receptor through the activation of chromatin was Waardenburg syndrome type II. Here, people heterozygous for the wild-type copy of microphthalmia (MITF) are deaf, have multicolored irises, and have a white forelock in their hair. Activation of this transcription factor through the protein tyrosine kinase cascade enables it to dimerize, to bind to the regulatory regions of particular genes, and to bind a histone acetyltransferase that opens a region of DNA for transcription (Figure 2; see Chapter 128).

Combinatorial Control of Transcription

The binding of a specific transcription factor to the enhancer or promoter does not ensure that that gene will be transcribed. Although "master regulatory genes" such as *PAX6* (eye) or *MYOD* (muscles) have been proposed, even they work in concert with other transcription factors to effect cell differentiation. The use of *PAX6* by different organs demonstrates the modular nature of transcriptional regulatory units. The PAX6 transcription factor is needed for mammalian eye, nervous system, and pancreatic development; mutations in the human *PAX6* gene cause severe nervous system, pancreatic, and optic abnormalities (Glaser et al., 1994; see Chapters 8 and 112). Pax6 binding sequences have been found in the enhancers of vertebrate lens crystallin genes and in the genes expressed in the endocrine cells of the pancreas (insulin, glucagon, and somatostatin).

Transcription factors work in concert with other transcription factors to activate a particular gene. For instance, in the chick δ1 lens crystallin gene, Pax6 works with Sox2, Maf-1, and Sp1. Sp1 is a general transcriptional activator found in all cells. Pax6 is found early in development throughout the head ectoderm. Sox2 is found only in those tissues that will become lens, and it is induced by the presumptive retina when the developing retinal cells contact the outer ectoderm. Thus only those cells that contain both Sox2 and Pax6 can express the lens crystallin gene. The association of the Pax6, Sox2, and Maf-1 transcription factors on the enhancer of the genes in lens cells recruits a histone acetyltransferase that can transfer acetyl groups to the histones and dissociate the nucleosomes in that area to activate the crystallin genes (Yang et al. 2007). In addition, there is a third site that can bind either an activator (the δEF3 protein) or a repressor (the δEF1 protein) of transcription. It is thought that the repressor may be critical in preventing crystallin expression in the nervous system.

Other regulatory regions that use Pax6 are the enhancers regulating the transcription of the insulin, glucagon, and somatostatin genes of the pancreas. Here, Pax6 cooperates 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; Lammert et al., 2001). New technologies (such as chromatin immunoprecipitation) have identified numerous important interactions whereby transcription factors stabilize each other's binding and function together to activate gene expression (see Wilson et al., 2010).

Pax6 also binds to the regulatory regions of the *PAX6* gene itself (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.

Thus 4 principles can be seen here: (1) transcription factors function in a combinatorial manner, wherein several work together to promote or inhibit transcription. (2) There are 2 major routes by which transcription factors become present in the nucleus: (a) through cell lineage, where the presumptive lens tissue acquires its Pax6 by being the head ectoderm and the presumptive pancreatic islets acquire Pdx1 through their being endodermal, and (b) through induction, as when the *Sox2* gene becomes expressed when the presumptive retina abuts the presumptive lens. (3) Transcription factors can continue to be synthesized after the original signal for their synthesis has ceased. (4) Another principle is seen in the example of MITF, mentioned previously: the mere presence of transcription factors in the cell is often insufficient for their binding to DNA and consequently functioning. Often, they have to be activated posttranslationally to function. MITF has to be phosphorylated in order to function as a transcription factor. When this activation occurs, it dimerizes and can bind an acetyltransferase which promotes gene expression by getting rid of nucleosomes in the vicity of the gene. The activation of dormant transcription factors is a major mechanism for the control of differentiation and morphogenesis; we will return to it later.

Although transcription factors are critical in activating genes, there appear to be two major classes of gene in the cell; the transcription factors may operate differently in these two classes. Low-CpG-content promoters (i.e., promoters characterized by a relatively low level of the nucleotide pair CpG) are usually found in those genes whose products characterize mature cells (e.g., the globins of red blood cells, the hormones of pancreatic cells, and the enzymes that carry out the normal maintenance functions of the cell). Their CpG sites are often methylated; this recruits enzymes that methylate the histones and keep the gene repressed. Therefore the default state of these promoters is "off." These genes can be activated when the methyl groups are removed

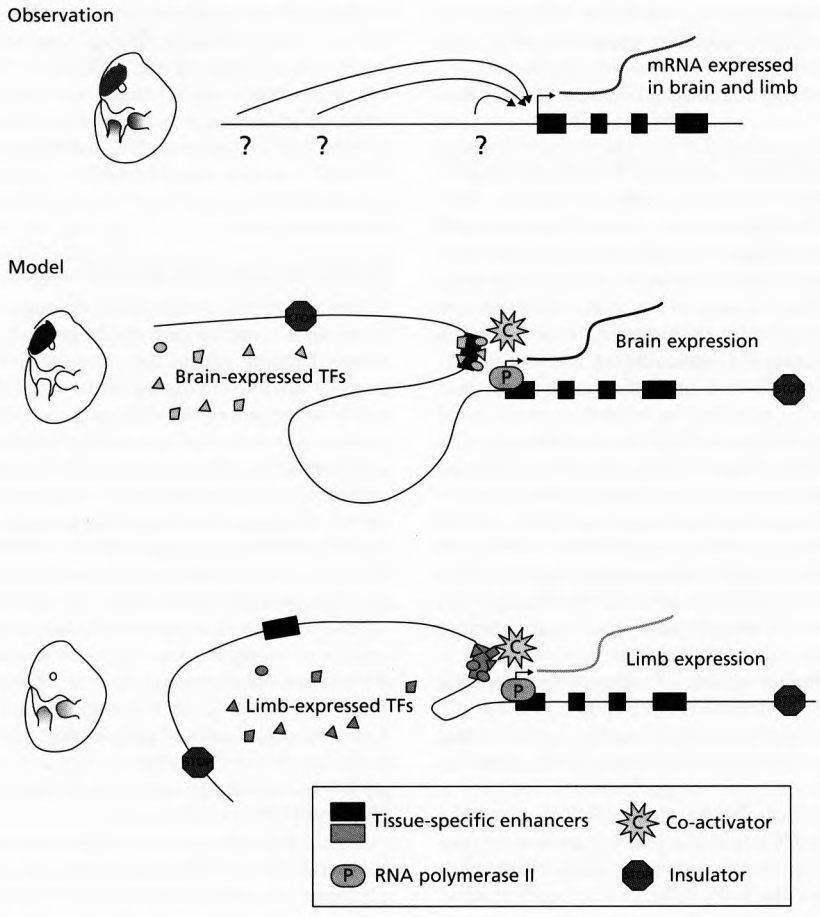

Fig. 2–1. Enhancer function and modularity. In this example, a gene is expressed in the brain and in the limbs (A). In developing brain cells (B), transcription factors found in the brain cells bind to the "brain enhancer," causing it to bind to the Mediator complex, stabilize RNA polymerase at the promoter, and modify the nucleosomes in the region of the promoter. The limb enhancer does not function, and the gene is transcribed in the brain cells. (C) In the limb, a similar process allows for transcription of the same gene in the limb cells. The gene is not transcribed in any cell type not containing transcription factors the enhancers of this gene can bind.
Source: After Visel et al., 2009.

from the DNA. This is very important in genomic imprinting, where DNA from the sperm or egg is differentially methylated such that transcription occurs only from the maternally or paternally derived gene.

High-CpG-content promoters are characteristic of the developmental regulatory genes that regulate cell fate. The DNA of these promoters is relatively unmethylated, and nucleosomes tend to be enriched with "activating" residues. As a result, these promoters usually have an RNA polymerase II protein already present on them (Hon et al 2009; Ernst and Kellis 2010). Indeed, there is often a small, truncated, transcript of nRNA already initiated (but not completed) at these promoters. DNA methylation does not appear to play a major role in regulating these promoters. Rather, the HCPs can be repressed by modifying the histone 3 to H327me3, which recruits polycomb repressive complex 2 (Peng et al., 2009; Li et al., 2010). This complex appears to inhibit further RNA polymerase binding as well as preventing the elongation of the existing nRNA transcripts. The rate-limiting step at these promoters is RNA elongation, which is regulated by another set of transcription factors.

OTHER MECHANISMS OF DEVELOPMENTAL GENE REGULATION

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 mRNA must be (1) processed into mRNA 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 (4) must therefore be posttranslationally modified to become active. Regulation can occur at any of these steps during development.

Differential Nuclear RNA Processing

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: control at the level of RNA processing and transport. There are 2 major ways in which differential RNA processing can regulate development. The first involves "censoring," determining which nuclear transcripts will be processed into cytoplasmic messages. Here different cells can select different nuclear transcripts to be processed and sent to the cytoplasm as mRNA. The same pool of nuclear transcripts can thereby give rise to different populations of cytoplasmic mRNAs in different cell types. The second mode of differential RNA processing is the splicing of the mRNA precursors into messages for different proteins, using different combinations of potential exons. If an mRNA precursor had 5 potential exons, 1 cell might use exons 1, 2, 4, and 5; a different cell might utilize exons 1, 2, and 3; and yet another cell type might use a different combination. Thus one gene can create a family of related proteins by alternative RNA splicing (Nilsen and Graveley, 2010).

This ability to create large numbers of proteins from 1 gene through differential exon splicing may be extremely important in human development. It is estimated that 92% of human genes have multiple types of mRNA. Therefore even though the human genome may contain 20,000 to 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). Whether a sequence of RNA is recognized as an exon or as an intron is a crucial step in gene regulation. What is an intron in one cell's nucleus may be an exon in another cell's nucleus.

In some instances alternatively spliced RNAs yield proteins that play similar yet distinguishable roles in the same cell. Pax6 has two splicing isoforms, and each is needed for different roles in the body. They cannot compensate for each other (Epstein et al., 1994). Similarly, 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 testis development, whereas the isoform containing the extra exon appears to be a splicing factor involved in kidney development (Hastie, 2001).

Thus proper development means not only that genes are transcribed at the appropriate time but also that the nuclear gene products are spliced appropriately. Mutations in the splice sites of genes can therefore prevent certain isoforms from arising, and it is estimated that 15% of all point mutations that result in human genetic disease are those creating splice site abnormalities (Krawczak et al., 1992; Cooper and Mattox, 1997). 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). Thus there is no β-globin from this gene, leading to 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). Congenital adrenal hypoplasia can be caused by a point mutation in the splice site for the second intron of the *CYP21* gene for 21-hydroxylase.

Some proteins and small nuclear RNAs are used throughout the body to effect differential pre-mRNA splicing, and there are some proteins that appear to be cell type–specific, regulating differential splicing in a manner characteristic for that cell. If the genes encoding these cell set–specific splicing factors are mutated, one could expect several cell-specific isoforms to be aberrant. This appears to be the case in the leading cause of hereditary infant mortality, spinal muscular atrophy. Here mutations in the gene encoding the survival of motor neurons (SMNs) protein prevent the maintenance of motor neurons. This protein is involved in splicing nRNAs in this subset of neurons (Pellizzoni et al., 1999).

RNA Translation

Once a message has been transcribed and properly spliced, it can enter the cytoplasm and be translated. However, translation is an intricately regulated mechanism that may also alter phenotypes. Some genetic diseases are due to mutations that create termination codons. For instance, a complete form of androgen insensitivity syndrome is caused by a guanine-to-adenine transition at nucleotide 2682, changing codon 717 from tryptophan to a translation stop signal (Sai et al., 1997). Codon 717 is in exon 4, and this truncated receptor thereby lacks most of its androgen-binding domain. Other mutations can alter the longevity of an mRNA, which can greatly affect the number of proteins synthesized from it. For example, hemoglobin α-Constant Spring is a naturally occurring mutation wherein the translation termination codon has been mutated to that of an amino acid codon, and the translation continues for 31 more codons (Wang et al., 1995). This readthrough results in the destabilization of the α-globin mRNA, a reduction of greater than 95% of α-globin gene expression from the affected locus and the resultant clinical disease (α-thalassemia).

In many instances, certain messages are stabilized or brought to the ribosomes by certain proteins. The most prevalent form of inherited mental retardation, fragile X syndrome, may result from the translational deficiency of certain neuronal messages. This disease usually results from the expansion and hypermethylation of CGG repeats in the 5′-untranslated region of the *FMR1* gene, which blocks transcription of this gene. *FMR1* encodes an RNA-binding protein, which appears to be critical for the translation of certain messages. Almost 85% of the FMR1 protein is associated with translating polysomes, whereas mutants in the RNA-binding domains produce severe forms of the syndrome and are not observed with the cytoplasmic polysomes (Feng et al., 1997a,b). Several studies (Brown et al., 2001; Darnell et al., 2001; Antar et al., 2006) have shown that a particular subset of mouse brain mRNA requires this protein for proper translation. Most of these genes are involved with synapse function or neuronal development. It is probable that fragile X mental retardation protein (FMRP) binds to specific mRNAs, either regulating their translation or targeting them to the dendrite, where they might await the signal for translation.

Posttranslational Modification

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," which integrates it into a relationship with numerous other proteins. Thus several changes can still take place that determine whether 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 to function. Proteins are often sequestered in certain regions, such as membranes, lysosomes, nuclei, or mitochondria; specific amino acid sequences are needed either as recognition sequences or as places for such tags. For instance, mucolipidosis II (I-cell disease) is characterized by a deficiency in the mannose-6-phosphate "address tag" put onto enzymes to target them to the lysosome. Here, there is a deficiency in GlcNAc-1-P transferase, which is involved in constructing the mannose-6-phosphate residues (Sly and Fischer, 1982).

Some proteins must 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. Diseases such as sickle cell anemia and certain types of osteogenesis imperfecta syndrome are caused by the improper assembly of protein subunits. Moreover, some proteins are not active unless they bind an ion, such as calcium, or are modified by the covalent addition of a phosphate or acetate group. This last type of protein modification will become very important in the following section as many important proteins in embryonic cells just sit there until some signal activates them.

EMBRYONIC INDUCTION

Induction and Competence

Organs are complex structures composed of numerous types of tissue. In the vertebrate eye, for example, 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 this organ 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 fate. This kind of interaction at close range between two or more cells or tissues of different history and properties is called *proximate interaction* or *induction*. There are at least two components to every inductive interaction. The first is the inducer, the tissue that produces a signal (or signals) that changes the cellular behavior of the other tissue, and the second is the responder, the tissue being induced.

Not all tissues can respond to the signal being produced by the inducer. For instance, if the optic vesicle (presumptive retina) of *Xenopus laevis* is placed in an ectopic location (i.e., in a different place from where it normally forms) underneath the head ectoderm, 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 the ectoderm in the flank or abdomen of the same organism, that ectoderm will not be able to respond. Only

the head ectoderm is competent to respond to signals from the optic vesicle by producing a lens (Saha et al., 1989; Grainger, 1992).

This ability to respond to a specific inductive signal is called *competence* (Waddington, 1940); this is not a passive state but an actively acquired condition. For example, in the developing chick and mammalian eye, the Pax6 protein appears to be important in making the ectoderm competent to respond to the inductive signal from the optic vesicle. Pax6 expression is seen in the head ectoderm, which can respond to the optic vesicle by forming lenses; and it is not seen in other regions of the surface ectoderm (Li et al., 1994). Moreover, the importance of Pax6 as a competence factor was demonstrated by recombination experiments using embryonic rat eye tissue (Fujiwara et al., 1994). Pax6 is required for the surface ectoderm to respond to the inductive signal from the optic vesicle; the inducing tissue does not need it. Pax6 is expressed in the anterior ectoderm of the embryo through the interaction of the lateral dorsal head ectoderm with the anterior neural plate (Zygar et al., 1998).

Induction is a multilevel and sequential series of activations, whereby an early activation may give the cell the competence to respond to later signals. Thus there is no single "inducer of the lens." Studies on amphibians suggest that the first lens inducers may be the pharyngeal endoderm and heart-forming mesoderm, which underlie the lens-forming ectoderm during the early and midgastrula stages (Jacobson, 1966). The anterior neural plate may produce the next signals, including a signal that promotes the synthesis of Pax6 in the anterior ectoderm. Thus, while the optic vesicle appears to be the inducer, the anterior ectoderm has already been induced by at least 2 other factors. (This situation is like that of the player who kicks the "winning goal" of a soccer match.) The optic vesicle appears to secrete 2 induction factors: 1 appears to be BMP4 (Furuta and Hogan, 1998), a paracrine factor protein that induces the transcription of the Sox2 and Sox3 transcription factors, and the other is thought to be FGF8, a paracrine factor that induces the appearance of the L-Maf transcription factor (Ogino and Yasuda, 1998; Vogel-Höpker et al., 2000; Ogino et al., 2012). The combination of Pax6, Sox2, Sox3, and L-Maf ensures the production of the lens.

Cascades of Induction: Reciprocal and Sequential Inductive Events

Another feature of induction is the reciprocal nature of many inductive interactions. Once the lens has begun forming, it can induce other tissues. One of these responding tissues is the optic vesicle itself. Now 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 (Cvekl and Piatigorsky, 1996). Such interactions are called *reciprocal inductions*.

At the same time, the developing lens also induces 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; Kanakubo et al 2006). Under the influence of the lens, the corneal ectodermal 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), which further differentiate the cornea. A third signal, the hormone thyroxine, dehydrates the tissue and makes it transparent (Hay, 1979; Bard, 1990). Thus, there are sequential inductive events and multiple causes for each induction.

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 to initiate a new gene expression in the responding cell. Without the inducing cell, the responding cell would not be 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, it is an instructive interaction. The second type of induction is *permissive interaction*. Here the responding tissue contains all the potentials that are to be expressed and needs only an environment that allows expression of these traits. For instance, many tissues need a solid substrate containing fibronectin or laminin to develop. The fibronectin or laminin does not alter the type of cell that is to be produced but only enables what has been determined to be expressed.

Regional Specificity of Induction

Some of the best-studied cases of induction are those involving the interactions of sheets of epithelial cells with adjacent mesenchymal cells. These interactions are called epithelial–mesenchymal interactions. Epithelia are sheets or tubes of connected cells; they can originate from any germ layer. *Mesenchyme* comprises loosely packed cells derived from the mesoderm or neural crest. All organs consist of an epithelium and an associated mesenchyme, so epithelial–mesenchymal interactions are among the most important phenomena in development. Some examples are listed in Table 2–2.

Using the induction of cutaneous 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 the ectoderm) and a dermis (a mesenchymal tissue derived from the mesoderm). The chick epidermis signals the underlying dermal cells to form condensations (probably by secreting sonic hedgehog and transforming growth factor (TGF)-β2 proteins, which will be discussed further on), and the condensed dermal mesenchyme responds by secreting factors that cause the epidermis to form regionally specific cutaneous structures (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. Researchers can separate the embryonic epithelium and mesenchyme from each other and recombine them in different ways (Saunders et al., 1957), demonstrating that the dermal mesenchyme is responsible for the regional specificity of induction in the competent epidermal epithelium. The same type of 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.

Tooth patterning is another set of processes driven by region-specific epithelial induction. Classical epithelial–mesenchymal transplantation studies have shown that early dental epithelium induces the underlying mesenchymal cells to become odontogenic (Mina and Kollar, 1987). Mammoto et al. (2011) studied the mechanisms involved in the early patterning of the tooth both on the level of the organ and that of the single cell. They found that the early dental epithelium secretes Fgf8, which attracts underlying mesenchymal cells and causes them to migrate toward the source of Fgf8. The epithelium also secretes Sema3f, which repulses the migrating mesenchymal cells. At a certain distance from the early dental epithelium, the opposing epithelial morphogens cause the underlying mesenchymal cells to condense. This compaction of mesenchymal cells and the associated change in their cell shape alone is enough to induce mesenchymal expression of odontogenic-specific genes such as *Pax9, Msx1,* and *Bmp4*. This is mediated by the suppression of RhoA, a cytoskeletal signaling molecule that responds to mechanical cues. Thus early tooth patterning is achieved via chemical cues sent by the inductive epithelium to the

Table 2–2. Examples of Organs and Their Epithelial and Mesenchymal Components.

Organ	Epithelial Component	Mesenchymal Component
Cutaneous appendages (hair, sweat glands)	Epidermis (ectoderm)	Dermis (mesoderm)
Gut organs	Endodermal epithelium	Mesodermal mesenchyme
Respiratory organs	Endodermal epithelium	Mesodermal mesenchyme
Kidney	Ureteric bud epithelium (mesoderm)	Metanephrogenic (mesodermal) mesenchyme
Tooth	Jaw epithelium (ectoderm)	Neural crest (ectodermal) mesenchyme

responding mesenchyme, causing it to condense. The condensed mesenchyme then continues independently to switch on the cell fate–specific gene expression by mechanotransduction. Once mesenchyme has acquired its odontogenic fate, tooth patterning continues via reciprocal inductions between the dental epithelium and mesenchyme.

AN INTRODUCTION TO SIGNAL-TRANSDUCTION PATHWAYS

Paracrine Factors

How are the signals between inducer and responder transmitted? While studying the mechanisms of induction that produce the kidney tubules and the teeth, Grobstein (1956) and others 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* (as 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*, and the diffusible proteins are called *paracrine factors* or *growth and differentiation factors* (GDFs). We will consider paracrine interactions first and then return to juxtacrine interactions later in this 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. During the past decade, developmental biologists have discovered that the induction of numerous organs is actually effected by a relatively small set of paracrine factors. The embryo inherits a rather compact "tool kit" and uses many of the same proteins to construct the heart, kidneys, teeth, eyes, and other organs. Moreover, the same proteins are utilized 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 these paracrine factors can be grouped into four major families on the basis of their structures: the fibroblast growth factor (FGF) family, the Hedgehog family, the Wingless (Wnt) family, and the TGF-β superfamily. Although most of the paracrine factors are members of these four families, some have few or no close relatives. Factors such as epidermal growth factor, hepatocyte growth factor, neurotrophins, and stem cell factor are not in the families mentioned previously, but each plays important roles during development. In addition, numerous factors— erythropoietin, the cytokines, and the interleukins —are involved almost exclusively with developing blood cells.

Classically, paracrine signals have been considered to diffuse passively in extracellular space. Recent years have shown that in several organ systems signalling proteins are transported along transient specialized signalling filopodia called cytonemes. Cytonemes take environmental cues from the extracellular matrix, and allow dynamic and precise secretion of signalling proteins to affect growth, differentiation and pattern formation (reviewed in Kornberg and Roy, 2014; Roy and Kornberg, 2014). Cytonemes extend to make contact with target cells, and disperse signalling proteins at sites of these cell-cell contacts. In *Drosophila* wing disc cytonemes have been shown to be responsible for both short and long distance decapentaplegic (Dpp, a TGF-β family member) signalling; wing disc-associated tracheal branch air sac primordium has been shown to extend two types of cytonemes, one specific for Dpp signalling and another for FGF signalling (Roy et al., 2014).

Another mechanism for controlled transportation of signalling proteins for short and long distances outside cells is their secretion in extracellular vesicles called exosomes that contain cytosol from the secreting cell surrounded by a lipid bilayer. Exosomes traffic proteins, lipids, and RNA molecules. They are more stable than soluble protein in the extracellular matrix, and they make it easier to traffic hydrophobic proteins like active Wnt proteins. Wnt proteins, at least in part, have been shown to be secreted and transported to recipient cells via exosomes in *C. elegans, D. melanocaster*, and in human cells (Kolotuev et al., 2009; Gross et al., 2012; Gross and Boutros, 2013).

Paracrine Factor Signaling Pathways

We now turn to the molecules involved in the response to induction. These molecules include the receptors in the membrane of the responding cell, which bind the paracrine factor, and the cascade of interacting proteins, which transmit a signal through a pathway from the bound receptor to the nucleus. These pathways between the cell membrane and the genome are called *signal-transduction pathways*; we will outline some of the major ones here. They appear to be variations on a common and rather elegant theme (Fig. 2): each receptor spans the cell membrane and has an extracellular region, a transmembrane region, and a cytoplasmic region. When a ligand (the paracrine factor) binds its receptor in the extracellular region, it induces a conformational change in the receptor's structure. This shape change is transmitted through the membrane and changes the shape of the cytoplasmic domains. The conformational change in the cytoplasmic domains gives them enzymatic activity, usually a kinase activity that can use ATP to phosphorylate proteins, including the receptor molecule itself. The active receptor can now 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, which activates (or represses) a particular set of genes. There are numerous modifications of this theme; some of the most important of these pathways are outlined in Fig. 2–2.

The RTK–Mitogen-Activated Protein Kinase Pathway

Fibroblast Growth Factors

Almost two dozen distinct *FGF* genes are known in vertebrates, and they can generate hundreds of protein isoforms by varying their RNA splicing or initiation codons in different tissues (see Chapter 60). FGFs can activate a set of receptor tyrosine kinases (RTKs), the FGF receptors (FGFRs). When the FGFR binds an FGF (and only when it does so), the dormant kinase is activated and it phosphorylates certain proteins within the responding cell. The proteins are now activated and can perform new functions. FGFs are associated with several developmental functions, including angiogenesis (blood vessel formation), mesoderm formation, and axon extension. Although FGFs can often substitute for one another, their expression patterns give them separate functions. FGF2 is especially important in angiogenesis, and FGF8 is important for the development of the midbrain, eyes, and limbs (Crossley et al., 1996).

The RTK signal-transduction pathway was one of the first pathways to unite various areas of developmental biology. Researchers studying *Drosophila* eyes, nematode vulvae, and human cancers found that they were all studying homologous genes. The RTK–mitogen-activated protein kinase (MAPK) pathway begins at the cell surface, where an RTK binds its specific ligand. Ligands that bind to RTKs include the FGFs, epidermal growth factors, platelet-derived growth factors, and stem cell factor. Each RTK can bind only one or a small set of these ligands (stem cell factor, for instance, will bind to only one RTK, the Kit protein). 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. Heterozygous loss-of-function alleles of FGF receptors 2 and 3, which abolish this kinase activity, have been associated with lacrimo-auriculo-dento-digital (LADD) syndrome of lacrimal duct aplasia, deafness, digital anomalies, and small teeth (Rohmann et al., 2006). 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 adapter protein complex (Fig. 2–1B). Such a complex serves as a bridge that links the phosphorylated RTK to a powerful intracellular

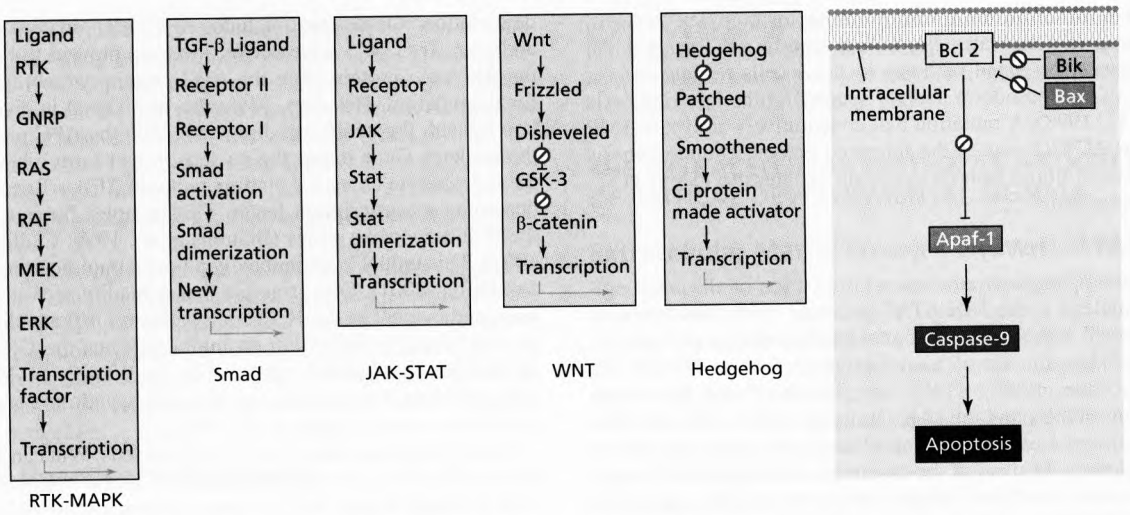

Fig. 2–2. Five of the major signal-transduction pathways through which signals on the cell surface are sent into the nucleus. (*A*) The receptor tyrosine kinase–mitogen-activated protein kinase (RTK-MAPK) pathway, (*B*) the Smad pathway used by transforming growth factor-β (TGF-β) superfamily proteins, (*C*) the JAK–STAT pathway, (*D*) the Wnt pathway, (*E*) the Hedgehog pathway, and (*F*) one of the apoptosis pathways used by mammalian neurons. *Abbreviations*: ERK, extracellular signal–regulated kinase; GNRP, guanine nucleotide–releasing protein; GSK, glycogen synthase kinase; JAK, Janus kinase; MEK, MAPK kinase; STAT, signal transducer and activator of transcription.

signaling system. While binding to the phosphorylated RTK through one of its cytoplasmic domains, the adapter protein also activates a G protein (see Chapter 60). Normally the G protein is in an inactive GDP-bound state. The activated receptor stimulates the adapter protein to activate the guanine nucleotide–releasing factor. 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. After delivering the signal, 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. Mutations in the genes encoding the small GTP-binding proteins and their regulators can lead to cancers or syndromes such as Aarskog-Scott syndrome, fibrous dysplasia, and neurofibromatosis 1 (see Chapters 180, 179, and 80). The Costello syndrome is caused by mutations in the HRAS small G protein (Aoki et al., 2005; Estep et al., 2006; Gripp et al., 2006), and Noonan syndrome is associated with missense mutations of the *PTPN11* gene (which encodes the SHP2 protein serving as part of the adapter complex in certain cell types (Tartaglia et al., 2001; see Chapter 81).

The active 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 MAPK kinase (MEK) protein by phosphorylating it. MEK is itself a kinase, which activates extracellular signal-regulated kinase (ERK) by phosphorylation, and phosphorylated ERK is a kinase that can enter the nucleus and phosphorylate certain transcription factors. This pathway is critical in numerous developmental processes, and mutations in the genes encoding MEK or RAF kinases can cause cardiofaciocutaneous syndrome (Rodriguez-Viciana et al., 2006; see Chapter 83).

In the migrating neural crest cells of humans and mice, the RTK pathway is important in activating the Mitf to produce the pigment cells (see Chapter 128). Mitf 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 this protein is not functional until it receives signals to become active. The clue to these signals lay in two mouse mutants whose phenotypes resemble those of mice homozygous for microphthalmia mutations. Like those mice, homozygous *White* mice and homozygous *Steel* mice are white because their pigment cells have failed to migrate. Perhaps 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 (Witte, 1990). This factor binds to and activates the Kit RTK encoded by the White gene (Spritz et al., 1992; Wu et al., 2000). The binding of the stem cell factor to the Kit RTK dimerizes the Kit protein, causing it to become phosphorylated. Phosphorylated Kit activates the pathway whereby phosphorylated ERK is able to phosphorylate Mitf (Hsu et al., 1997; Hemesath et al., 1998). Only the phosphorylated form of Mitf is able to bind the p300/cyclic AMP response element–binding protein (CBP) coactivator protein, which enables it to activate transcription of the genes encoding tyrosinase and other proteins of the melanin-formation pathway (Price et al., 1998).

The Smad Pathway

TGF-β Superfamily Proteins

There are more than 30 structurally related members of the TGF-β superfamily, and they regulate some of the most important interactions in development (see Chapter 31). 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 (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). The members of the BMP family were originally discovered by their ability to induce bone formation; hence they are the BMPs. Bone formation, however, is only one of their many functions; they regulate cell division, apoptosis (programmed cell death), cell migration, and differentiation (Hogan, 1996).

Members of the TFG-β superfamily activate members of the Smad family of transcription factors (Heldin et al., 1997). 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. (Researchers named the Smad proteins by eliding the names of the first identified members of this family: the *Caenorhabditis elegans* Sma protein and the *Drosophila* Mad protein.) Smads 1 and 5 are activated by the BMP family of TGF-β factors, while the receptors binding activin and the TGF-β family phosphorylate Smads 2 and 3. These phosphorylated Smads

bind to Smad 4 and form the transcription factor complex that will enter the nucleus. In vertebrates, the TGF-β superfamily ligand Nodal appears to activate the Smad pathway in those cells responsible for the formation of the mesoderm and for specifying the left–right axis (Nomura and Li, 1998). A mutation that constitutively activates BMP type I receptor ACVR1 causes the inherited ectopic chondrogenesis and osteogenesis of fibrodysplasia ossificans progressiva (Shore et al., 2006, see Chapter 56).

The JAK-STAT Pathway

Another important pathway transducing information on the cell membrane to the nucleus is the JAK-STAT pathway (with JAK standing for "Janus kinase" and STAT for "signal transducer and activator of transcription"). Here the set of transcription factors consists of the STAT proteins (Ihle, 1996). STATs are phosphorylated by certain RTKs, including FGFRs and the JAK family of tyrosine kinases. The JAK–STAT pathway is extremely important in the differentiation of blood cells and the activation of the casein gene during milk production. In casein production, for instance, the endocrine factor prolactin binds to the extracellular regions of prolactin receptors, causing them to dimerize. A JAK protein kinase is bound to each of the receptors (in their respective cytoplasmic regions), and these JAK proteins are now brought together, where they can phosphorylate the receptors at several sites. The receptors are now activated and have their own protein kinase activity. Therefore the JAK proteins convert a receptor into an RTK. The activated receptors can now phosphorylate particular inactive STATs and cause them to dimerize. These dimers are the active form of the STAT transcription factors; they are translocated into the nucleus, where they bind to specific regions of DNA. In this case, they bind to the upstream promoter elements of the casein gene, causing it to be transcribed.

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 genetic lesion resides in the gene encoding FGFR3, a receptor expressed in the cartilage precursor cells in the growth plates of the long bones (Rousseau et al., 1994; Shiang et al., 1994). Normally the FGFR3 protein is activated by an FGF; it signals the chondrocytes to stop dividing and begin differentiating into cartilage. This signal is mediated by the STAT1 protein, which is phosphorylated by the activated FGFR3 and then translocated into the nucleus. Inside the nucleus, this transcription factor activates the genes encoding a cell cycle inhibitor, the p21 protein (Su et al., 1997). The mutations causing thanatophoric dwarfism result in 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, 1996).

The Wnt–β-Catenin Pathway

Wnt-Family Proteins

The Wnts constitute a family of cysteine-rich glycoproteins. There are at least 15 members of this family in vertebrates (see Chapter 36). Their name comes from fusing the name of the *Drosophila* segment polarity gene *wingless* with int-1, the first cloned integration site for mouse mammary tumor virus. Although 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 (McMahon and Bradley, 1990; Stern et al., 1995). Wnt proteins are also critical in establishing the polarity of insect and vertebrate limbs, and they are used in several steps of urogenital system development.

Members of the Wnt family of paracrine factors interact with transmembrane receptors of the Frizzled family (see Chapter 36). In the classical Wnt pathway, the binding of Wnt by the Frizzled protein causes the Frizzled protein to activate the Dishevelled protein. Once the Dishevelled protein is activated, it inhibits the activity of the glycogen synthase kinase-3 (GSK-3) enzyme. GSK-3 is part of a protein degradation complex that includes APC (adenomatous polyposis coli) and axin. If GSK-3 were active, it 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 given and GSK-3 is inhibited, β-catenin can dissociate from the APC protein and enter the nucleus. Once inside the nucleus, it can form a heterodimer with the lymphocyte enhancer-binding factor (LEF) or T-cell factor (TCF), becoming a transcription factor. This complex binds to and activates the Wnt-responsive genes (Behrens et al., 1996; Cadigan and Nusse, 1997). This model is undoubtedly an oversimplification (see McEwen and Peifer, 2001). One principle that is readily seen in the Wnt pathway (and evident in the Hedgehog pathway) is that activation is often accomplished by inhibiting an inhibitor. Thus the GSK-3 protein is an inhibitor that is itself repressed by the Wnt signal. Mutation of the gene encoding Axin-2 causes an intriguing syndrome of tooth agenesis and colon cancer (Lammi et al., 2004).

In addition to sending signals to the nucleus, Wnt can affect the actin and microtubular cytoskeleton. Here, the Dishevelled protein interacts with a protein kinase that initiates a cascade that will phosphorylate cytoskeletal proteins and thereby alter cell shape, cell polarity (where the upper and lower portions of the cell differ), or motility (Witte et al., 2010; Sepich et al., 2011; Ho et al., 2012).

The Hedgehog Pathway

Hedgehog Proteins

The Hedgehog proteins constitute a family of paracrine factors that are often used by the embryo to induce particular cell types and to create boundaries between tissues (see Chapter 27). Vertebrates have at least 3 homologues of the *Drosophila* hedgehog gene: *Sonic hedgehog* (*Shh*), *Desert hedgehog* (*Dhh*), and *Indian hedgehog* (*Ihh*). Desert hedgehog is expressed in the Sertoli cells of the testes, and mice homozygous for a null allele of *Dhh* exhibit defective spermatogenesis. Indian hedgehog is expressed in the gut and cartilage and is important in postnatal bone growth.

Sonic hedgehog is the most widely used of the three vertebrate homologues. Made by the notochord, it is processed so that only the amino-terminal two-thirds of the molecule is secreted. This peptide is responsible for patterning the neural tube such that motor neurons are formed from the ventral neurons and sensory neurons are formed from the dorsal neurons (Yamada et al., 1993). Sonic hedgehog is also responsible for patterning the somites so that the portion of the somite closest to the notochord becomes the cartilage of the spine (Fan and Tessier-Lavigne, 1994; Johnson et al., 1994). Sonic hedgehog is critical for the formation and maintenance of the facial midline, and several mutations of the *Sonic hedgehog* gene (or its receptor) cause holoprosencephaly (Maity et al., 2005). Sonic hedgehog mediates the formation of the left-right axis in several vertebrates, initiates the anterior-posterior axis in limbs, induces the regionally specific differentiation of the digestive tube, and induces hair formation. Sonic hedgehog often works with other paracrine factors, such as Wnt and FGF proteins. In the developing tooth, Sonic hedgehog, FGF4, and other paracrine factors are concentrated in the region where cell interactions create the cusps of the teeth (Vaahtokari et al., 1996).

Members of the Hedgehog protein 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 (Smo), and the Patched protein prevents Smo from functioning. In the absence of the Hedgehog protein binding to the Patched protein, Smo, inhibited by Patched, is inactive and the Cubitus interruptus (Ci) protein (Glis 1to 3 in vertebrates) is tethered to the primary cilium of the responding cell. The primary cilium is a structure that protrudes from the cell surface and is formed from microtubules. While it is on the primary cilium, Ci is cleaved in such a way that a portion of it enters the nucleus and acts as a transcriptional repressor. This portion of the Ci protein binds to the promoters and enhancers of particular genes and acts as an inhibitor of transcription. When Hedgehog binds to Patched, the Patched protein's shape is altered such that it no longer inhibits Smo. Smo accumulates in the primary cilium

and complexes with Evc2, a ciliary protein; this restricts Smo to localizing to a distinct ciliary compartment, the Evc zone. The localization of the Smo-Evc2 complex leads to the dissociation of inhibitors Suppressor of Fused (SuFu) and protein kinase A from Ci. Active Smo signaling helps to transport Ci and SuFu to the tip of the primary cilium. This, in turn, allows Ci to dissociate from SuFu, and the intact Ci protein can now enter the nucleus, where it acts as a transcriptional activator of the same genes it used to repress (Aza-Blanc et al., 1997; Dorn et al. 2012).

The Hedgehog pathway is extremely important in limb and neural differentiation in vertebrates (see Chapters 28 to 35). Here the homologues of Ci are the Gli proteins. Both the transport of the Hedgehog protein and its reception by the target cell require cholesterol. Therefore mutations involving cholesterol metabolism and teratogens that block cholesterol synthesis can cause the same spectrum of defects as mutations in Sonic hedgehog.

Juxtaposed Ligands and Receptors

The Notch Pathway

Although 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 protein in their membranes activate neighboring cells that contain the Notch protein in cell membranes (see Chapter 72). Notch extends through the cell membrane; 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 it to be cut by the presenilin-1 protease. The cleaved portion enters the nucleus and binds to a dormant transcription factor of the CSL (CBF1, Suppressor of Hairless, Lag-1) family. When bound to the Notch protein, the CSL transcription factors activate their target genes (Lecourtois and Schweisguth, 1998; Schroeder et al., 1998; Struhl and Adachi, 1998).

Notch proteins 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).

Eph Receptors and Ephrins

Eph receptors are the largest subfamily of RTK receptors. Their ligands, ephrins, form a family of related proteins. Both the receptors and ligands are bound to plasma membrane. Eph/ephrin signaling is bidirectional: both receptors and ligands can induce signal transduction. Eph receptor activation of ephrin ligand is called *forward signaling*, and ligand-activation of Eph receptor is called *reverse signaling*. Eph receptors and ephrins can interact between two opposing cells, termed *trans* interactions, or it can happen within one cell, termed *cis* interactions (Arvanitis and Davy, 2008).

Eph receptors and ephrins are expressed widely in the developing embryo, and they play important physiological roles in the adult as well. What makes their biological functions wide and varied is the ability of Eph/ephrins to interact with other cell surface receptors, adhesion proteins, and channels and pores such as gap junctions and the ion channels found in synapses (Arvanitis and Davy, 2008).

In *Xenopus* FGF and EphrinB1 signaling regulated the migration of retinal progenitor cells into the eye domain within the anterior neural plate. FGF signaling represses the cell migration into the eye domain, and EphrinB1 reverse signaling promotes it (Moore et al., 2004). Moore and colleagues also show that *Fgfr2* and *EphrinB1* are coexpressed at the anterolateral borders of the eye domain, and that EphrinB1 reverse signaling can rescue the FGF repression of retinal cell fate caused, for instance, by constitutively active *Fgfr2*. This Fgf and Ephrin signaling may be due to direct interactions, because Fgfr can associate with the cytoplasmic tail of the EphrinB1 in *Xenopus* and induce phosphorylation of EphrinB1 (Chong et al., 2000). The progenitor cell movements are critical for retinal cell fate as the ectopic expression of eye-specifying transcription factor *Pax6* in blastomere cells; these are not normally destined to become retinal cells do not adopt a retinal cell fate unless they have migrated into the eye domain (Kenyon et al., 2001; Moody, 2004).

THE EXTRACELLULAR MATRIX AS A SOURCE OF CRITICAL DEVELOPMENTAL SIGNALS

Extracellular Matrix Proteins and Functions

The extracellular matrix consists of macromolecules secreted by cells into their immediate environment (see Chapter 163). 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. These large glycoproteins are responsible for organizing the matrix and cells into an ordered structure. Fibronectin plays an important role in cell migration. The "roads" over which certain migrating cells travel are paved with this protein. Fibronectin paths lead germ cells to the gonads and lead 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 and type IV collagen are major components of a type of extracellular matrix called the *basal lamina*, which is characteristic of the closely knit sheets that surround epithelial tissue. 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).

Bissell and colleagues (1982; Martins-Green and Bissell, 1995) have shown that the extracellular matrix is also capable of inducing specific gene expression in developing tissues—especially those of the liver, testis, and mammary gland—in which the induction of specific transcription factors depends on cell–substrate binding (Liu et al., 1991; Streuli et al., 1991; Notenboom et al., 1996). Often the presence of bound integrin (the cell membrane receptor for fibronectin and other extracellular matrix molecules) 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 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 matrix, they shrivel up and die. Endo et al. (2012) have shown that direct interactions between the extracellular matrix protein anosmin and Fgf8, Bmp5 and Wnt3a in anterior neural folds ensure correct local doses of these paracrine factors to form the cranial neural crest and thus ultimately normal craniofacial morphogenesis. In humans, mutations in *anosmin* cause X-linked Kallmann syndrome which includes anosmia and craniofacial defects (Legouis et al., 1991)

The extracellular matrix may play especially important roles in mediating the branching of parenchymal organs. Lin and colleagues (2001) showed that murine kidney and lung mesenchymes induce collagen XVIII in different places on the epithelium of murine kidneys and lungs and that this expression is predictive of the branching pattern of the epithelia.

Recent experiments (Cota and Davidson, 2015) demonstrate that the extracellular matrix can be critical in creating divergent cell types from a common precursor cell. In the tunicate *Ciona*, Fgf signals appear to tell the matrix-bound cell of a dividing pair to become a heart precursor, but its sister cell is destined to become a pharynx cell. It appears that the adherent cell membrane is able to keep and trap Fgf receptors via integrin and caveolin. In this way, the matrix-bound cell is able to respond to the Fgf, while the non-adherent sister cell is not.

CELL DEATH PATHWAYS

Programmed cell death, or apoptosis, is a normal part of development. 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 of 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.) While a human individual is still within the uterus, he or she is constantly making and destroying cells, generating about 3 times as many neurons as he or she eventually ends up with at birth.

> 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. (Thomas, 1992)

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; Rodrigez et al., 1997; Roberts and Miller, 1998). Apoptosis prunes away unneeded structures, controls the number of cells in particular tissues, and sculpts complex organs. Different tissues use different signals for apoptosis. One of the signals often used in vertebrates is BMP4. Some tissues, such as connective tissue, respond to BMP4 by differentiating into bone. Others, such as the surface 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 GDFs are secreted by the enamel knot. After the cusp has grown, the enamel knot synthesizes BMP4 and shuts itself down by apoptosis (Vaahtokari et al., 1996).

Cells are programmed to die in several tissues; they 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 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 (Bittorf et al., 2000; Socolovsky et al., 2001). In this way the amount of erythropoietin that is 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*. It was found that the cell death abnormal (ced) proteins encoded by the *ced-3* and *ced-4* genes were essential for apoptosis; however, in the cells that did not undergo apoptosis, those genes were turned off by the product of the *ced-9* gene (Hengartner et al., 1992). The CED-4 protein is a protease activating factor that activates CED-3, a protease that initiates the destruction of the cell. Mutations that inactivate the CED-9 protein cause numerous cells that would normally survive to activate their *ced-3* and *ced-4* genes and die. This leads to the death of the entire embryo. Conversely, gain-of-function mutations of *ced-9* cause the CED-9 protein to be made in cells that would otherwise die. 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 and that those cells that survive are rescued by activation of the *ced-9* gene.

The CED-3 and CED-4 proteins form 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 the CED-3 or CED-4 protein or inactivate the CED-9 molecules. In mammals, the homologues of the CED-9 protein are members of the Bcl-2 family of genes. This family includes *BCL-2, Bcl-X*, and similar genes. The functional similarities are so strong that if an active human *BCL-2* gene is placed into *C. elegans* embryos, it prevents normally occurring cell deaths in the nematode embryos (Vaux et al., 1992). In vertebrate red blood cell development (mentioned previously), the Stat5 transcription factor activated by erythropoietin functions by binding to the promoter of the *Bcl-X* gene, where it activates the synthesis of that antiapoptosis protein (Socolovsky et al., 1999).

The mammalian homologue of CED-4 is called Apaf-1 (apoptotic protease activating factor-1); it participates in the cytochrome *C*–dependent activation of the mammalian CED-3 homologues caspase-9 and caspase-3 (Shaham and Horvitz, 1996; Cecconi et al., 1998; Yoshida et al., 1998). Activation of the caspases causes autodigestion of the cell. Caspases are strong proteases, and they digest the cell from within. The cellular proteins are cleaved and the DNA is fragmented.

Although apoptosis-deficient nematodes deficient for CED-4 are viable (despite their 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 (Kuida et al., 1996, 1998; Jacobson et al., 1997). Mice homozygous for targeted deletions of Apaf-1 have severe craniofacial abnormalities, brain overgrowth, and webbing between their toes.

In mammals there is more than one pathway to apoptosis. Apoptosis of the lymphocytes, for instance, is not affected by the deletion of Apaf-1 or caspase-9 and works by a separate pathway initiated by the CD95 protein. Different caspases may function in different cell types to mediate the apoptotic signals (Hakem et al., 1998; Kuida et al., 1998). One of the most interesting involves the "death domain," containing receptors of the tumor necrosis factor (TNF) family. These receptors can induce apoptosis in several cell systems, and they appear to accomplish this by blocking the antiapoptosis signals sent by other factors. It is likely that the death domain binds a phosphatase that cleaves the phosphates from the RTKs, which would be activated by the antiapoptotic signal (Daigle et al., 2002). This prevents their activation and allows apoptosis to commence.

Just as developmental signals (e.g., BMP4) can be used by some cells for apoptosis, so the death domain receptors can, in some instances, be used for nonapoptotic development. One of the developmentally important TNF receptors with a death domain is Edar, a protein required for the development of hair, teeth, and other cutaneous appendages. Mutations of this or of its ligand, Eda, cause hypohidrotic epidermal dysplasia, a syndrome characterized by lack of sweat glands, sparse hair, and poorly formed teeth. An identical syndrome is also produced by a deficiency of the adapter protein that binds the death domain of this receptor. In this instance, instead of producing cell death, activation of the receptor enables continued development (Headon et al., 2001).

THE NATURE OF GENETIC SYNDROMES

Pleiotropy

Research on the expression patterns of transcription factors and paracrine factors has suggested mechanisms that explain some of the genetic syndromes wherein mutations at a single genetic locus can cause numerous different malformations. The production of several conditions by mutations at one locus is called *pleiotropy*. For instance, in humans and mice, heterozygosity for MITF deficiency causes a condition that involves iris defects, pigmentation abnormalities, deafness, and an inability to produce the normal number of mast cells (see Fig. 2–1B). The skin pigment, the iris of the eye, the inner ear tissue, and the mast cells of the immune system are not related to one another

in such a way that the absence of one would produce the absence of the others. Rather, all 4 parts of the body independently use the MITF protein as a transcription factor. This type of pleiotropy has been called *mosaic pleiotropy* because the relevant organ systems are separately affected by the abnormal gene function.

Whereas the eye pigment, body pigment, and mast cell features of MITF deficiency are separate events, other features of the syndrome are not. For instance, the failure of MITF expression in the pigmented retina prevents this structure from fully differentiating. This, in turn, causes a malformation of the choroid fissure of the eye, resulting in drainage of the vitreous humor fluid. Without this fluid, the eye fails to enlarge (hence the name *microphthalmia*, which means "small eye"). This phenomenon, in which several developing tissues are affected by the mutation even though they do not express the mutated gene, is called *relational pleiotropy* (see Grüneberg, 1938, 1962).

Genetic Heterogeneity

Another important feature of syndromes is that mutations in different genes can produce the same phenotype. If the genes are part of the same signal-transduction pathway, a mutation in any of them can give a similar result. The phenomenon whereby mutations in different genes produce similar phenotypes is called genetic heterogeneity. For example, cyclopia can be produced by mutations in the *Sonic hedgehog* gene or by mutations in cholesterol synthesis genes. Since they are in the same pathway, mutations in one gene generate a phenotype similar or identical to mutations in the other genes. Similarly, as we saw earlier, mutations in *EDA* (ligand), *EDAR* (receptor) or *EDARADD* (adapter protein) gave the same phenotype of hypohidrotic ectodermal dysplasia (Headon et al., 2001).

Mechanisms of Dominance

Whether a syndrome is dominant or recessive can now be explained at the molecular level. First, many syndromes are referred to as "dominant" only because the homozygous condition is lethal to the embryo and the fetus never survives. Therefore the homozygous condition never exists. Second, there are at least 4 ways of achieving a dominant phenotype.

The first mechanism of dominance is haploinsufficiency. This merely means that one copy of the gene (the haploid condition) is not enough to produce the required amount of product for normal development. For example, individuals with Waardenburg syndrome type II have roughly half the wild-type amount of MITF. This is not enough for full pigment cell proliferation, mast cell differentiation, or inner ear development. Thus an aberrant phenotype results when only 1 of the 2 copies of this gene is absent or nonfunctional.

The second mechanism of dominance is gain-of-function mutations. As mentioned earlier, thanatophoric dysplasia (as well as milder forms of dwarfism, such as achondroplasia) results from a mutation causing the FGFR to be constitutively active. This activity is enough to cause an anomalous phenotype to develop.

The third mechanism of dominance is a dominant-negative allele. This situation can occur when the active form of the protein is a multimer and all proteins of the multimer have to be of the wild type for the protein to function. A dominant-negative allele is the cause of Marfan syndrome, a disorder of the extracellular matrix. This syndrome is associated with anomalies of the joints and connective tissues, not all of which are necessarily disadvantageous. Increased height, disproportionately long limbs and digits, and mild-to-moderate joint laxity are characteristic. However, patients with Marfan syndrome also experience vertebral column deformities, myopia, loose lenses, and (most importantly) aortic problems that may lead to an aneurysm (bursting of the aorta) later in life. The mutation is in the gene for fibrillin, a secreted glycoprotein that forms multimeric microfibrils in elastic connective tissue. The presence of even small amounts of mutant fibrillin prohibits the association of wild-type fibrillin with microfibrils. Eldadah and colleagues (1995) have shown that when a mutant human gene for fibrillin is transfected into fibroblast cells that already contain two wild-type genes, incorporation of fibrillin into the matrix is inhibited.

The fourth mechanism of dominance involves subunit interactions wherein the dimer made from the products of 2 different alleles is superior in function to dimers made exclusively of the product of either allele (Trehan and Gill, 2002).

MODULARITY AND CONTEXTUALITY

During the development of mammals, there are domains set off by the expression of transcription factors. These regions can interact, thereby activating the expression or function of another set of transcription factors. These can further subdivide the domains, or they can initiate the expression of those batteries of genes that cause the differentiated proteins of the particular cell type to emerge. This concept is known as *modularity*. Development occurs through a series of discrete and interacting modules (Riedl, 1978; Gilbert et al., 1996; Raff, 1996; Davidson, 2001). In development, such modules include physical modules (e.g., hair follicles and teeth), morphogenetic fields (e.g., those described for the limb or eye), and physical structures for which there are no adult counterparts (e.g., rhombomeres). Modular units allow certain parts of the body to change without interfering with the functions of other parts.

These fields often provide the genes and proteins with their context. For instance, as already mentioned, BMP4 can be a signal for bone development (in somites), apoptosis (in the neural and derivatives), epidermis formation (in the induction of the ectoderm), or lens formation (within the eye). What it does depends on the field in which it is expressed. Unlike the genes for globin or chymotrypsin, the genes expressed early in development are often used for multiple functions. In some instances these functions can depend on very different properties of the molecule. β-Catenin can play a role on the cell membrane as part of an adhesion complex, or it can play a role in the nucleus as a transcription factor. Similarly a protein that functions as an enolase or alcohol dehydrogenase enzyme in the liver can function as a structural crystallin protein in the lens (Piatigorsky and Wistow, 1991). Developmental genes work within specific contexts.

Phenotypic Heterogeneity

These contexts can also determine the effect of a particular gene. It is not uncommon in the context of clinical genetics to identify a mutant allele that produces a mildly abnormal phenotype in one generation and a more abnormal phenotype in another generation. A mutant gene that produces limblessness in one generation can produce only a mild thumb abnormality in the next (Freire-Maia, 1975). Indeed, by following the phenotypes produced in different members of the same family, one can see that the same gene can produce different phenotypes depending on the other genes that are present (Wolf, 1995; Nijhout and Paulsen, 1997). There are even cases where the phenotype depends on whether the mutant allele is passed through the father or the mother.

MORPHOGENETIC PROCESSES

The Morphogenetic Repertoire

Morphogenesis involves changes in cell behavior. There are 2 main groups 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 each other and operate as individual units. Morphogenesis is brought about through a limited repertoire of cellular processes in these 2 classes of cells: (1) the direction and number of cell divisions, (2) changes in cell shape, (3) cell movement, (4) cell growth, (5) cell death, and (6) changes in the composition of the cell membrane and extracellular matrix. How these processes are accomplished can differ between mesenchymal and epithelial cells (Table 2–3; see Larsen, 1997).

Three important morphogenetic events—the orientation of cell division, the orientation of cell polarity, and cell migration—are regulated through the cytoskeleton. The orientation and number of cell divisions are tightly regulated. Given the similarity of facial appearance within a family, it seems that such regulation is under genetic control.

Table 2–3. The Morphogenetic Repertoire.

Cell Type	Property	Mechanism	Example
Mesenchyme (and nerves)	Movement	Contact guidance	Amphibian gastrulation
		Haptotaxis	Cell movement in vitro
			Retinal cells mapping to tectum
		Contact inhibition/ space	NC invasion of cornea
		Cell repulsion	NC blockage in posterior somite
		Chemotaxis	Neuronal migration in spinal cord
	Apoptosis		Loss of müllerian duct
	Condensation	Migration	Ganglion formation
		Adhesion	Dermal condensations
		Growth	Theoretical
		Traction	Cell aggregation in vitro
		Loss of ECM	Limb precartilage condensations
	Stability		
Epithelia	Polarization		Nephron lumen formation
	Forming folds	Buckling	Ciliary body formation
		Bending of sheet	Neural fold formation
	Forming tubes	Hollowing	Fish neural tube
		Growth	Blood vessels
	Tube branching	Signaling	Trachea
	Cell sorting		*Drosophila* egg chamber formation
	Movement	Passive/growth	Serosal migration in honeybee
			Surface ectoderm
		Rearrangement	Convergent extension
		Cessation	Completion of wound healing
	Tissue integrity		Retinal stability
	Growth		Throughout body

ECM, extracellular matrix; NC, neural cell

Source: J. Bard, personal communication.

The direction of cell divisions is controlled by the orientation of the mitotic spindle. This, in turn, is regulated by the cytoskeleton of the cell, especially by the dynein-rich cortex (O'Connell and Wang, 2000; Dujardin and Vallee, 2002). The positioning of the cytoskeleton can be effected by cell-cell contacts or by paracrine factors such as Wnt proteins (Goldstein, 2000; Le Borgne et al., 2002). Both appear to work by causing changes in the cytoskeleton.

Changes in cell shape allow for morphogenesis in several ways. Many morphogenetic processes, such as neurulation, are caused by changes in the shape of cells. In neurulation, this is especially important for the formation of the neural tube (Smith and Schoenwolf, 1997; Zolessi and Arruti, 2001). In mammals, changes in cell shape are also critical for the formation and polarity of the trophectoderm (Kidder, 2002). Such changes are also mediated through the cytoskeleton.

There are several mechanisms that enable cells to migrate from one region of the body to another. These mechanisms involve interactions of the cell surfaces with molecules that give cues as to when to start migrating, where to go, and when to cease migration. Migration is dependent on both the actin cytoskeleton and the ATPases (e.g., dynein) that drive it. Defects in the regulation of these systems, such as in lissencephaly-1, where the regulation of dynein in certain neuronal precursors is defective (Vallee et al., 2001), result in defective migration. There are two major modes of cell migration, and they are often combined. One mode is to follow a gradient of chemotactic molecules to its source. The other mode is to follow a particular substrate pathway. In addition to attractive signals, repulsive signals also can occur.

Chemotaxis is defined as cellular locomotion directed in response to a concentration gradient of a chemical factor in solution. Cells sense the chemical and migrate toward higher concentrations of it until they reach the source secreting it. In the vertebrate lung, brain, and limb, FGFs function as chemotactic proteins (Park et al., 1998; Li and Muneoka, 1999; Kubota and Ito, 2000). Failure of chemotaxis is found in several systems, especially those affecting the cytoskeleton of neural cells and lymphocytes.

Gradients do not have to be in solution. An adhesive molecule could be present in increasing amounts along an extracellular matrix. A cell that was constantly making and breaking adhesions with such a molecule would move from a region of low concentration to an area where that adhesive molecule was more highly concentrated. Such a phenomenon is called *haptotaxis* (Curtis, 1969). Migration of the pronephric duct cells in salamanders is regulated by haptotaxis (Zackson and Steinberg, 1987; Drawbridge et al., 1995). Moreover, certain human genetic conditions appear to have bases in the haptotactic mode of migration. In Kallmann syndrome, for instance, the protein anosmin is absent. This protein plays a key role in the migration of gonadotropin-releasing hormone neurons and olfactory nerves to the hypothalamus, and it is thought to be part of the extracellular matrix (Soussi-Yanicostas et al., 1996).

RANDOMNESS AND CHANCE IN MORPHOGENESIS

There are numerous sources of randomness in the human phenotype. In addition to somatic mutation and recombination, the major sources of chance are (1) X-chromosome inactivation, (2) stochastic interactions, and (3) environmental induction.

X-Chromosome Inactivation

In female mammals, including humans, one X chromosome in each cell is inactive while the other X chromosome is active. Very early in the development of human females, both X chromosomes are active; but as development proceeds, one X chromosome is turned off in each cell. Moreover, this inactivation is random. In some cells, the paternally derived X chromosome is inactivated; in other cells, the maternally derived X chromosome is shut off. This process is irreversible. Once an X chromosome has been inactivated, the same X chromosome is inactivated in all of that cell's progeny. As X inactivation happens relatively early in development, an entire region of cells derived from a single cell may have the same X chromosome inactivated. Thus, all tissues in female mammals are mosaics of 2 cell types (see Migeon, 1994).

If one of the X chromosomes contains a mutant allele and the other does not, the pattern of inactivation can produce different phenotypes. There are several cases of monozygous twins who are heterozygous for an X-linked form of muscular dystrophy. Most heterozygous women do not express any symptoms because the cells expressing the wild-type allele can compensate for the cells expressing the mutant allele. However, if by chance the wild-type allele is on the inactivated X chromosome in a large proportion of a woman's muscle cells, she will manifest the disease. There have been several instances where one girl showed the symptoms of the disease whereas her monozygous twin sister did not (Pena et al., 1987; Norman and Harper, 1989; Richards et al., 1990; Tremblay et al., 1993). Similarly, there are cases where monozygous female twins are discordant with respect to color blindness or Hunter syndrome owing to X-chromosome inactivation (Jorgensen et al., 1992; Winchester et al., 1992).

Stochastic Variation

Many events in the body are the products of both chance and necessity. The vascularization of the body differs even between identical twins, as do the intimate details of neural connection, lymphocyte repertoire, and iris pattern (see Daugman and Downing 2001). Even in relatively simple animals that are genetically identical and are given the same homogeneous environment, differences can be seen. The Bristol N2 strain of *C. elegans* has an invariant cell lineage (the cell divisions that occur between the fertilized egg and the adult are largely identical and

always produce the same set of tissues), an invariant nervous system whose 302 neurons have reproducible synaptic connectivity, and an invariant genotype. Moreover, this strain of *C. elegans* has a repertoire of behaviors that it performs in a very limited environment, a flat agar surface supplied with a uniform pad of identical bacteria. However, this behavior is not uniform. Jorgensen et al. (1992) isolated mutants most of which were seen to lie straight in a paralyzed manner. However, a fraction of the worms consistently took on a quite different "curly" posture. Such differences in a cloned population might be caused by slight differences in neural connectivity, which could be caused by stochastic developmental effects (see Schnabel, 1997). The nervous system seems to encourage such differences by the "winner take all" mechanisms caused by Hebbian rules of neuronal connections.

Environmental Determination of Phenotype

In addition to its role in immunocyte differentiation, the environment plays other significant roles in phenotype production. Obviously starvation or overeating will change an individual's phenotype, and there are likely to be genes that respond to dietary factors and produce diseases such as diabetes and coronary artery disease. These are diseases wherein genes interact with environmental conditions to create the pathological states. There are some genetic conditions that can be completely abolished by dietary supplementation. Foremost among these is gulonolactone oxidase deficiency (hypoascorbemia, OMIM 240400), a mutation in the gulonolactone oxidase gene on the short arm of chromosome 8, which causes a syndrome that produces death in childhood due to connective tissue malfunction. Gulonic acid oxidase is the final enzyme in the pathway leading to ascorbic acid, and although most mammals have this enzyme and can synthesize vitamin C, our genes for it are mutated and we cannot make this necessary compound. Without vitamin C replacement therapy from the environment, each of us would die.

Environment can also be extremely important in modulating the effects of mutant alleles. The fetuses most at risk for neural tube defects appear to be those with mutations in genes associated with folate metabolism (Whitehead et al., 1995; De Marco et al., 2000). Folate is a critical substrate for the methylation of homocysteine to methionine, but the mechanisms by which folate deficiency interferes with neural tube closure are not known. Smith-Lemli-Opitz syndrome and phenylketonuria are 2 other conditions wherein the environment interacts with genetics to produce the phenotype. In Smith-Lemli-Opitz syndrome (Chapter 28), dietary cholesterol can offset the mutant alleles that prevent proper cholesterol synthesis. In phenylketonuria, the behavioral and cognitive deficiencies associated with this syndrome can be ameliorated by dietary restriction of phenylalanine.

It is also possible that interactions between the environment and our genome control some part of our facial phenotype. Physical stress from the environment is needed to produce bones such as the mammalian patella and the bird fibular crest (Müller and Steicher, 1989; Wu, 1996). Corruccini (1984) and Varrela (1992) have speculated that the reason that nearly a quarter of our population needs orthodontic appliances is that our children's mild-textured diet causes the lower jaw to develop abnormally. Increased chewing causes tension that stimulates growth of the mandible bone and muscle (Kiliardis et al., 1985; Weijs and Hillen, 1986). The placement of young primates on a soft diet will cause malocclusions in their jaws, similar to that seen in humans (Corruccini and Beecher, 1982, 1984).

One of the most important changes that the environment can effect in mammalian development concerns DNA methylation. The "Barker hypothesis" (Barker, 1995) postulates that certain anatomical and physiological parameters get "programmed" during embryonic and fetal development and that changes in nutrition during this time can produce permanent changes in the pattern of metabolic activity. These changes can predispose the adult to particular diseases. Specifically epidemiological studies have suggested that the adult offspring of mothers who experienced protein deprivation during certain months of pregnancy (owing to wars, famines, or migrations) were at high risk for developing hypertension, heart disease, and diabetes as adults.

Although some of these changes may be anatomical (a low-protein diet might not allow the kidney to construct as many nephrons as it would build on a protein-rich diet), other changes may involve methylation differences that "set" the metabolic levels of the adult. Lillycrop and colleagues (2005) have shown that rats born of mothers having a low-protein diet had a different pattern of liver gene methylation than did the offspring of mothers fed a diet with a normal amount of protein and that these differences in methylation changed the metabolic profile of the rat's liver. For instance, the methylation of the promoter region of the *PPARα* gene (a gene that is critical in the regulation of carbohydrate and lipid metabolism) is 20% lower in those rats fed protein-restricted diets and its transcriptional activity is 10-fold greater. Moreover, the difference between these methylation patterns could be abolished by including folic acid in the protein-restricted diet. Thus the difference in methylation results from changes in folate metabolism caused by the limited amount of protein available to the fetus. It appears, then, that prenatal nutrition can induce long-lasting gene-specific alterations in transcriptional activity and metabolism. Similarly, genetically identical mice, each having the *viable Agouti* allele, had different patterns of pigmentation and obesity depending on the methyl supplementation of their mother's diet during pregnancy (Waterland and Jirtle, 2003). This was found to be regulated by the methylation of a regulatory region of the *viable Agouti* gene.

DNA methylation differences arise during the lifetimes of monozygotic twins. Moreover, such differences increase with time and are reflected in different gene expression patterns (Fraga et al., 2005). Thus different phenotypes can arise from identical genotypes through environmental interactions. This was demonstrated experimentally by depriving rats of maternal care during the first week of postnatal life. Those rats that had low maternal care were more anxious as adults. This was found to be correlated with the failure of the glucocorticoid receptor to become expressed in the hippocampus of these rats. This hippocampal expression failure, in turn, was found to be caused by the methylation of a particular cytosine in the regulatory region of the glucocorticoid receptor gene. This methylation prevented the binding of the Egr-1 transcription factor found in the brain and the prevention of transcription from these brain cells (Weaver et al., 2004, 2005).

The production of phenotype from genotype is regulated at numerous gene expression levels: at the levels of gene transcription, mRNA processing, mRNA translation, and posttranslational modification. It is further controlled by cell-cell and cell-substrate interactions and by environmental influences. At all of these levels, the vectors of regulation work both ways. A cell's fate is determined both by the gene expression within it and the community of cells in which it resides. Even the environment can alter patterns of gene expression; in the production of the human phenotype, experience is added to endowment (Childs, 1999).

ACKNOWLEDGMENTS

Support for this work came from grants from Swarthmore College and the Academy of Finland.

References

Complete references for this chapter are available online at: www.oxfordmedicine.com/inbornerrorsofdevelopment3e