From: Sterelny, K., Griffiths, P.E., 1999, Sex and Death: An Introduction to Philosophy of Biology, Chicago University Press, Chicago.

Chapter

6

Mendel and Molecules

6.1 How Theories Relate: Displacement, Incorporation, and Integration

One problem in philosophy of science concerns the relationship between apparently different theories of the same domain. For example, in psychology, we have three apparently different ways of explaining human behavior. Cognitive psychology explains human behavior by seeing it as the result of information processing. Its program is to explain, say, our ability to predict others' behavior by characterizing the information about others we possess, the form in which that information is stored, and the techniques we use to process and deploy that information. But the neurosciences are also in the business of explaining human behavior. Those disciplines are gradually developing an account of the physiological mechanisms on which our behavioral abilities depend. Furthermore, we were not wholly incapable of explaining human behavior before the scientific developments of the twentieth century. For thousands of years we have had at our disposal a "folk psychology" through which we have explained the behavior of others. These explanations are couched in terms of beliefs, goals, emotions, moods, and the like. How do the explanations of folk psychology relate to those developed in the natural sciences? How do the two scientific programs relate to each other?

This general problem arises in biology as well. As we saw in section 2.2, heredity—parent/offspring similarity—is central to evolution. Unless offspring tend to resemble their parents more than they resemble some randomly chosen member of their parents' generation, natural selection is powerless to change the character of a population over time. But there seem to be two different theoretical programs through which this central phenomenon can be studied. The first of these dates back to Gregor Mendel's

work in the mid-nineteenth century; the second began when the rediscovery of his work at the beginning of the twentieth century prompted a search for its cellular and molecular basis. What follows is a cartoon version of these programs; we go into more detail in section 6.2.

It was Mendel who hit on the idea of genes as discrete units of inheritance while studying the results of pea breeding experiments in the 1860s. When he focused on two states of a single character, round versus wrinkled seeds in true-breeding pea lineages, he noted that first-generation hybrids were all round, but that second-generation hybrids were not. Some—about ¾—were round, but ¼ were not. When he considered not just one, but two traits, seed texture and flower color, once again the first-generation hybrids were uniformly round-seeded, yellow-flowered peas. But the second-generation hybrids were not. Roughly 9/16 of the second generation were like the first-generation hybrids. But about 3/16 were yellow-flowered, wrinkled-seeded peas; about 3/16 were green-flowered, round-seeded peas; and about 1/16 were both green-flowered and wrinkle-seeded.

Mendel realized that these results fell into place with the following assumptions:

1. Phenotypic traits such as color and texture are determined by a unitary hereditary factor. These factors can exist in alternative forms, or *alleles*.

2. The gametes of an organism (the pollen or the ova) carry just one of the alternate character states of these traits (one of the factors for yellow or green; round or wrinkled).

3. When an organism is formed from two gametes that carry rival factors for one trait, one dominates the other. In this case, the factor for round is *dominant* over the factor for wrinkled. In other words, the factor for wrinkled is *recessive*.

4. When a first-generation hybrid organism (the *first filial* or F_1 generation) forms gametes, about 50% of the gametes carry one factor, and about 50% carry the other.

5. The factors for traits that are not alternatives to one another—in this case, flower color and seed shape—are inherited independently of one another. From the fact that a gamete carries the *wrinkled* factor, we can tell nothing about whether it carries the *yellow* factor, and vice versa.

As we shall see in section 6.2, after the rediscovery of Mendel's work around 1900, much was added to this picture, and it was altered in important ways. But biologists have continued to investigate heredity by studying the

Box 6.1 What Is an Allele?

Mendelian genetics defines genes, and hence variants of the same gene, through their effects on phenotypes rather than by appeal to their intrinsic physical structures. So when do we have two genes, each of which may exist in a variety of forms? When do we have different alleles of one gene? Since genes can affect more than one trait, we cannot assume that a gene that affects, say, antenna structure in fruit flies is distinct from one that affects their wing length.

Genetic complementation was a central technique in answering this question. Suppose we have two mutant flies: one with short wings, and another with wrinkled antennae. We wish to know whether we have two different mutated alleles of the same gene or mutant forms of two different genes. Mutant forms of different genes (typically) *complement* one another. That is, if we cross the short-winged fly with the wrinkled-antenna fly, and the result is phenotypically normal offspring, we can infer that the mutations are of distinct genes at different loci. We have discovered that the genes are *complementary*. The phenotypically normal offspring result because the gametes from the parent with wrinkled antennae have an unmutated, *wild-type* allele for wing length, and the gametes from the short-winged parent have an unmutated, wild-type allele for antenna form. So the offspring get one unmutated allele for each gene, and are hence phenotypically normal. The offspring are heterozygotes at both loci, with the normal (wild-type) allele dominant over the mutant allele. Clearly, this explanation of why the offspring are normal assumes that the mutations were of separate genes, hence the inference from complementation to alleles of distinct genes. On the other hand, if the hybrid generation is phenotypically unusual, we can infer that we have two mutations of the same gene, and hence two different alleles of the one gene.

patterns of parent/offspring similarity manifested in an organism's phenotype. This program is sometimes known as *transmission genetics*. The debate about human intelligence is one particularly controversial example of such studies.

Shortly after the rediscovery of Mendel's work, a second closely related program developed: an investigation into first the cellular and then the molecular basis of heredity. While the molecular basis of hereditary factors—protein versus nucleic acid—remained in dispute until the mid-twentieth century, their cellular basis in chromosomes was soon discovered. As early as 1903, Walter Sutton showed that meiosis explains our second principle, Mendel's *law of segregation*. For meiosis results in each gamete receiving just one of a homologous pair of chromosomes. Somewhat later, in T. H. Morgan's famous fly lab, the discovery of the physical location of genes on chromosomes undercut principle 5, the *law of independent assortment*. When genes are located on the same chromosome, the inheritance of one is not independent of the inheritance of the other. Further on down the track it was discovered that nucleic acids were the critical molecules making up the genes. Then, in 1953, James D. Watson and Francis Crick developed the famous double helix model of the structure of DNA. Since then, discoveries have come thick and fast.

How, then, might these theoretical programs be related? One possibility is the *displacement* of one program by another—that is, one program can show that another is simply mistaken. The geological program of plate tectonics displaced the conception of earth history in which the position of the continents was taken to be fixed. Much more controversially, Paul and Patti Churchland argue that folk psychology is being displaced by the neurosciences. It was once expected that folk psychological explanations of behavior could be "reduced" to neurophysiological explanations. The idea was to define the concepts of folk psychology—moods, emotions, and cognitive states—in neurophysiological terms. Fear, for example, might turn out to be a specific form of arousal of the autonomic nervous system. Most philosophers of mind are physicalists and think that there is nothing to the mind except the physical brain and the wider physical context it inhabits. However, it is now generally accepted that though the emotions do depend on the physiology of the nervous system, they do so in complex ways that vary from individual to individual and over time. So there is wide agreement that psychological concepts like belief and desire cannot be defined in neuroscientific terms. The Churchlands take this to be a symptom that there is something wrong with folk psychology. In their view, the failure of reduction suggests that the neurosciences should displace folk psychology (P. Churchland 1986; P. M. Churchland 1989).

A second possibility is that one program *incorporates* or absorbs the other—that the first is shown to be just a special case of the second. Planetary motions in the solar system are well described by Kepler's three laws of planetary motion:

1. The orbits of the planets are ellipses with the sun at a common focus.

2. The line joining a planet to the sun sweeps out equal areas in equal periods of time.

3. The squares of the periods of any two planets' orbits are proportional to the cubes of their mean distance from the sun.

Reduction takes place when such laws are shown to be a special case of a more general system of laws. Thus Kepler's laws were shown (with minor corrections) to be a special case of Newton's laws of motion. They can be deduced from, and hence are reduced to, those more general laws. As we shall see, "reduction" is an ambiguous notion, but construed this way, it explains why nothing is lost in the move from the old theoretical framework to the new one. The first theoretical framework is shown to have limited validity by its successor framework; it is incorporated within its successor.

Displacement and incorporation should probably be seen as two ends of a continuum rather than two sharply distinct fates. The fate of Newton's theory is often seen as intermediate between incorporation and displacement. Newton's theory correctly predicts how objects move in space and time at low speeds. At these speeds, the predictions of a theory in which an object has an absolute location in space and time are almost exactly the same as those of a theory in which an object's location is relative to the observer's frame of reference. Relativistic physics is both more accurate and covers a wider array of cases than Newtonian mechanics, but Newton's framework is shown to have some partial validity by its successor.

A third possibility is that two programs can be *integrated*. The classic theory of gases describes the lawlike relationships between observable quantities such as pressure, volume, and temperature. The kinetic theory of gases explains these relationships as the effect of random movements of large ensembles of molecules, each with a quantity of kinetic energy, which it can transfer by impact to other molecules. The explanation of the laws in terms of molecular motion supports the claim that gases are "nothing but" ensembles of molecules in motion. The ontology of the first theory—gases, heat, and pressure—is reduced to the ontology of the second theory—molecules and kinetic energy. We have here a second concept of "reduction": the objects described by one theory are "reduced to" the apparently very different entities postulated by another theory. The classic theory of gases relating pressure, volume, and temperature is sometimes called the *phenomenological* theory of gases because the properties it deals with are observable phenomena. A reduction in this second sense explains the regularities among these observable properties by appeal to the properties of their unobservable constituents.

The distinction between incorporation and integration is not sharp. If the ontological reduction is simple—if there are definitions or bridge laws linking the concepts of a reduced theory to the concepts of a reducing theory—then integration can turn into incorporation. The chemical property of valency, which measures the capacity of an element to form compounds with other elements, turns out to have a straightforward physical basis in an atom's configuration of electrons. Valency is definable in physical terms. So some chemical generalizations about the combinatory power of atoms will turn out to be special cases of physical principles about electron bonds. They can be deduced from physical generalizations via these bridge laws or definitions. Usually, however, it is at least practically necessary to continue to use phenomenological theories. Trying to calculate the efficiency of a heat pump in a freezer by tracking individual molecules would be a thankless task. And, as we shall see, there can be more fundamental reasons that block incorporation.

Prima facie, the relationship between molecular and Mendelian genetics includes elements of both incorporation and integration. Molecular mechanisms, we might suppose, explain the regularities in parent/offspring similarity revealed in Mendelian genetics. Molecular genetics seems to be a superior and more general successor to Mendelian genetics. Mendel's original laws are reasonably accurate in a limited range of cases because some of the DNA segments described by modern molecular biology are passed on from one generation to another in roughly the way Mendel postulated. When Mendel's laws are not honored, the new theory can explain what is happening instead. These considerations suggest partial incorporation. Molecular genetics also seems to reduce earlier genetic theories ontologically. Surely there is nothing more to genes than the DNA studied by molecular biologists? Classic Mendelian genetics is a phenomenological theory, for it involves observable patterns in the inheritance of phenotypic characteristics. Just as the phenomenological theory of gases, relating the observable quantities of heat, pressure, and volume, is explained by features of their microscopic constituents, so too are the generalizations of classic Mendelian genetics explained by microscopic constituents of genes. Yet for the same reasons that the phenomenological theory of gases remains useful in practice, transmission genetics retains some practical value.

No one doubts that there is something right about this picture of the relationship between Mendelian and molecular genetics. Everyone agrees that the genetic material is made up of DNA and associated molecular structures, and that the behavior of these molecular structures underlies the regularities observed by earlier geneticists. However, there is an influential group

of philosophers of biology, starting with Hull (1974), who think that the relationship between classic genetics and molecular biology is vastly more complicated than the parallels with heat, valency, or planetary motion suggest. Over this chapter and the next we shall focus on the relations between molecular and Mendelian genetics. In this discussion, the following themes will all be prominent:

1. To what extent does molecular biology vindicate the central ideas of Mendelian genetics, explaining the molecular mechanisms that underlie the patterns of similarity and difference among relatives? To what extent does molecular biology require a revision of these ideas?

2. To what extent can transmission genetics and molecular genetics be developed independently of each other? The chemical property of valency is linked via a bridge law or definition to the configuration of an atom's electrons. According to the antireductionists, the concepts of transmission genetics are not definable in any comparable way. Molecular biology illuminates many aspects of earlier genetic theory, but in complex and indirect ways. Mendelian genetics contains theoretical concepts, such as the idea that one allele is "dominant" to another, whose explanation in molecular biology varies case by case. The idea of dominance has no single, natural correlate at the molecular level. Furthermore, molecular biological explanations often refer to the wider cellular context in which molecular events occur. This seems to run counter to the idea that the behavior of larger entities is being explained in terms of their smaller constituents. So, although the transmission of similarity from parent to offspring depends on molecular mechanisms and their context, these patterns can be studied in relative independence from molecular biology. The two theories are linked by the fact that in any given case, we can explain the observable similarity between parent and offspring in molecular terms, but since these explanations vary from case to case, their integration is not tight.

3. Entwined with these specifically biological themes are more general ones about the right way to conceive of the relationship between scientific programs. Here the general issue of *reduction* looms large. As we have already noted, "reduction" is a many ways ambiguous notion. Three ideas, at least, are in play:

   a. An idea that historically has been very prominent in the discussion of reduction is the idea of *theoretical unification*. According to this conception, the aim of science is to develop systems of laws or generalizations. Particular branches of science are characterized by the laws or generalizations that they discover. We have already seen an example in planetary science, Kepler's three laws of motion. Theoretical unification was achieved when these laws were shown to be, with minor corrections, a special case of Newton's laws of motion. More controversially, and with much more correction, Newton's laws are seen as a special case of relativistic laws. Many philosophers of science interpret the relations between the generalizations of chemistry and those of physics in the same way. The generalizations of chemistry are shown to be special cases of those of physics with the aid of various bridge laws defining chemical properties in physical terms. A definition of valency in terms of electron shells is an example of such a bridge law. So theoretical unification involves the incorporation of the laws of a reduced theory into those of the reducing theory, either directly or via the aid of bridge laws. Thus one aspect of scientific progress is the construction of an increasingly general, unified conception of nature's laws.

As we shall see, it is this sense of reduction that is most under the gun in the antireductionist consensus. Hull and the other antireductionists have raised doubts about the existence of suitable bridge laws. But as we shall see in section 15.2, it is not at all clear that we should think of the branches of biology as being in the business of formulating laws or generalizations. This whole conception of reduction and the nature of science, based as it is on physics and chemistry, may not fit biology well.

   b. An important "reductive" research strategy in contemporary science is explanation by *decomposition*. How do we work out what is going on in some domain? By taking it apart and studying the components in isolation. If the system cannot be decomposed physically, we can decompose it methodologically. We do this by keeping every component but one constant, and studying the behavior of the system when that one component changes. For instance, we can establish a *norm of reaction* for a genotype by studying how a clone of plants grows when we vary different aspects of the environment, one by one. Variation in the system as a whole is studied by controlling potential sources of variation and allowing only one focal component to vary.

Those who argue for the importance of *holistic* approaches to science and against reductionism often have this conception of reduction in mind. They oppose it by arguing for the importance of *emergent* phenomena. For example, it is common to suggest that ecosystems cannot be understood by decompositional methods because crucial ecological phenomena arise only out of the interaction of many components of a system. Whatever the merits of this idea,

it is important to realize that it is quite different from the view that Hull and his allies put forward. There are, however, echoes of this idea in the view that the cellular context in which a gene acts is so important that the strategy of explanation by decomposition is undermined (7.3).

c. A third sense of reduction is the idea that a scientific explanation must include an identifiable mechanism—it cannot depend on "miracles." One reason why the proponents of continental drift remained in the minority in the period between the two world wars was that it was impossible to see how the continents *could* shift. The mechanisms proposed were unworkable. So continental drift was unpopular as a scientific theory because it depended on a spooky mechanism, a process that could not be understood as a concatenation of ordinary physical and chemical processes. The objects, mechanisms, and processes of a scientific theory must involve nothing spooky: no additions to the standard mechanical processes of the world.

We take this third idea to be an uncontroversial version of reductionism. For instance, a standard puzzle about memory is posed by the fact that humans are very good at recognizing human faces in their normal orientation, but not if the face is inverted. Explaining this phenomenon by detailing the physical changes in the parts of the brain involved in memory is in this sense a reductive process, however complex the relation between a psychological description of what we can remember and a neuroscientific description of changes in neural connectivity might be, for an account of the neural substrate would show that memory involves nothing spooky or occult. In this sense, molecular explanations of dominance or of the independent assortment of traits are reductive explanations, however complex they are, for they show that nothing spooky is in play.

So one sense of reduction clearly involves the incorporation of the reduced theory into the reducing theory. But the two other senses may not: they are compatible with the two theories being integrated without one being incorporated within the other. Consider, for example, the fact that genes are often *pleiotropic;* that is, they have effects on more than one trait. Explanation by decomposition may be an effective strategy for studying this phenomenon even if the relationship between pleiotropy and the molecular mechanisms that explain it is too complex and varied for there to be a bridge law defining it in molecular terms.

We have no interest in haggling over which of these various ideas deserves to be called "reduction." The important point is to recognize their differences, and the fact that the relationship between real theories in science will

rarely fit exactly one of these definitions cleanly. So the reader is warned: in this and the next chapter, a number of balls are in the air. We first sketch the empirical background of this controversy, and then proceed to the theoretical upshot.

6.2 What Is Mendelian Genetics?

Mendelian genetics is the theory that grew by elaboration and development of the laws of segregation and independent assortment after these were rediscovered at the beginning of the twentieth century. The first Mendelians realized that the pattern of inheritance of some biological traits could be explained by postulating a pair of factors underlying each trait—a pair of *alleles* occupying a *locus* on a chromosome. The law of segregation says that the two alleles are separated in the formation of the gametes (sex cells), with each gamete receiving only one allele. Although the alleles from two gametes are united in the *zygote* (the fertilized egg), they do not mix together, and they are separated again to form the next generation. The law of independent assortment says that the probability of a gamete receiving a particular allele at one locus is independent of which allele it receives at another locus. This second "law" was subsequently discovered to be widely violated. There are *linkages* of varying strength between loci: the stronger the linkage, the more likely the alleles are to be inherited together.

Both the original Mendelian "laws" and the exceptions to them were discovered through breeding experiments. In his seminal presentation of the antireductionist consensus, Hull followed the geneticist Theodosius Dobzhansky in using this methodological fact to distinguish the new molecular genetics. Molecular genetics is concerned with the intrinsic nature of the hereditary material; it proceeds by looking inside the cell. In contrast, "genetics is concerned with gene differences; the operation employed to discover a gene is hybridization: parents differing in some trait are crossed and the distribution of the trait in hybrid progeny is observed" (Dobzhansky 1970, 167; quoted in Hull 1974, 23).

The outcomes of breeding experiments, however, were very quickly related to *cytology*—the study of the structure and activity of cells. The discovery of chromosomes provided an explanation for the phenomenon of gene linkage. The genetic material in the cell nucleus consists of several chromosomes. If we assume that genes occur in a line along each chromosome, then genes on different chromosomes will assort independently, while those on the same chromosome will be linked together. A further cytological observation explains the fact that the links between genes can differ in strength. Chromosomes come in homologous pairs, and one of the pair is passed on

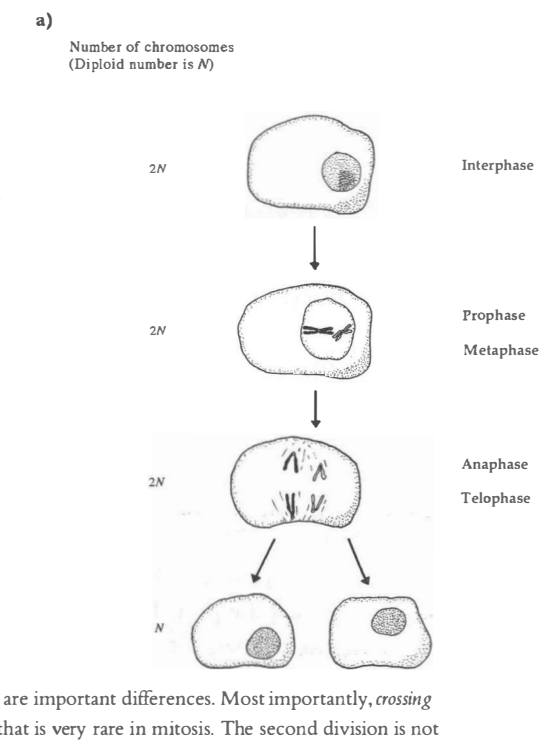


Figure 6.1 Mitosis, meiosis, and crossing-over. (a) Mitosis is the process by which cells multiply and organisms grow. It is represented here for one pair of homologous chromosomes (the two copies of the same chromosome contributed by the organism's two parents). During *interphase* the cell's DNA is replicated, so that when the chromosomes become condensed and visible in *prophase,* each consists of two *chromatids* connected by a *centromere.* During *metaphase* the nuclear membrane disintegrates, and microtubules from the centromeres join to those of the *spindle.* During *anaphase,* the chromatids are drawn apart by the spindle. During *telophase,* two new nuclear membranes form. The cell can then split into two. (b) Meiosis, or reduction division, forms four haploid sex cells by two successive divisions of one diploid cell. The process is represented here for two pairs of homologous chromosomes. The first division resembles mitosis, although there are important differences. Most importantly, *crossing over* occurs during prophase I, something that is very rare in mitosis. The second division is not preceded by DNA replication, and so produces haploid cells with half the diploid chromosome number. (c) (Adapted from Alberts et al. 1994, 100.) Crossing-over is a process in which pairs of homologous chromosomes line up with one another and exchange segments. Where the mother and father were not genetically identical, this can create new gene combinations.

to each gamete. During meiosis, homologous chromosomes cross over and recombine, so that a part of each chromosome is exchanged with the other (see figure 6.1c). The probability of two linked genes being separated by *crossing-over,* thus breaking the link between them, can be greater or smaller depending on how close together they are on a chromosome.

Two other important elements of Mendelian genetics are its account of the relations between genes and phenotypes and its account of the relations between the pairs of alleles that occupy a locus. It was natural for early Mendelians to adopt the hypothesis that there is a single gene for each phenotypic trait. It soon became clear, however, that this hypothesis could not be defended in the face of pleiotropic genes and polygenic traits. *Pleiotropy* refers to the phenomenon of one gene having many effects. Hull gives the nice example of an allele that affects both the eye color of *Drosophila* (fruit flies) and the shape of the spermatheca (an organ in females for storing sperm). *Polygenic* traits, such as human height, are affected by many different genes. Furthermore, some genes interact *epistatically:* the effect of an allelic substitution at one locus depends on which alleles are present at one or more other loci. The relation between genes and phenotypes is thus not one-to-one, but many-to-many.

The way in which the two alleles at a single locus interact to create their distinctive effect is similarly complex. An allele can be characterized as *dominant* or *recessive* relative to some other allele that can occupy the same locus. When two different alleles occur together, if the heterozygote, *Aa,* has a phenotype identical to that of an organism with two copies of one of the alleles—say, *AA*—then *A* is dominant and *a* is recessive. Numerous other categories of dominance were defined by classic geneticists. When the heterozygote expresses a trait more extremely than either homozygote, the alleles are said to be *overdominant.* When the heterozygote expresses the traits of both homozygotes, the alleles are said to be *codominant.* An allele of a pleiotropic gene may be dominant with respect to some of its effects and recessive with respect to others.

Box 6.2 Genetic Atomism

In the growth of theories of heredity and development, the gene has been pressed into service to play a number of distinct biological roles. One is transmission: the production of offspring/parent similarity. But another is mutation: the creation of an unheralded phenotypic form in offspring. Yet a third is recombination: the reshuffling of traits in the phenotype of the next generation that occurred separately in the last, and vice versa. Recombination thus defines the "grain" of inheritance. Finally, genes must somehow function in the development of the organisms that carry them.

The simplest hypothesis is that the gene is the fundamental unit of all four processes. This hypothesis was developed by Morgan and his school in the 1920s. One way of interpreting the further developments in both transmission genetics and molecular genetics since that time is that these roles have been separated. For example, the fundamental unit of mutation (the single base) is distinct from that of function (the codon, a three-base sequence), and that is different again from the unit of recombination (Portin 1993, 781).

Mendelian genetics discovers phenomena that are revealed through breeding experiments, so the explanation of dominance, overdominance, codominance, and similar effects lies outside its scope. Genes interact with one another to determine the norm of reaction of a genotype, and this interacts with environmental variables to determine a phenotype. Mendelian genetics can describe the differences made to this process when one allele is substituted for another at a particular locus on a chromosome, but it does not explain the mechanical bases of these differences. It is part of the role of molecular genetics to uncover these underlying mechanisms. The theorists who expected to reduce Mendelian genetics to molecular biology expected to find one or a few molecular mechanisms that would explain how gene substitutions cause phenotypic differences. This would have allowed them, for example, to identify the phenomenon of dominance with one or a few specific molecular mechanisms. The antireductionist consensus is generated by the fact that expectations of this sort have not been fulfilled.

6.3 Molecular Genetics: Transcription and Translation

The phrase *molecular genetics* refers to the study of the chemical nature of the hereditary material and its molecular surroundings. Chromosomes had long

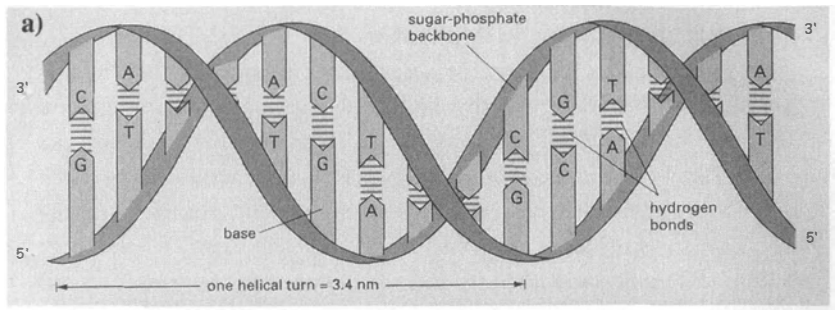


Figure 6.2 (a) The double-stranded helical structure of DNA. (b) The single-stranded structure of RNA, which is the genetic material in viruses and some bacteria. (Adapted from Alberts et al. 1994, 101.)

been known to contain nucleic acids, such as DNA, and proteins, such as histones. It finally became clear at the beginning of the 1950s that DNA was the critical ingredient of the genes. In 1953 Watson and Crick produced a successful model of the molecular structure of DNA. Since then, much has been discovered about its molecular machinery. In this context, these discoveries all contribute to a common theme: they highlight the critical role of the cellular environment in structuring the effect of DNA sequences on an organism's phenotype. The causal chain between DNA and phenotype is indirect and complex not just in having many links; it also has many branches. As we shall see, different cellular environments link identical DNA sequences to quite different phenotypic outcomes.

It was clear as soon as the structure of DNA was elucidated that this structure explains some of the phenomena observed by transmission geneticists. DNA plays its central role in life because it can be both replicated and read.

Box 6.3 DNA as Code and Replicator

DNA can be reliably replicated because guanine and adenine form hydrogen bonds with cytosine and thymine, respectively, and only with them. When the double helix is split apart, each half specifies how to reconstruct the other by forming G-C and A-T bonds. Later research has revealed how DNA functions in the formation of the proteins that make up the structural and functional elements of cells. A single strand of messenger RNA (mRNA) is transcribed by *RNA polymerase enzymes* from one half of the double strand of DNA. The DNA sequence specifies the mRNA transcript by means of the same complementary pairing that allows DNA replication (except that in the mRNA transcript, the base uracil replaces thymine). Within the DNA sequence there is a region beginning with a start and ending with a stop signal. These signals form a *reading frame.* Within the reading frame, the bases divide into three-base sequences, counting from the start signal. Each of these triples is a codon. Hence *frameshift mutations* can cause transcription of the sequence to begin at a new point by redefining the reading frame. A sequence that had been segmented into the codons, say, __/AAG/AGG/GUU/_ can become redivided into __A/AGA/GGG/UU_/.

The critical feature underlying its replicability is its *complementarity*—the fact that when the double helix splits into two single strands, each uniquely specifies the other. Each base in the sequence will pair with only one other base.

DNA reading depends on two main mechanisms, *transcription* and *translation.* First, DNA specifies *messenger RNA (mRNA)* by the same unique pairing mechanism involved in its replication. The resulting mRNA transcript, like its DNA template, is organized into three-base sequences called *codons.* This *primary transcript* plays a central role in protein synthesis, as the codons specify particular amino acids. These amino acids, in turn, are the constituents of proteins. However, it would be wrong to suppose that DNA specifies proteins in the sense of uniquely determining a particular protein. Different primary RNA transcripts can be transcribed from the same DNA sequences. It is also possible for sequences transcribed as different mRNAs to overlap one another (see box 6.3). So the relation between a given DNA sequence and the mRNA input to the protein-making system is one-to-many. When we consider the reading mechanisms of eukaryotic cells, this basic message gets further support.

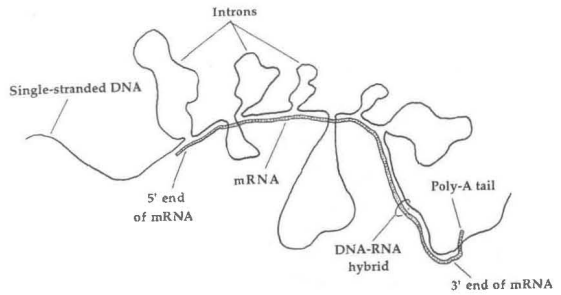


Figure 6.3 Introns can be located by artificially inducing an edited mRNA transcript to bind to a single strand of the DNA from which it was transcribed. Each section of the mRNA *hybridizes* with the section of the DNA from which it was transcribed. The leftover loops of DNA are the introns; the corresponding sections of the mRNA were spliced out during posttranscriptional processing. (Redrawn from Arms and Camp 1987, 205.)

In eukaryotic cells, such as those of plants, animals, and fungi, the primary transcript of mRNA is further processed by the enzymatic machinery of the cell. "Tails" and "caps" are added to the mRNA transcript, and extensive portions are cut out and discarded. These discarded segments are referred to as *introns.* The segments that are retained and spliced together to form the final mRNA are known as *exons.* Alternative splicing patterns, of which there are many examples, make it possible to produce several final mRNA transcripts from the same DNA sequence. Finally, it has recently been discovered that some primary mRNA transcripts may be edited in detail, one base at a time, before proceeding to the translation phase. Some mRNAs are edited (by converting a C into a U) so as to produce a *stop codon* in the middle of the transcript so that it codes for a different, shorter protein. Notice, already, the complex, indirect, and equivocal nature of the relationship between the DNA sequences in chromosomes and their phenotypic consequences. In what follows, this message gets yet more support.

Translation from mRNA to protein occurs with the help of devices called *ribosomes* and a second form of RNA, *transfer RNA (tRNA),* which acts as a physical link between the amino acids that are the constituents of proteins and the final mRNA transcript. The ribosome moves along the mRNA, creating chains of amino acids that are then folded into proteins. The genetic code is *degenerate*—different codons specify the same amino acid—but it is never *ambiguous:* the same codon is never linked via its various intermediaries to more than one amino acid.

Even in the accompanying technical boxes we have barely scratched the surface of the complex machinery that mediates between DNA and protein construction. But the take-home message is simple: One DNA sequence can

Box 6.4 The Genetic Code

In a rather dubious metaphor, the genome of an organism is often regarded as a coded description of the organism as a whole. But there is a sense in which it really is a code for the proteins in the organism. Proteins are made from a stock of twenty different amino acids. So the basic function of the genetic code is to specify those amino acids in the right sequence. Each amino acid is specified by a three-base sequence drawn from the mRNA bases uracil, adenine, guanine, and cytosine. But since there are sixty-four ($4 \times 4 \times 4$) possible three-base sequences, there are sixty-four different codons, and hence there is *degeneracy* in the coding system. That is, more than one three-base sequence can code for the same amino acid. AUG codes for the amino acid methionine, and since all newly synthesized proteins start with methionine, AUG functions as the *start codon.* But there are three *stop codons* (UGA, UAA, and UAG), and sixty-one codons that code for amino acids. The degree of redundancy ranges from leucine, coded by six sequences (UUA, UUG, CUU, CUC, CUA, CUG) to tryptophan, coded only by UGG. An additional source of degeneracy is the differences between the coding mechanism of the genes in the cell nucleus and those in the mitochondria. UGA is not a stop codon for mitochondrial DNA. But though the code is degenerate, it is never *ambiguous:* one codon is always mapped onto one, and only one, amino acid.

be input to mechanisms that yield different protein sequences. So though the RNA codon/tRNA anticodon/amino acid system is not ambiguous in that anticodons always attach to the same codon and are always attached to the same amino acid, this is merely an unambiguous subsystem within a system fraught with ambiguity. It is a system that maps the same DNA sequences onto different proteins and, further, to different phenotypic outcomes. The one-to-many character of the DNA/phenotype relationship is even more apparent when we consider the regulation of genes—the mechanisms that turn them on and off.

6.4 Gene Regulation

A skin cell and a brain cell are very different from each other—and they and their descendants will probably remain that way. Tissue differentiation is often a one-way street. Once a cell lineage has become a lineage of one particular tissue type, it usually does not revert to some earlier, more plastic form. Early cell biologists took very seriously indeed the idea that the hereditary material was divided up between the different tissue types, so that the hereditary material for skin went to skin cells and the hereditary material for nerves went to nerve cells, and only the sex cells retained a full copy (the *mosaic theory*). But this hypothesis was disproved. In fact, most cells have the complete genome. The differences between them are due to mechanisms of gene regulation and cell line heredity. These mechanisms are being discovered at an impressive rate, and any attempt to summarize them here would be quickly out of date. Furthermore, even the mechanisms already known are far too varied and complex to describe in a text of this kind. So we offer here some very general observations about these mechanisms, which will play a role in the arguments over reductionism.

Box 6.5 Reading the Code

Only one strand of the DNA double helix is read, since DNA can be read from only one end, the 5′ end. From this strand, an mRNA strand is constructed as each base in the 5′ strand is paired with its complementary base. The codons of the genetic code are sequences in this mRNA strand.

Actual protein synthesis takes place at structures called *ribosomes* in the cell cytoplasm. Transfer RNAs (tRNA) are chunks of RNA in the cell cytoplasm, each consisting of three bases. Each tRNA binds at one end to a specific amino acid and at the other, again via the base pairing mechanism in which each base has a unique partner, to the mRNA at the ribosomes. So each codon of the mRNA is recognized by a tRNA *anticodon* with an amino acid attached. As the amino acids are lined up and attached by tRNA to mRNA at the ribosome, they form bonds with their neighbors, and a sequence of amino acids is built. This sequential order, in the right molecular context, specifies the protein.

As we have noted, the genetic code is degenerate. Where it is degenerate, it is usually so at the third position in a codon. So mutations that affect the third position are often *silent:* they have no effect on the amino acid being made. But they can affect the rate at which it is made. For the rate at which the code is read depends on the stock of available reading chemicals. The building of a protein depends on the supply of tRNA in the cell cytoplasm. The range of tRNAs that a cell synthesizes helps to determine the assembly of amino acids into proteins.

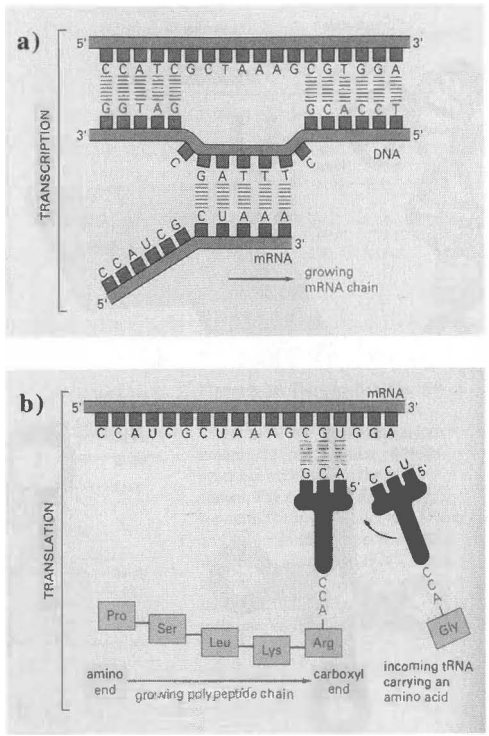


Figure 6.4 Transcription and translation. (a) Each base of DNA is transcribed into the corresponding RNA base, producing a strand of messenger RNA. (b) Each codon of the mRNA transcript matches the anticodon on one end of a transfer RNA. The other end of each tRNA carries a specific amino acid. Ribosomes (not illustrated) move along the mRNA, translating it into a chain of amino acids—one of the polypeptide chains of which proteins are composed. (Adapted from Alberts et al. 1994, 108.)

The expression of a DNA sequence can be controlled at almost every stage of the process between the sequence itself and the functional protein it produces. Various posttranscriptional mechanisms operate on the mRNA transcript, as we have already described. Each of these offers a point of intervention affecting the final protein. Splicing and editing affect the type of protein translated, and other processes affect the quantity translated. Since two forms of RNA play an essential role in this process, the rate of translation of mRNA to protein is affected by the availability of tRNAs (which are synthesized from other regions of the genome) and by the rate at which mRNAs are degraded so that they become unavailable for translation.

Gene regulation through control of transcription has been known for much longer than these posttranscriptional processes. The most intensively studied and best understood form of gene regulation involves *regulatory sequences,* short stretches of DNA that bind to certain characteristic classes of *regulatory proteins.* Transcription of DNA depends on an enzyme called *RNA polymerase,* which splits the double helix and begins the transcription process.

Regulatory proteins affect the ability of RNA polymerase to bind to the regulatory sequences and initiate transcription.

The DNA sequences that are transcribed into mRNA are preceded by *promoter* sequences, to which RNA polymerase attaches itself. In prokaryotic cells, such as the bacterium *E. coli,* regulation is relatively simple. Regulatory sequences lie adjacent to the promoters. Some of these bind *repressors,* negative regulatory proteins that interfere with RNA polymerase binding. Others bind *transcription factors,* positive regulatory proteins that facilitate RNA polymerase binding. In eukaryotic cells, such as those of plants and animals, things are much more complex. The RNA polymerases that transcribe eukaryotic genes typically require a whole complex of transcription factors to be present for them to initiate transcription. This complex machinery enables the overall rate of transcription to be influenced by many different factors, contributing to the ability of eukaryotic cells to create many different cell types from the differential activation of a single genome.

Transcription in eukaryotes is also affected by the organization of DNA into chromosomes. Chromosomes are composed of a material called *chromatin,* which consists mainly of DNA and structural molecules called *histones.* The long DNA molecule can be condensed in various ways in chromatin structures. The most compressed forms are known as *heterochromatin,* and DNA in these forms cannot usually be transcribed into mRNA. This form of gene regulation plays a well-known role in female mammals. Females have two X chromosomes, one of which is rendered inactive by being compressed into a dense, heterochromatic *Barr body.*

A cell's pattern of gene activity is frequently passed on to descendant cells that originate from it by mitosis. Some cells pass on to their descendants not only the genome, but a complex of extragenomic factors that they have acquired during the process of tissue differentiation and which cause them to express those genes, and only those genes, needed in that tissue. The inactivation of the second X chromosome just described is a case in point. One or the other X chromosome is randomly chosen to become a dense, inactive Barr body in the founding cells of certain cell lineages. All cells in the lineage inherit the same pattern of inactivation. So female organisms are genetic mosaics, with different sets of X chromosome genes acting in different tissues.

Another mechanism of cell line heredity is *DNA methylation,* in which parents attach methyl groups to the DNA of their sperm or eggs. In vertebrates and some invertebrates, additional methyl groups can be attached to the bases cytosine or guanine. Heavily methylated sequences are not transcribed. An enzyme called *DNA methyltransferase* copies the methylation

pattern when DNA is replicated. A gene that was turned off by methylation in the parent cell is thus turned off in daughter cells.

Overall, then, the same lesson as before applies: the connection between DNA sequence and phenotype is not just indirect, it's many-to-many. The effect of DNA sequences on phenotype is modulated by mechanisms that turn genes on and off, mechanisms that affect the rate at which "on" genes are transcribed and translated, and mechanisms that determine which proteins are eventually built from a transcribed sequence. So the relationship between DNA sequence and phenotype is many-to-many with a vengeance.

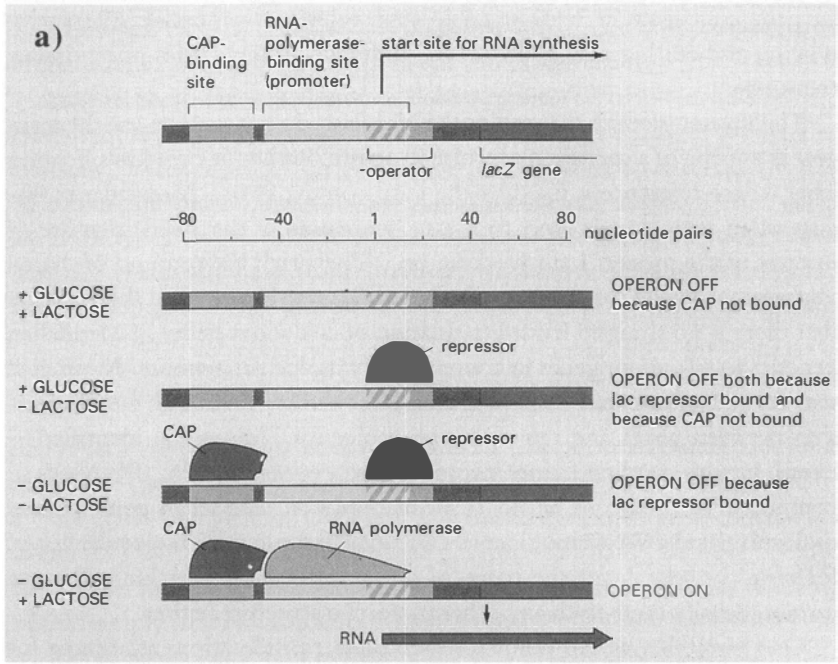
6.5 Are Genes Protein Makers?

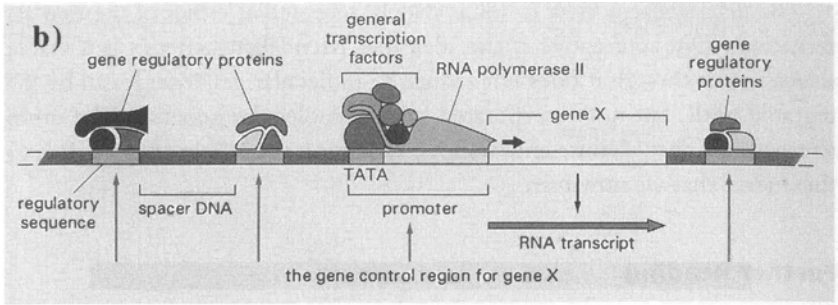
Just as early research in genetics was guided by the ultimately untenable "one gene–one trait" concept, early research in molecular genetics was guided by a "one gene–one protein" concept. The classic molecular gene concept is a stretch of DNA that codes for a single polypeptide chain. We have not tied any of the foregoing discussion to this important gene concept, referring instead simply to DNA sequences. That is because the classic molecular gene is a highly problematic unit in light of the very processes of transcription and translation that we have just described. The original intent of the classic molecular gene concept was to identify a gene with the DNA sequence from which a particular protein is transcribed, via mRNA. But even ignoring the fact that reading frames may overlap, the relationship between DNA sequences and protein chains is many-to-many, not one-to-one. To see this, consider the role of regulatory sequences. These sequences do not themselves code for a protein (so, if they are independent genes, the classic molecular gene concept is already in trouble). But unless at least some noncoding regulatory machinery is included along with the transcribed sequence, the presence of a gene does not explain the presence of the relevant protein. If all regulatory and promoter sequences were adjacent to the transcribed sequences they regulate, we could regard the whole sequence as a single gene. Bacterial genetics more or less works this way. The *operon* of bacterial genetics consists of one or more transcribed sequences and their immediately adjacent promoter and regulatory sequences (see figure 6.5a). In eukaryote gene regulation, however, regulatory sequences may be distant from the sequences they regulate and may be involved in regulating many sequences. Genes coding for transcription factors may be arbitrarily distant from the genes transcribed, perhaps because eukaryote DNA can loop around to bring transcription factors bound to distant regulatory sites close to a gene being transcribed (see figure 6.5b,c). Other problems for the classic molecular gene concept arise because of posttranscriptional processes. Alternative splicing and editing may make several different proteins from one primary transcript.

The upshot, then, is that molecular biologists do not seem to use the term *gene* as a name of a specific molecular structure. Rather, it's used as a floating label whose reference is fixed by the local context of use. Molecular biologists often seem to use *genes* to mean "sequences of the sort(s) that are of interest in the process I am working on." Their rich background of shared assumptions makes this usage perfectly satisfactory. However, it then follows that there is no straightforward translation of talk about genes in Mendelian genetics to talk about genes in contemporary molecular genetics. As we shall see, the antireductionist consensus makes the further point that the relationship between genes and the structures molecular biology has identified—exons, introns, reading frames, promoters, repressors, mRNA, tRNA—is so complex that there can be no clean mapping of Mendelian genes to *any* molecular kinds. We cannot identify Mendelian genes with molecular genes, for *molecular gene* is not the name of one specific molecular kind. But we cannot identify them with any other molecular structure, either.

One possibility at this point is to see these considerations as arguing for the displacement of Mendelian genetics by molecular biology. Contemporary geneticists have proposed, for example, that the dominant/recessive distinction be replaced by a gain of function/loss of function distinction. Recessive phenotypes, according to this idea, are typically the result of an organism being saddled with two copies of a defective gene. The recessive phenotype develops because something does *not* happen. Moreover, though genes can lose function for more than one reason, this would still be a more cohesive molecular-level explanation than the dominant/recessive one. One problem with this revisionary idea is that the gain of function/loss of function distinction depends on how wild-type gene functions are defined. *Oncogenes,* for example, are dominant and represent an inappropriate (from the organism's point of view) gain of function leading to cancer. However, it might be argued that the true "function" of an oncogene is to remain silent in certain cell types, and it is a *loss* of function in its control system that leads to its *gaining* the ability to be expressed at the wrong time (Chambers, personal communication). A more straightforward problem is that some loss of function mutations are dominant; for example, in cases in which the loss of one allele lowers protein production below a critical threshold level.

Classic accounts of reduction acknowledged that the old theory would often have to be "corrected" before it could be reduced. The old theory might contain elements unconnected with its explanatory successes (but





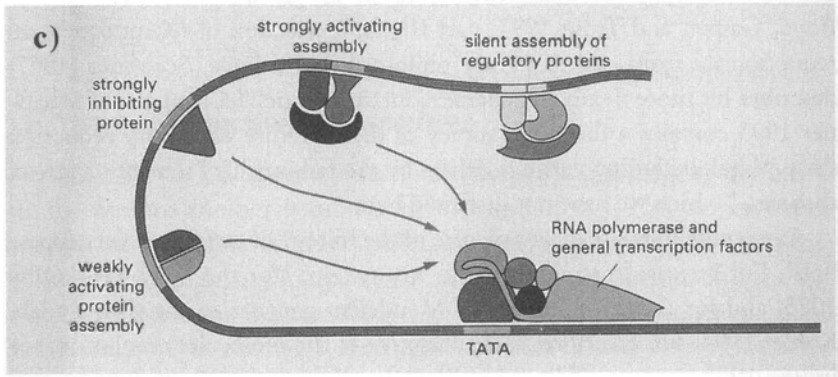


Figure 6.5. Gene regulation. (a) The *lac* operon in the bacterium *E.coli* was the first gene regulatory mechanism to be understood. The operon consists of a transcribed sequence plus one *promoter* site and one *repressor* site adjacent to the start site for mRNA transcription. The regulatory proteins bound at these sites respond to glucose and lactose concentrations. The regulatory factor CAP (catabolite activator protein) helps the enzyme RNA polymerase to open the double helix and initiate transcription of the DNA. The repressor protein stops this process from proceeding. This causes RNA polymerase to be bound and transcription to commence only when there is a low concentration of glucose and a high concentration of lactose. The resulting gene product metabolizes lactose into glucose. (b) Gene regulation in eukaryotes is much more complicated than in bacteria. The TATA box is a sequence of T-A and A-T base pairs close to the start site for mRNA transcription. This sequence binds a collection of general transcription factors (involved in the same process for many other genes). Regulatory regions specific to the particular gene may exist far upstream of the TATA box, or even downstream of the transcribed sequence. (c) The regulatory proteins bound to these distant regulatory regions are thought to be brought into contact with those bound to the TATA box by looping of the DNA. (Adapted from Alberts et al. 1994, 420, 424, 429.)

perhaps responsible for its explanatory failings) that could not be derived from the new theory and the bridge principles. However, if too much correction were required to effect a reduction, this process would no longer be one of theory reduction, but of theory replacement—that is, of displacement rather than incorporation or integration. No one would dream of "correcting" the phlogiston theory of combustion to say that phlogiston is taken up in combustion rather than lost in combustion and then claiming to reduce the phlogiston theory to the oxygen theory. The phlogiston theory was just wrong, and the oxygen theory displaced it. In one view, the "corrections" in Mendelian genetics that would be required in order to reduce it to molecular genetics are so large that this project resembles the frivolous proposal to "reduce" phlogiston to oxygen. So, just as the Churchlands take the irreducibility of psychological kinds to neural kinds to show that there really are no such things as beliefs, Rosenberg takes the irreducibility of classic genes to molecular genes to show that molecular genetics displaces Mendelian genetics:

> Molecular genetics reveals that there is no one single kind of thing that in fact does what Classical genetics tells us (classical) genes do. In this respect of course molecular genetics replaces classical Mendelian genetics. (Rosenberg 1997, 447)

One of the best current texts offers a summary review of "classical genetics," beginning with the claim that in classic genetics a gene is "a functional unit of inheritance usually corresponding to the segment of DNA coding for a single protein product" (Alberts et al. 1994, 1072). This, of course, is the classic *molecular* gene concept; Mendelian genes have disappeared from the map altogether.

The displacement view is not as widely accepted as either of the two alternatives. One alternative is the idea that Mendelian genetics is a viable science even though it does not reduce to molecular genetics: it can be integrated with, but not incorporated within, molecular genetics. The other alternative is that, despite appearances, reduction is possible after all. It is to these ideas that we now turn.

Further Reading

6.1 The classic account of theory reduction is given by Nagel (1961). See Boyd, Gasper, and Trout 1991, part III, for a selection of recent papers on reductionism from contemporary philosophy of science. Schaffner (1967) describes his more flexible "general reduction model." Chapter 9 of Schaffner 1993 contains a thorough survey of the literature on theory reduction since Nagel, including versions driven by the fashionable "semantic view of theories," which we have not discussed here.

As we note in the text, our picture of the history of genetics is very superficial. For serious treatments of this history, see (for the early days) Olby 1985, and for the development of Mendelian genetics in the fruit fly lab, Kohler 1994. For a very readable narrative of the molecular revolution, see Judson 1997. A more philosophically focused account of the history is given by Depew and Weber (1995). Mayr wears a historian's hat too: part III of Mayr 1982a is his account of the development of genetics. Dupré (1993) and Rosenberg (1994) present an interesting contrast. They essentially agree in thinking that the classic accounts of theoretical unification fail to fit biology. But whereas Dupré develops a case for thinking that the program of unification and the metaphysics that underlies it is wrong-headed, Rosenberg argues that unreduced biology cannot be regarded as an objective account of the way the world is. So their work is relevant throughout this and the next chapter.

6.2–6.5 The history of the gene concept is complex and controversial. Falk (1984, 1986) discusses its many transformations. Portin (1993) presents a good recent treatment. As usual, Keller and Lloyd (1992) provide a good entrée into the literature; Maienschein overviews the history of the concept, and Kitcher surveys its current uses. An authoritative source on modern molecular biology is Alberts et al. 1994. For accessible introductions to these difficult issues, see Moore 1993 or Mayr 1982a.