

A Brief History



Garden pea flowers. Flower color (purple or white) was one of the traits Mendel studied in his classic examination of inheritance in the pea plant. © Shape'n'colour/Alamy, RF.

What is molecular biology? The term has more than one definition. Some define it very broadly as the attempt to understand biological phenomena in molecular terms. But this definition makes molecular biology difficult to distinguish from another well-known discipline, biochemistry. Another definition is more restrictive and therefore more useful: the study of gene structure and function at the molecular level. This attempt to explain genes and their activities in molecular terms is the subject matter of this book.

Molecular biology grew out of the disciplines of genetics and biochemistry. In this chapter we will review the major early developments in the history of this hybrid discipline, beginning with the earliest genetic experiments performed by Gregor Mendel in the mid-nineteenth century.

In Chapters 2 and 3 we will add more substance to this brief outline. By definition, the early work on genes cannot be considered molecular biology, or even molecular genetics, because early geneticists did not know the molecular nature of genes. Instead, we call it transmission genetics because it deals with the transmission of traits from parental organisms to their offspring. In fact, the chemical composition of genes was not known until 1944. At that point, it became possible to study genes as molecules, and the discipline of molecular biology began.

1.1 Transmission Genetics

In 1865, Gregor Mendel (Figure 1.1) published his findings on the inheritance of seven different traits in the garden pea. Before Mendel's research, scientists thought inheritance occurred through a blending of each trait of the parents in the offspring. Mendel concluded instead that inheritance is particulate. That is, each parent contributes particles, or genetic units, to the offspring. We now call these particles **genes**. Furthermore, by carefully counting the number of progeny plants having a given **phenotype**, or observable characteristic (e.g., yellow seeds, white flowers), Mendel was able to make some important generalizations. The word *phenotype*, by the way, comes from the same Greek root as *phenomenon*, meaning *appearance*. Thus, a tall pea plant exhibits the tall phenotype, or appearance. *Phenotype* can also refer to the whole set of observable characteristics of an organism.

Mendel's Laws of Inheritance

Mendel saw that a gene can exist in different forms called **alleles**. For example, the pea can have either yellow or green seeds. One allele of the gene for seed color gives rise to yellow seeds, the other to green. Moreover, one allele can be **dominant** over the other, **recessive**, allele. Mendel demonstrated that the allele for yellow seeds was dominant when he mated a green-seeded pea with a yellow-seeded pea. All of the progeny in the first filial generation (F_1) had yellow seeds. However, when these F_1 yellow peas were allowed to self-fertilize, some green-seeded peas reappeared. The ratio of yellow to green seeds in the second filial generation (F_2) was very close to 3:1.

The term *filial* comes from the Latin: *filius*, meaning son; *filia*, meaning daughter. Therefore, the first filial generation (F_1) contains the offspring (sons and daughters) of the original parents. The second filial generation (F_2) is the offspring of the F_1 individuals.

Mendel concluded that the allele for green seeds must have been preserved in the F_1 generation, even though it did not affect the seed color of those peas. His explanation

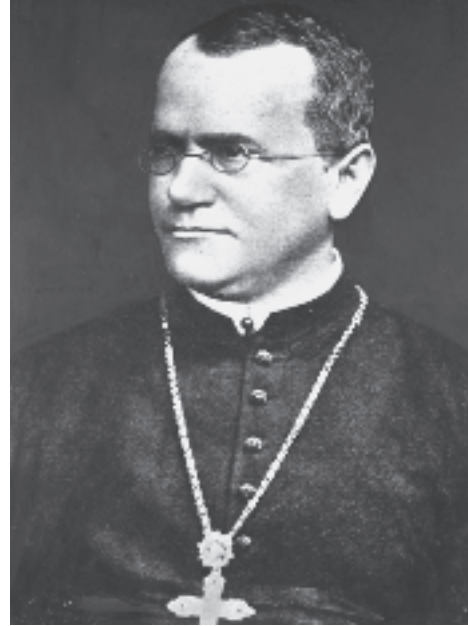


Figure 1.1 Gregor Mendel. (Source: © Pixtal/age Fotostock RF.)

was that each parent plant carried two copies of the gene; that is, the parents were **diploid**, at least for the characteristics he was studying. According to this concept, **homozygotes** have two copies of the same allele, either two alleles for yellow seeds or two alleles for green seeds. **Heterozygotes** have one copy of each allele. The two parents in the first mating were homozygotes; the resulting F_1 peas were all heterozygotes. Further, Mendel reasoned that sex cells contain only one copy of the gene; that is, they are **haploid**. Homozygotes can therefore produce sex cells, or **gametes**, that have only one allele, but heterozygotes can produce gametes having either allele.

This is what happened in the matings of yellow with green peas: The yellow parent contributed a gamete with a gene for yellow seeds; the green parent, a gamete with a gene for green seeds. Therefore, all the F_1 peas got one allele for yellow seeds and one allele for green seeds. They had not lost the allele for green seeds at all, but because yellow is dominant, all the seeds were yellow. However, when these heterozygous peas were self-fertilized, they produced gametes containing alleles for yellow and green color in equal numbers, and this allowed the green phenotype to reappear.

Here is how that happened. Assume that we have two sacks, each containing equal numbers of green and yellow marbles. If we take one marble at a time out of one sack and pair it with a marble from the other sack, we will wind up with the following results: one-quarter of the pairs will be yellow/yellow; one-quarter will be green/green; and the remaining one-half will be yellow/green. The alleles for yellow and green peas work the same way.

Recalling that yellow is dominant, you can see that only one-quarter of the progeny (the green/green ones) will be green. The other three-quarters will be *yellow* because they have at least one allele for yellow seeds. Hence, the ratio of yellow to green peas in the second (F_2) generation is 3:1.

Mendel also found that the genes for the seven different characteristics he chose to study operate independently of one another. Therefore, combinations of alleles of two different genes (e.g., yellow or green peas with round or wrinkled seeds, where yellow and round are dominant and green and wrinkled are recessive) gave ratios of 9:3:3:1 for yellow/round, yellow/wrinkled, green/round, and green/wrinkled, respectively. Inheritance that follows the simple laws that Mendel discovered can be called **Mendelian inheritance**.

SUMMARY Genes can exist in several different forms, or alleles. One allele can be dominant over another, so heterozygotes having two different alleles of one gene will generally exhibit the characteristic dictated by the dominant allele. The recessive allele is not lost; it can still exert its influence when paired with another recessive allele in a homozygote.

The Chromosome Theory of Inheritance

Other scientists either did not know about or uniformly ignored the implications of Mendel's work until 1900 when three botanists, who had arrived at similar conclusions independently, rediscovered it. After 1900, most geneticists accepted the particulate nature of genes, and the field of genetics began to blossom. One factor that made it easier for geneticists to accept Mendel's ideas was a growing understanding of the nature of chromosomes, which had begun in the latter half of the nineteenth century. Mendel had predicted that gametes would contain only one allele of each gene instead of two. If chromosomes carry the genes, their numbers should also be reduced by half in the gametes—and they are. Chromosomes therefore appeared to be the discrete physical entities that carry the genes.

This notion that chromosomes carry genes is the **chromosome theory of inheritance**. It was a crucial new step in genetic thinking. No longer were genes disembodied factors; now they were observable objects in the cell nucleus. Some geneticists, particularly Thomas Hunt Morgan (Figure 1.2), remained skeptical of this idea. Ironically, in 1910 Morgan himself provided the first definitive evidence for the chromosome theory.

Morgan worked with the fruit fly (*Drosophila melanogaster*), which was in many respects a much more



Figure 1.2 Thomas Hunt Morgan. (Source: National Library of Medicine.)

convenient organism than the garden pea for genetic studies because of its small size, short generation time, and large number of offspring. When he mated red-eyed flies (dominant) with white-eyed flies (recessive), most, but not all, of the F_1 progeny were red-eyed. Furthermore, when Morgan mated the red-eyed males of the F_1 generation with their red-eyed sisters, they produced about one-quarter white-eyed males, but no white-eyed females. In other words, the eye color phenotype was **sex-linked**. It was transmitted along with sex in these experiments. How could this be?

We now realize that sex and eye color are transmitted together because the genes governing these characteristics are located on the same chromosome—the X chromosome. (Most chromosomes, called **autosomes**, occur in pairs in a given individual, but the X chromosome is an example of a **sex chromosome**, of which the female fly has two copies and the male has one.) However, Morgan was reluctant to draw this conclusion until he observed the same sex linkage with two more phenotypes, miniature wing and yellow body, also in 1910. That was enough to convince him of the validity of the chromosome theory of inheritance.

Before we leave this topic, let us make two crucial points. First, every gene has its place, or **locus**, on a chromosome. Figure 1.3 depicts a hypothetical chromosome and the positions of three of its genes, called A, B, and C. Second, diploid organisms such as human beings normally have two copies of all chromosomes (except sex chromosomes). That means that they have two copies of most genes, and that these copies can be the same alleles, in which case the organism is **homozygous**, or different

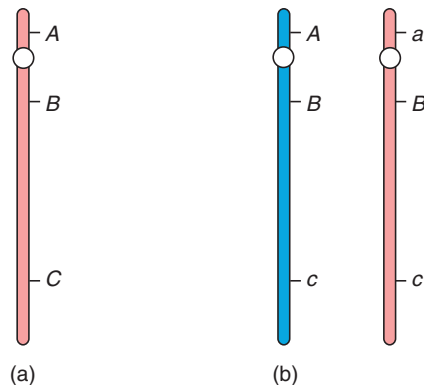


Figure 1.3 Location of genes on chromosomes. (a) A schematic diagram of a chromosome, indicating the positions of three genes: A, B, and C. (b) A schematic diagram of a diploid pair of chromosomes, indicating the positions of the three genes—A, B, and C—on each, and the genotype (A or a; B or b; and C or c) at each locus.

alleles, in which case it is **heterozygous**. For example, Figure 1.3b shows a diploid pair of chromosomes with different alleles at one locus (*Aa*) and the same alleles at the other two loci (*BB* and *cc*). The **genotype**, or allelic constitution, of this organism with respect to these three genes, is *AaBBcc*. Because this organism has two different alleles (*A* and *a*) in its two chromosomes at the *A* locus, it is heterozygous at that locus (Greek: *hetero*, meaning different). Since it has the same, dominant *B* allele in both chromosomes at the *B* locus, it is homozygous dominant at that locus (Greek: *homo*, meaning same). And because it has the same, recessive *c* allele in both chromosomes at the *C* locus, it is homozygous recessive there. Finally, because the *A* allele is dominant over the *a* allele, the phenotype of this organism would be the dominant phenotype at the *A* and *B* loci and the recessive phenotype at the *C* locus.

This discussion of varying phenotypes in *Drosophila* gives us an opportunity to introduce another important genetic concept: **wild-type** versus **mutant**. The wild-type phenotype is the most common, or at least the generally accepted standard, phenotype of an organism. To avoid the mistaken impression that a wild organism is automatically a wild-type, some geneticists prefer the term **standard type**. In *Drosophila*, red eyes and full-size wings are wild-type. Mutations in the *white* and *miniature* genes result in mutant flies with white eyes and miniature wings, respectively. Mutant alleles are usually recessive, as in these two examples, but not always.

Genetic Recombination and Mapping

It is easy to understand that genes on separate chromosomes behave independently in genetic experiments, and that genes on the same chromosome—like the genes for miniature wing (*miniature*) and white eye (*white*)—behave

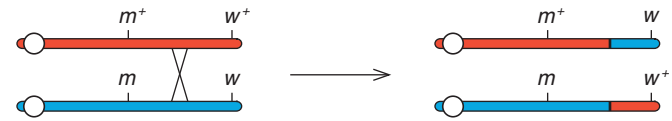


Figure 1.4 Recombination in *Drosophila*. The two X chromosomes of the female are shown schematically. One of them (red) carries two wild-type genes: (m^+), which results in normal wings, and (w^+), which gives red eyes. The other (blue) carries two mutant genes: *miniature* (m) and *white* (w). During egg formation, a recombination, or crossing over, indicated by the crossed lines, occurs between these two genes on the two chromosomes. The result is two recombinant chromosomes with mixtures of the two parental alleles. One is $m^+ w$, the other is $m w^+$.

as if they are linked. However, genes on the same chromosome usually do not show perfect **genetic linkage**. In fact, Morgan discovered this phenomenon when he examined the behavior of the sex-linked genes he had found. For example, although *white* and *miniature* are both on the X chromosome, they remain linked in offspring only 65.5% of the time. The other offspring have a new combination of alleles not seen in the parents and are therefore called **recombinants**.

How are these recombinants produced? The answer was already apparent by 1910, because microscopic examination of chromosomes during meiosis (gamete formation) had shown crossing over between **homologous chromosomes** (chromosomes carrying the same genes, or alleles of the same genes). This resulted in the exchange of genes between the two homologous chromosomes. In the previous example, during formation of eggs in the female, an X chromosome bearing the *white* and *miniature* alleles experienced crossing over with a chromosome bearing the red eye and normal wing alleles (Figure 1.4). Because the crossing-over event occurred between these two genes, it brought together the *white* and normal wing alleles on one chromosome and the red (normal eye) and *miniature* alleles on the other. Because it produced a new combination of alleles, we call this process **recombination**.

Morgan assumed that genes are arranged in a linear fashion on chromosomes, like beads on a string. This, together with his awareness of recombination, led him to propose that the farther apart two genes are on a chromosome, the more likely they are to recombine. This makes sense because there is simply more room between widely spaced genes for crossing over to occur. A.H. Sturtevant extended this hypothesis to predict that a mathematical relationship exists between the distance separating two genes on a chromosome and the frequency of recombination between these two genes. Sturtevant collected data on recombination in the fruit fly that supported his hypothesis. This established the rationale for **genetic mapping** techniques still in use today. Simply stated, if two loci recombine with a frequency of 1%, we say that they are separated by a map distance of one **centimorgan** (named for Morgan



Figure 1.5 Barbara McClintock. (Source: Bettmann Archive/Corbis.)

himself). By the 1930s, other investigators found that the same rules applied to other **eukaryotes** (nucleus-containing organisms), including the mold *Neurospora*, the garden pea, maize (corn), and even human beings. These rules also apply to **prokaryotes**, organisms in which the genetic material is not confined to a nuclear compartment.

Physical Evidence for Recombination

Barbara McClintock (Figure 1.5) and Harriet Creighton provided a direct physical demonstration of recombination in 1931. By examining maize chromosomes microscopically, they could detect recombinations between two easily identifiable features of a particular chromosome (a knob at one end and a long extension at the other). Furthermore, whenever this physical recombination occurred, they could also detect recombination genetically. Thus, they established a direct relationship between a region of a chromosome and a gene. Shortly after McClintock and Creighton performed this work on maize, Curt Stern observed the same phenomenon in *Drosophila*. So recombination could be detected both physically and genetically in animals as well as plants. McClintock later performed even more notable work when she discovered transposons, moveable genetic elements (Chapter 23), in maize.

SUMMARY The chromosome theory of inheritance holds that genes are arranged in linear fashion on chromosomes. The reason that certain traits tend to be inherited together is that the genes governing these traits are on the same chromosome. However, recombination between two homologous chromosomes during meiosis can scramble the parental alleles to give nonparental combinations. The farther apart two genes are on a chromosome the more likely such recombination between them will be.



Figure 1.6 Friedrich Miescher. (Source: National Library of Medicine.)

1.2 Molecular Genetics

The studies just discussed tell us important things about the transmission of genes and even about how to map genes on chromosomes, but they do not tell us what genes are made of or how they work. This has been the province of molecular genetics, which also happens to have its roots in Mendel's era.

The Discovery of DNA

In 1869, Friedrich Miescher (Figure 1.6) discovered in the cell nucleus a mixture of compounds that he called nuclein. The major component of nuclein is **deoxyribonucleic acid (DNA)**. By the end of the nineteenth century, chemists had learned the general structure of DNA and of a related compound, **ribonucleic acid (RNA)**. Both are long polymers—chains of small compounds called nucleotides. Each nucleotide is composed of a sugar, a phosphate group, and a base. The chain is formed by linking the sugars to one another through their phosphate groups.

The Composition of Genes By the time the chromosome theory of inheritance was generally accepted, geneticists agreed that the chromosome must be composed of a polymer of some kind. This would agree with its role as a string of genes. But which polymer is it? Essentially, the choices were three: DNA, RNA, and **protein**. Protein was the other major component of Miescher's nuclein; its chain is composed of links called **amino acids**. The amino acids in protein are joined by **peptide bonds**, so a single protein chain is called a **polypeptide**.

Oswald Avery (Figure 1.7) and his colleagues demonstrated in 1944 that DNA is the right choice (Chapter 2). These investigators built on an experiment performed earlier by Frederick Griffith in which he transferred a genetic trait from one strain of bacteria to another. The trait was virulence, the ability to cause a lethal infection,



Figure 1.7 Oswald Avery. (Source: National Academy of Sciences.)

and it could be transferred simply by mixing dead virulent cells with live avirulent (nonlethal) cells. It was very likely that the substance that caused the transformation from avirulence to virulence in the recipient cells was the gene for virulence, because the recipient cells passed this trait on to their progeny.

What remained was to learn the chemical nature of the transforming agent in the dead virulent cells. Avery and his coworkers did this by applying a number of chemical and biochemical tests to the transforming agent, showing that it had the characteristics of DNA, not of RNA or protein.

The Relationship Between Genes and Proteins

The other major question in molecular genetics is this: How do genes work? To lay the groundwork for the answer to this question, we have to backtrack again, this time to 1902. That was the year Archibald Garrod noticed that the human disease alcaptonuria seemed to behave as a Mendelian recessive trait. It was likely, therefore, that the disease was caused by a defective, or mutant, gene. Moreover, the main symptom of the disease was the accumulation of a black pigment in the patient's urine, which Garrod believed derived from the abnormal buildup of an intermediate compound in a biochemical pathway.

By this time, biochemists had shown that all living things carry out countless chemical reactions and that these reactions are accelerated, or catalyzed, by proteins called **enzymes**. Many of these reactions take place in sequence, so that one chemical product becomes the starting material, or substrate, for the next reaction. Such sequences of reactions are called **pathways**, and the products or substrates within a pathway are called



(a)



(b)

Figure 1.8 (a) George Beadle; (b) E. L. Tatum. (Source: (a, b) AP/Wide World Photos.)

intermediates. Garrod postulated that an intermediate accumulated to abnormally high levels in alcaptonuria because the enzyme that would normally convert this intermediate to the next was defective. Putting this idea together with the finding that alcaptonuria behaved genetically as a Mendelian recessive trait, Garrod suggested that a defective gene gives rise to a defective enzyme. To put it another way: A gene is responsible for the production of an enzyme.

Garrod's conclusion was based in part on conjecture; he did not really know that a defective enzyme was involved in alcaptonuria. It was left for George Beadle and E. L. Tatum (Figure 1.8) to prove the relationship between genes and enzymes. They did this using the mold *Neurospora* as their experimental system. *Neurospora* has an enormous advantage over the human being as the subject of genetic experiments. By using *Neurospora*, scientists are not limited to the mutations that nature provides, but can use **mutagens** to introduce mutations into genes and then observe the effects of these mutations on biochemical pathways. Beadle and Tatum found many instances where they could create *Neurospora* mutants and then pin the defect down to a single step in a biochemical pathway, and therefore to a single enzyme (see Chapter 3). They did this by adding the intermediate that would normally be made by the defective enzyme and showing that it restored normal growth. By circumventing the blockade, they discovered where it was. In these same cases, their genetic experiments showed that a single gene was involved. Therefore, a defective gene gives a defective (or absent) enzyme. In other words, a gene seemed to be responsible for making one enzyme. This was the one-gene/one-enzyme hypothesis. This hypothesis was actually not quite right for at least three reasons: (1) An enzyme can be composed of



Figure 1.9 James Watson (left) and Francis Crick.
(Source: © A. Barrington Brown/Photo Researchers, Inc.)



(a)



(b)

Figure 1.10 (a) Rosalind Franklin; (b) Maurice Wilkins. (Sources: (a) From *The Double Helix* by James D. Watson, 1968, Atheneum Press, NY. © Cold Spring Harbor Laboratory Archives. (b) Courtesy Professor M. H. F. Wilkins, Biophysics Dept., King's College, London.)

more than one polypeptide chain, whereas a gene has the information for making only one polypeptide chain. (2) Many genes contain the information for making polypeptides that are not enzymes. (3) As we will see, the end products of some genes are not polypeptides, but RNAs. A modern restatement of the hypothesis would be: Most genes contain the information for making one polypeptide. This hypothesis is correct for prokaryotes and lower eukaryotes, but must be qualified for higher eukaryotes, such as humans, where a gene can give rise to different polypeptides through an *alternative splicing* mechanism we will discuss in Chapter 14.



(a)



(b)

Figure 1.11 (a) Matthew Meselson; (b) Franklin Stahl. (Sources: (a) Courtesy Dr. Matthew Meselson. (b) Cold Spring Harbor Laboratory Archives.)

Activities of Genes

Let us now return to the question at hand: How do genes work? This is really more than one question because genes do more than one thing. First, they are replicated faithfully; second, they direct the production of RNAs and proteins; third, they accumulate mutations and so allow evolution. Let us look briefly at each of these activities.

How Genes Are Replicated First of all, how is DNA replicated faithfully? To answer that question, we need to know the overall structure of the DNA molecule as it is found in the chromosome. James Watson and Francis Crick (Figure 1.9) provided the answer in 1953 by building models based on chemical and physical data that had been gathered in other laboratories, primarily x-ray diffraction data collected by Rosalind Franklin and Maurice Wilkins (Figure 1.10).

Watson and Crick proposed that DNA is a **double helix**—two DNA strands wound around each other. More important, the bases of each strand are on the inside of the helix, and a base on one strand pairs with one on the other in a very specific way. DNA has only four different bases: adenine, guanine, cytosine, and thymine, which we abbreviate A, G, C, and T. Wherever we find an A in one strand, we always find a T in the other; wherever we find a G in one strand, we always find a C in the other. In a sense, then, the two strands are complementary. If we know the base sequence of one, we automatically know the sequence of the other. This complementarity is what allows DNA to be replicated faithfully. The two strands come apart, and enzymes build new partners for them using the old strands as templates and following the Watson–Crick base-pairing rules (A with T, G with C). This is called **semiconservative replication** because one strand of the parental double helix is conserved in each of the daughter double helices. In 1958, Matthew Meselson and Franklin Stahl (Figure 1.11)

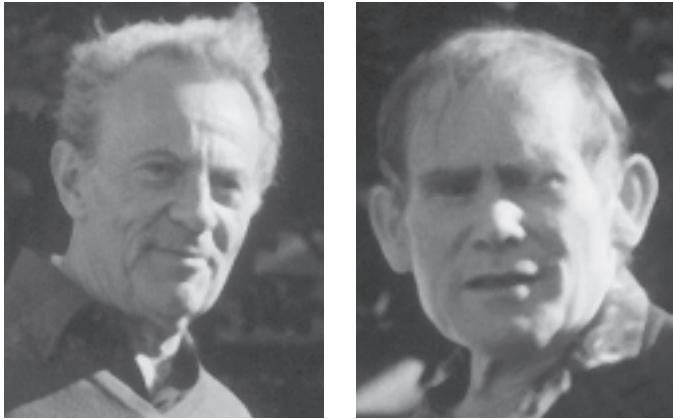


Figure 1.12 (a) François Jacob; (b) Sydney Brenner. (Source: (a, b) Cold Spring Harbor Laboratory Archives.)

proved that DNA replication in bacteria follows the semi-conservative pathway (see Chapter 20).

How Genes Direct the Production of Polypeptides **Gene expression** is the process by which a cell makes a gene product (an RNA or a polypeptide). Two steps, called **transcription** and **translation**, are required to make a polypeptide from the instructions in a DNA gene. In the transcription step, an enzyme called RNA polymerase makes a copy of one of the DNA strands; this copy is not DNA, but its close cousin RNA. In the translation step, this RNA (**messenger RNA**, or **mRNA**) carries the genetic instructions to the cell's protein factories, called **ribosomes**. The ribosomes “read” the **genetic code** in the mRNA and put together a protein according to its instructions.

Actually, the ribosomes already contain molecules of RNA, called **ribosomal RNA (rRNA)**. Francis Crick originally thought that this RNA residing in the ribosomes carried the message from the gene. According to this theory, each ribosome would be capable of making only one kind of protein—the one encoded in its rRNA. François Jacob and Sydney Brenner (Figure 1.12) had another idea: The ribosomes are nonspecific translation machines that can make an unlimited number of different proteins, according to the instructions in the mRNAs that visit the ribosomes. Experiment has shown that this idea is correct (Chapter 3).

What is the nature of this genetic code? Marshall Nirenberg and Gobind Khorana (Figure 1.13), working independently with different approaches, cracked the code in the early 1960s (Chapter 18). They found that 3 bases constitute a code word, called a **codon**, that stands for one amino acid. Out of the 64 possible 3-base codons, 61 specify amino acids; the other three are stop signals.



Figure 1.13 Gobind Khorana (left) and Marshall Nirenberg. (Source: Corbis/Bettmann Archive.)

The ribosomes scan a messenger RNA 3 bases at a time and bring in the corresponding amino acids to link to the growing polypeptide chain. When they reach a stop signal, they release the completed polypeptide.

How Genes Accumulate Mutations Genes change in a number of ways. The simplest is a change of one base to another. For example, if a certain codon in a gene is GAG (for the amino acid called glutamate), a change to GTG converts it to a codon for another amino acid, valine. The protein that results from this mutated gene will have a valine where it ought to have a glutamate. This may be one change out of hundreds of amino acids, but it can have profound effects. In fact, this specific change has occurred in the gene for one of the human blood proteins and is responsible for the genetic disorder we call sickle cell disease.

Genes can suffer more profound changes, such as deletions or insertions of large pieces of DNA. Segments of DNA can even move from one locus to another. The more drastic the change, the more likely that the gene or genes involved will be totally inactivated.

Gene Cloning Since the 1970s, geneticists have learned to isolate genes, place them in new organisms, and reproduce them by a set of techniques collectively known as **gene cloning**. Cloned genes not only give molecular biologists plenty of raw materials for their studies, they also can be induced to yield their protein products. Some of these, such as human insulin or blood clotting factors, can be very useful. Cloned genes can also be transplanted to plants and animals, including humans.

These transplanted genes can alter the characteristics of the recipient organisms, so they may provide powerful tools for agriculture and for intervening in human genetic diseases. We will examine gene cloning in detail in Chapter 4.

SUMMARY All cellular genes are made of DNA arranged in a double helix. This structure explains how genes engage in their three main activities: replication, carrying information, and collecting mutations. The complementary nature of the two DNA strands in a gene allows them to be replicated faithfully by separating and serving as templates for the assembly of two new complementary strands. The sequence of nucleotides in a gene is a genetic code that carries the information for making an RNA. Most of these are messenger RNAs that carry the information to protein-synthesizing ribosomes. The end result is a new polypeptide chain made according to the gene's instructions. A change in the sequence of bases constitutes a mutation, which can change the sequence of amino acids in the gene's polypeptide product. Genes can be cloned, allowing molecular biologists to harvest abundant supplies of their products.

1.3 The Three Domains of Life

In the early part of the twentieth century, scientists divided all life into two kingdoms: animal and plant. Bacteria were considered plants, which is why we still refer to the bacteria in our guts as intestinal “flora.” But after the middle of the century, this classification system was abandoned in favor of a five-kingdom system that included bacteria, fungi, and protists, in addition to plants and animals.

Then in the late 1970s, Carl Woese (Figure 1.14) performed sequencing studies on the ribosomal RNA genes of many different organisms and reached a startling conclusion: A class of organisms that had been classified as bacteria have rRNA genes that are more similar to those of eukaryotes than they are to those of classical bacteria like *E. coli*. Thus, Woese named these organisms *archaeobacteria*, to distinguish them from true bacteria, or *eubacteria*. However, as more and more molecular evidence accumulated, it became clear that the archaeobacteria, despite a superficial resemblance, are not really bacteria. They represent a distinct domain of life, so Woese changed their name to archaea. Now we recognize three **domains of life: bacteria, eukaryota, and archaea**. Like bacteria, archaea are **prokaryotes**—organisms without nuclei—but their molecular biology is actually more like that of eukaryotes than that of bacteria.



Figure 1.14 Carl Woese. (Source: Courtesy U. of Ill at Urbana Champaign.)

The archaea live in the most inhospitable regions of the earth. Some of them are **thermophiles** (“heat-lovers”) that live in seemingly unbearably hot zones at temperatures above 100°C near deep-ocean geothermal vents or in hot springs such as those in Yellowstone National Park. Others are **halophiles** (halogen-lovers) that can tolerate very high salt concentrations that would dessicate and kill other forms of life. Still others are **methanogens** (“methane-producers”) that inhabit environments such as a cow’s stomach, which explains why cows are such a good source of methane.

In this book, we will deal mostly with the first two domains, because they are the best studied. However, we will encounter some interesting aspects of the molecular biology of the archaea throughout this book, including details of their transcription in Chapter 11. And in Chapter 24, we will learn that an archaeon, *Methanococcus jannaschii*, was among the first organisms to have its genome sequenced.

SUMMARY All living things are grouped into three domains: bacteria, eukaryota, and archaea. Although the archaea resemble the bacteria physically, some aspects of their molecular biology are more similar to those of eukaryota.

This concludes our brief chronology of molecular biology. Table 1.1 reviews some of the milestones. Although it is a very young discipline, it has an exceptionally rich history, and molecular biologists are now adding new knowledge at an explosive rate. Indeed, the pace of discovery in molecular biology, and the power of its techniques, has led many commentators to call it a revolution. Because some of the most important changes in medicine and agriculture over the next few decades are likely to depend on the manipulation of genes by molecular biologists, this revolution will touch everyone’s life in one way or another. Thus, you are

Table 1.1 Molecular Biology Time Line

1859	Charles Darwin	Published <i>On the Origin of Species</i>
1865	Gregor Mendel	Advanced the principles of segregation and independent assortment
1869	Friedrich Miescher	Discovered DNA
1900	Hugo de Vries, Carl Correns, Erich von Tschermak	Rediscovered Mendel's principles
1902	Archibald Garrod	First suggested a genetic cause for a human disease
1902	Walter Sutton, Theodor Boveri	Proposed the chromosome theory
1910, 1916	Thomas Hunt Morgan, Calvin Bridges	Demonstrated that genes are on chromosomes
1913	A.H. Sturtevant	Constructed a genetic map
1927	H.J. Muller	Induced mutation by x-rays
1931	Harriet Creighton, Barbara McClintock	Obtained physical evidence for recombination
1941	George Beadle, E.L. Tatum	Proposed the one-gene/one-enzyme hypothesis
1944	Oswald Avery, Colin McLeod, Maclyn McCarty	Identified DNA as the material genes are made of
1953	James Watson, Francis Crick, Rosalind Franklin, Maurice Wilkins	Determined the structure of DNA
1958	Matthew Meselson, Franklin Stahl	Demonstrated the semiconservative replication of DNA
1961	Sydney Brenner, François Jacob, Matthew Meselson	Discovered messenger RNA
1966	Marshall Nirenberg, Gobind Khorana	Finished unraveling the genetic code
1970	Hamilton Smith	Discovered restriction enzymes that cut DNA at specific sites, which made cutting and pasting DNA easy, thus facilitating DNA cloning
1972	Paul Berg	Made the first recombinant DNA in vitro
1973	Herb Boyer, Stanley Cohen	First used a plasmid to clone DNA
1977	Frederick Sanger	Worked out methods to determine the sequence of bases in DNA and determined the base sequence of an entire viral genome (ϕ X174)
1977	Phillip Sharp, Richard Roberts, and others	Discovered interruptions (introns) in genes
1993	Victor Ambros and colleagues	Discovered that a cellular microRNA can decrease gene expression by base-pairing to an mRNA
1995	Craig Venter, Hamilton Smith	Determined the base sequences of the genomes of two bacteria: <i>Haemophilus influenzae</i> and <i>Mycoplasma genitalium</i> , the first genomes of free-living organisms to be sequenced
1996	Many investigators	Determined the base sequence of the genome of brewer's yeast, <i>Saccharomyces cerevisiae</i> , the first eukaryotic genome to be sequenced
1997	Ian Wilmut and colleagues	Cloned a sheep (Dolly) from an adult sheep udder cell
1998	Andrew Fire and colleagues	Discovered that RNAi works by degrading mRNAs containing the same sequence as an invading double-stranded RNA
2003	Many investigators	Reported a finished sequence of the human genome
2005	Many investigators	Reported the rough draft of the genome of the chimpanzee, our closest relative
2007	Craig Venter and colleagues	Used traditional sequencing to obtain the first sequence of an individual human (Craig Venter).
2008	Jian Wang and colleagues	Used "next generation" sequencing to obtain the first sequence of an Asian (Han Chinese) human.
2008	David Bentley and colleagues	Used single molecule sequencing to obtain the first sequence of an African (Nigerian) human.

embarking on a study of a subject that is not only fascinating and elegant, but one that has practical importance as well. F. H. Westheimer, professor emeritus of chemistry at Harvard University, put it well: "The greatest intellectual

revolution of the last 40 years may have taken place in biology. Can anyone be considered educated today who does not understand a little about molecular biology?" Happily, after this course you should understand more than a little.

SUMMARY

Genes can exist in several different forms called alleles. A recessive allele can be masked by a dominant one in a heterozygote, but it does not disappear. It can be expressed again in a homozygote bearing two recessive alleles.

Genes exist in a linear array on chromosomes. Therefore, traits governed by genes that lie on the same chromosome can be inherited together. However, recombination between homologous chromosomes occurs during meiosis, so that gametes bearing nonparental combinations of alleles can be produced. The farther apart two genes lie on a chromosome, the more likely such recombination between them will be.

Most genes are made of double-stranded DNA arranged in a double helix. One strand is the complement of the other, which means that faithful gene replication requires that the two strands separate and acquire complementary partners. The linear sequence of bases in a typical gene carries the information for making a protein.

The process of making a gene product is called gene expression. It occurs in two steps: transcription and

translation. In the transcription step, RNA polymerase makes a messenger RNA, which is a copy of the information in the gene. In the translation step, ribosomes “read” the mRNA and make a protein according to its instructions. Thus, a change (mutation) in a gene’s sequence may cause a corresponding change in the protein product.

All living things are grouped into three domains: bacteria, eukaryota, and archaea. The archaea resemble bacteria physically, but their molecular biology more closely resembles that of eukaryota.

SUGGESTED READINGS

- Creighton, H.B., and B. McClintock. 1931. A correlation of cytological and genetical crossing-over in *Zea mays*. *Proceedings of the National Academy of Sciences* 17:492–97.
- Mirsky, A.E. 1968. The discovery of DNA. *Scientific American* 218 (June):78–88.
- Morgan, T.H. 1910. Sex-limited inheritance in *Drosophila*. *Science* 32:120–22.
- Sturtevant, A.H. 1913. The linear arrangement of six sex-linked factors in *Drosophila*, as shown by their mode of association. *Journal of Experimental Zoology* 14:43–59.