Understanding Genetics:
DNA, Genes, and Their Real-World Applications
Parts I & II
Professor David Sadava

THE TEACHING COMPANY ®

Table of Contents
Understanding Genetics: DNA, Genes, and Their Real-World Applications

David Sadava, Ph.D.

Pritzker Family Foundation Professor of Biology
The Claremont McKenna, Pitzer, and Scripps Colleges

David Sadava is the Pritzker Family Foundation Professor of Biology at Claremont McKenna, Pitzer, and Scripps, three of The Claremont Colleges. Professor Sadava graduated from Carleton University as the science medalist, with a B.S. with first-class honors in Biology and Chemistry. A Woodrow Wilson Fellow, he received a Ph.D. in Biology from the University of California at San Diego.

Following postdoctoral research at the Scripps Institution of Oceanography, he joined the faculty at Claremont, where he has twice won the Huntoon Award for Superior Teaching, as well as receiving numerous other faculty honors. He teaches undergraduate courses in general biology, biotechnology, and cancer biology, and has been a visiting professor at the University of Colorado and at the California Institute of Technology.

A visiting scientist in oncology at the City of Hope Medical Center, Professor Sadava has held numerous research grants and written more than 55 peer-reviewed scientific research papers, many with his undergraduate students as coauthors. His research concerns resistance to chemotherapy in human lung cancer, with a view to developing new, plant-based medicines to treat this disease.

He is the author or coauthor of five books, including *Plants, Genes, and Crop Biotechnology* and the recently published eighth edition of a leading biology textbook, *Life: The Science of Biology*.

Understanding Genetics: DNA, Genes, and Their Real-World Applications

Scope:

Perhaps no branch of knowledge has been as exciting over the past 50 years as genetics, the scientific study of heredity. The DNA double helix, discovered in 1953, is one of the great icons of science in our society, rivaling the atom in its pervasiveness in our culture. Like the atom, DNA symbolizes not just scientific knowledge that in this case doubles every few years, but immense implications for humanity. Knowledge of DNA and genetics is radically impacting the two important applications of biology to human welfare—medicine and agriculture. In addition, studies of genes are changing the way we look at ourselves and the other organisms with which we share the Earth.

Lectures One through Three describe genetics as we knew it before DNA. People have long wondered how characteristics are passed on through generations. Before the mechanism of inheritance was investigated with the methods of experimental science, there were many ideas. Some scientists and philosophers thought that only the male (or female) contributed inheritance to offspring. Others proposed that, while the sexes contributed equally to the offspring, whatever it was that each contributed blended together permanently after the union of male and female. The Austrian monk and scientist Gregor Mendel put an end to these notions in 1866 when he published the results and interpretation of years of deliberate and careful experiments on pea plants. He clearly showed not only that the sexes contribute equally to offspring, but also that the genetic determinants, or genes, were particulate and retained their individuality after mating. Almost 40 years later, his results and conclusions were independently verified by other scientists. As biologists began to study life at the microscopic level, in the tiny cells that make up every organism, the genes were located in structures inside of every cell called chromosomes. Mendel and his successors, and the cell biologists looking at chromosomes, gave geneticists the tools to work out the rules of inheritance. But the exact nature of what determined inherited characteristics remained unknown.

The nature of genes and how they are arranged and expressed is described in Lectures Four through Nine. The search for what the gene really is made of quickly focused on DNA. Circumstantial evidence favored it: DNA was in the right place at the right times in the right amounts. But these were correlations—and as such are not valid scientific evidence. A set of experiments on many different organisms provided the proof that DNA was the molecule of heredity. Soon afterward, the double-helix model of DNA was described, as was the elegant way in which it duplicates itself when cells reproduce. The next issue was to determine how DNA as the gene is expressed. The information in each gene is usually expressed as a protein. These complex molecules have vital roles in the organism. They provide structure and can act as enzymes to speed up chemical transformations inside cells. With thousands of such proteins, there are thousands of genes. Coming full circle, the gene-protein relationship was described in a genetic code that is virtually universal in all life on Earth. By the end of the 20th century, biologists were able to determine the information content of every gene of an organism (the genome) and were on the way to describing the functions of these genes. Genome projects, ranging from bacteria to humans to rice plants, are accumulating information at a dizzying rate.

Lectures Ten through Fifteen describe how our knowledge of DNA and its expression can be used to manipulate it for our own purposes. The creation of gene splicing (recombinant DNA) in 1973 was an epochal event in the history of genetics. It is now possible to take any gene from any organism and transfer it to another organism or cell. We are no longer confined to breeding within a species. The applications of this technology range from coaxing bacteria to make human proteins for use as pharmaceuticals (for example, insulin) to inserting genes into “super bugs” that can clean up environmental pollutants (for example, oil spills). The development of new ways to manipulate DNA, including the rapid amplification of DNA from just a single cell, the ability to make and sequence DNA quickly and efficiently in the lab, and the ability to shut down gene expression at will have opened up new horizons in research and application that were unheard of until just recently. The well-publicized use of DNA in forensic identification is just one example.

Lectures Sixteen and Seventeen deal with evolution, the great unifying idea of biology. DNA and its expression can explain Charles Darwin’s idea that there are changes in genes that create variations, some of which have an advantage for reproduction, and that natural selection results in those variants being passed on to the next generation. So there is change through time—evolution. Comparisons of organisms’ DNAs are now possible, and these show clear relationships between organisms and descent with modification. DNA changes even provide an evolutionary “molecular

clock" through which organisms can be related.

Lectures Eighteen through Twenty-Two describe how knowledge of DNA and genetics is leading to a new kind of medicine called molecular medicine. With increasing precision, we are gaining an understanding of what goes wrong in illness and devising specific ways of dealing with it. Genetic screening is allowing for timely interventions, before a disease caused by abnormal genes takes effect. Our understanding of how the immune system fights invaders to the body has been enhanced and harnessed. Even cancer, the great scourge of modern societies, is under intense investigation—and new therapies are either here or on the horizon. At the frontier of molecular medicine are gene therapy, in which deliberate genetic changes to human cells can be used for treatment, and cloning and stem cells, in which new tissues can be made to replace ones that are damaged. The need for these approaches is there, the potential is impressive, and both knowledge and applications are progressing.

Lectures Twenty-Three and Twenty-Four deal with agriculture, that other application of biology to human welfare. Genetics has been part of food production for thousands of years, and increasingly effective ways to breed better crops have been developed as knowledge about genetics has accumulated. This culminated in the impressive achievements of agriculture in the last 50 years, in which food production has more than kept pace with population growth. DNA and biotechnology have the potential to take this progress even further. These applications are just starting, and they include introducing specific genes to crops to make them grow better, need fewer pesticides, and produce more nutritious foods.

As with any technology that is so radically different, biotechnology has its critics, on both philosophical and practical grounds. There is great potential, but also risk. Biological scientists have opened up the "genetics bottle"; it is up to us if and how we use it.

Lecture One
Our Inheritance

Scope: Genetics is the science of heredity. It seeks to explain how the characteristics of living things are passed on from parent to offspring. Genetics is not genealogy, which describes family relationships. Nor does genetics fully explain all the characteristics of a plant or animal: The environment plays an important role. From our earliest days, humans have wondered about genetics, and when we began selectively breeding plants and animals for desirable characteristics, understanding the rules of inheritance became more important. At first, scientists thought that genetic determinants blended after mating, like inks. But later, it was shown by experiments that genetic determinants have a particulate nature, and that this is a chemical substance called DNA. These discoveries in chemical genetics have led to powerful new ways not just to understand how genes work, but to manipulate them for our purposes. These methods have real-world applications in medicine and agriculture.

Outline

I. Opening story: a "perfect crime."

A. We begin with the story of a "perfect crime"—or so the bank robber thought. A discarded cigarette butt led to the identification of the robber as a "cold hit" because the robber's DNA was in a database, as he had been arrested previously for armed robbery. In the past he had been arrested because of his fingerprints—a phenotype, or outward expression of an inherited characteristic, or gene. This time, his genes—DNA—gave him away.

B. There is a lot of genetics and biotechnology behind this police work. We can look at it as a series of questions:

1. What are the rules of inheritance?
2. How do we know that DNA is the gene, the material that is inherited?
3. What is the basis of the differences between individuals that are inherited?
4. How often do these differences occur, and how can that be a basis for identification?
5. How can a tiny amount of DNA on a cigarette butt be amplified for analysis in the lab?

II. Introduction to the science of genetics.

A. Genetics is the science of heredity.

1. When you consider the diversity of organisms on Earth, from the tiniest bacteria, invisible to the human eye; to the tallest redwood tree; to yourself; the diversity of living things is obvious.

2. But behind this diversity is a unity at the basic chemical level.

3. Three major ideas tie biological science together.

a. Mechanism. Many societies have believed, and many people continue to believe, that the rules that govern life are different from the inanimate universe. They propose a "vital force" to explain this. Examples of this are the *chi* of Chinese philosophy and traditional medicine. Biological science rejects this idea, proposing instead that the same rules of physics and chemistry that govern the lifeless world (e.g., rocks, the air) govern life.

b. Cell theory. Since the microscope was first used to visualize living things, biologists have agreed that cells are the building blocks of life (just as atoms are the building blocks of chemistry). All living things are made up of cells, and all cells come from other cells.

c. Evolution. Organisms are related by common ancestry, and there has been and continues to be change through time, or descent with modification. In 1859, Charles Darwin proposed natural selection as a way to explain how organisms with different characteristics change through many generations.

4. Genetics explains the rules by which characteristics, such as hair color in people, come to differ and how these differences are passed on to the next generation (inherited). The last half of the 20th century saw the nature of the genetic determinants, and how they are expressed in organisms, described in chemical terms. This is the science of molecular biology, including DNA.

B. Genetics is not destiny or genealogy.

1. Genetics is not destiny. Genes do not always predict what will happen in an organism. Is my hair really dark brown? Only my hairdresser knows for sure. What any person—or other organism—ends up looking like is determined by both heredity and environment.

2. Genetics is not genealogy. Genealogy deals with relationships between families of organisms: A family tree is genealogy. Genetics deals with the rules of inheritance: The patterns of inherited biological characteristics in the family are genetics.

III. A very short history of genetics.

A. People began selectively breeding other creatures for human purposes long ago.

1. The date palm was bred by ancient Egyptians 4,000 years ago to improve the quality and amount of the fruit.

2. Horses were bred in the Near East and Asia 3,500 years ago for speed in war and racing.

B. Deliberate breeding and human curiosity led people to ask: What are the rules of inheritance?

1. The Greek philosopher Aristotle (382–322 B.C.) proposed that the male semen had imperfect ingredients that were "organized" by the menstrual fluid during intercourse. His ideas had great influence on scientific thinking in Western cultures for over 1,500 years.

2. The Dutch scientist Antonie Van Leeuwenhoek in the 17th century looked at human sperm under the microscope and thought he saw little people, curled up. This was in accord with Aristotle's view.

3. Soon, other microscopists thought they saw little people curled up in the human egg cell. This set up an intellectual battle between "spermists" and "ovists" as to which gender passed on the hereditary characteristics to the offspring.

4. Careful observations from crosses by Dutch tulip breeders in the next 100 years showed that both sexes contribute equally to the inheritance of the offspring.

C. Does the genetic material blend, or is it particulate?

1. Attention turned to what happened to the genetic material when the sperm and egg cells unite.

2. Initially, plant breeders felt that the genetic determinants blended when put together in the fertilized

egg: A sperm from a red-flowered tulip combining with an egg of a white-flowered tulip would often produce pink offspring. And like inks that blend, these hereditary determinants would never be seen again in that offspring. It would pass on only the genetic determinant for pink color.

3. Blending reigned as the explanation for heredity until the late 1800s, when Gregor Mendel performed careful experiments using garden pea plants that showed that hereditary determinants were particulate. Mendel's results led to a rejection of blending.

4. Mendel's results were consistent with the idea that the genetic material is carried on particles in the cell called chromosomes; these had been discovered about the same time he did his work.

5. In the mid-20th century, the genetic material was identified as a chemical substance—DNA.

IV. The goals of this course.

A. The first goal is to understand the idea of genes as determinants of inheritance.

1. This begins with the rules of inheritance (Lectures Two and Three) and the chemical nature of genes as DNA (Lectures Four and Five).

2. This discussion is followed by the relationship between genes and their expression as some outward characteristics (Lectures Six through Nine).

B. The second goal is to learn about how our understanding of genes has led to powerful tools to manipulate them for our purposes.

1. This includes splicing genes together as recombinant DNA (Lectures Ten and Eleven) and using these tools to make products such as drugs and to clean up the environment (Lectures Twelve and Thirteen).

2. New ways to identify organisms have been devised, and these are used not only in forensics, as we described earlier, but also to study evolution—changes in organisms through time (Lectures Fourteen through Sixteen).

C. The third goal is to learn how our knowledge of genetics and tools of manipulating genes are being used in real-world applications.

1. These applications include molecular medicine (Lectures Seventeen through Twenty-Two) and genetically modified plants in agriculture (Lectures Twenty-Three and Twenty-Four).

2. These applications, while they hold great promise for improving life for people, are not without controversy.

Essential Reading:

Ricki Lewis, *Human Genetics*, 7th ed. (New York: McGraw-Hill, 2006), chap. 1.

David Sadava, Craig Heller, Gordon Orians, William Purves, and David Hillis, *Life: The Science of Biology*, 8th ed. (Sunderland, MA: Sinauer Associates; New York: W. H. Freeman and Co., 2008), pt. 3.

Supplemental Reading:

Anthony Griffiths, Susan Wessler, Richard Lewontin, William Gelbart, and David Suzuki, *An Introduction to Genetic Analysis*, 8th ed. (New York: W. H. Freeman and Co., 2007).

Edwin McConkey, *How the Human Genome Works* (Sudbury, MA: Jones and Bartlett, 2004).

Questions to Consider:

1. Consider yourself as you read through this course. Can you describe some of your characteristics that are clearly hereditary and some that are environmentally determined? What are some characteristics that are determined by a combination of the two?

2. What are some nonscientific beliefs about inheritance? Survey your nonscientist friends to find out what people think about the rules of inheritance.

Lecture Two
Mendel and Genes

Scope: Gregor Mendel, a monk from what is now the Czech Republic, did the experiments and came to the conclusions that founded the modern science of genetics. As the son of a farmer, who took courses in science

and mathematics at the University of Vienna, Mendel was well prepared to undertake his investigations, which occurred in 1856–1863 while he was teaching at a monastery. He wisely chose pea plants as his experimental subject, as they had clearly defined patterns of inheritance. He made careful crosses and counted the offspring plants, expressing the data mathematically. He concluded that the factors that determine inheritance, which we now call genes, are particulate and that each individual has two copies of every gene, one from each parent. Only one of these copies ends up in each sex cell (sperm or egg) to be passed on to the offspring. Genes that determine different characteristics (such as seed color and height) are inherited independently. Mendel's conclusions apply to most organisms, including humans. But since we cannot do deliberate genetic crosses as experiments, analysis of human genetics is done by pedigrees. Genes are not destiny. While genes determine the capability for a certain characteristic, the environment often has an important role in their expression.

Outline

I. Opening story.

A. It does not take an advanced degree and a well-funded position to do important science.

1. Gregor Mendel, who made one of the two key discoveries in genetics, was not a university professor or a member of a prestigious research institute. He was a monk and a teacher. His laboratory was not a large room with lots of equipment and a cadre of assistants. He worked alone on a small plot of land adjacent to the monastery.

2. Nevertheless, Mendel's background prepared him to make the discoveries he did. Born in 1822 in what is now the eastern Czech Republic, he was the son of a farmer and a mother whose father was a gardener. So he knew a lot about plants.

3. In 1843, he joined an Augustinian monastery at Brno. Assigned to teach, he was so good that the abbot sent him to the University of Vienna for further education. There he took courses in mathematics, chemistry, and biology. He learned the scientific method, which had been developed centuries before.

a. Science begins with a hypothesis: a testable idea.

b. A hypothesis is tested by an experiment under controlled conditions.

c. The experiment either proves that the hypothesis is valid or not; if the latter, the hypothesis must be revised.

d. Scientific investigations must be published so that they can be verified and extended. This takes the form of a scientific paper that is a complete description of the work. In the fields of genetics and DNA, hundreds of research papers are published every day.

4. Mendel returned to the monastery from Vienna, and from 1856 to 1863, he taught and did experimental work on garden pea plants that laid the foundation for the modern science of genetics.

5. In 1865, Mendel gave a public lecture at the local medical society on his findings and published them a year later in the society's journal. Unfortunately for science, he was "kicked upstairs" and promoted to abbot a year later, and the burden of administration precluded further investigations.

6. Also unfortunately, his paper was published in a journal that was not read by most scientists in the field, and its conclusions were so new that few would accept them.

B. Mendel's work was not appreciated until 1901, when three botanists working separately came to the same conclusions from genetics experiments that Mendel did, and then found his paper, which had just been translated from German to English.

II. Mendel performed experiments on pea plants.

A. Mendel used pea plants for several good reasons.

1. They were easy to grow in the monastery garden.

2. He could control which plants mated. The flowers have both male and female reproductive organs, so cutting off the pollen-producing male organs from a plant's flowers allowed him to brush pollen from another plant onto that one, which became the female for the cross.

3. There were well-defined strains of pea plants that were true-breeding for that characteristic, or phenotype. So, true-breeding plants that formed seeds that were spherical (plump) always formed

spherical seeds when crossed among each other; true-breeding plants with wrinkled seeds formed only wrinkled seeds, etc.

B. Mendel's experiments led to laws of heredity.

1. When Mendel crossed two true-breeding strains differing only in one characteristic (e.g., spherical × wrinkled), the first-generation offspring always showed just one of the two characteristics (in this case, all spherical). He called this the dominant characteristic and the one not there the recessive characteristic. There was no blending!

2. This was confirmed when he crossed the first-generation plants among themselves. He got the spherical and wrinkled plants back in the second generation!

3. Now Mendel's mathematics education came into the picture: He counted the plants in the second generation and found that there were 5474 spherical and 1850 wrinkled, a 3:1 ratio. He got this ratio with six other pairs of characteristics, such as tall and short plant height.

4. On the basis of his data, Mendel proposed that:

a. The genetic determinants are particulate; they are not lost after fertilization.

b. Each individual plant has two copies of the determinant (he called them "elementen," and we call them "genes"). One comes from each parent.

c. The genes for a particular characteristic (e.g., seed shape) can exist in different forms. We call these alleles. In a true breeding plant, the two alleles were the same (e.g., two alleles that result in spherical seeds—we call them "spherical" alleles). In the first-generation plants, the two alleles were different (e.g., one "spherical" and one "wrinkled"). One allele is dominant ("round") and expressed in the offspring when it is present together with the recessive allele ("wrinkled"). But the alleles never disappear, as would be the case if there was blending.

d. Only one allele is passed on to the next generation in the sperm or egg. It is random which allele a particular sperm or egg contains for that characteristic. So on average, each plant in the first generation of his seed shape cross would make half of the sperm or egg cells with the "round" allele and half with the "wrinkled" allele. This is called the "law of segregation." It is a law in science because it was shown after 1900 to be general and to apply to most other organisms, not just peas.

5. Mendel crossed peas with two sets of differing characteristics (e.g., seed shape and height) and showed that these characteristics behaved as if they were independent of each other. This led to the law of independent assortment.

6. Mendelian laws of genetics follow simple laws of probability. This can be illustrated by tossing two coins. The probability of tossing one heads is 1/2. The probability of tossing two heads at once is 1/2 × 1/2, or 1/4.

III. Mendelian laws of genetics apply to most organisms, including humans.

A. Because we cannot do genetic crosses in humans to determine genes and alleles, we rely on the past—a pedigree (or genealogy with inherited characteristics). In a pedigree following a phenotype that is expressed by a dominant allele—for example, with Huntington's disease—every affected individual has one parent who is affected. Since most unusual alleles are rare, the affected parent is most likely heterozygous (one dominant and the other recessive allele). The zygote is the biological word for a fertilized egg. So there is a 50% chance of passing this allele on to a child.

B. In a pedigree following a phenotype that is expressed by a recessive allele, an individual must have two identical alleles (that is, be homozygous) for the recessive alleles. This is the case for albinism, for example. Again, since the unusual allele is rare, an individual with albinism must have inherited one "albinism" allele from each parent, and both parents must be carriers (they are heterozygous, but normal). Recall the two first-generation heterozygous pea plants that all had spherical seeds but when crossed among themselves produced plants with wrinkled seeds.

IV. Not all inheritance follows the Mendelian ratios.

A. In codominance (e.g., the ABO blood groups in humans), both alleles are expressed in heterozygotes.

B. The ABO blood groups also are an example of multiple (more than two) alleles for a characteristic. But note that an individual only has two alleles (e.g., AB, AO, etc.).

C. Many complex characteristics are determined by the interactions of numerous genes (e.g., height in humans).

V. Genes are not destiny. The environment plays a vital role.

Essential Reading:

Ricki Lewis, *Human Genetics*, 7th ed. (New York: McGraw-Hill, 2006), chap. 4.

Benjamin Pierce, *Genetics: A Conceptual Approach* (New York: W.H. Freeman and Co., 2005), chap. 3.

Supplemental Reading:

Anthony Griffiths, Susan Wessler, Richard Lewontin, William Gelbart, and David Suzuki, *An Introduction to Genetic Analysis*, 8th ed. (New York: W. H. Freeman and Co., 2007).

Robin Henig, *The Monk in the Garden: The Lost and Found Genius of Gregor Mendel* (Boston: Houghton-Mifflin, 2001).

Questions to Consider:

1. Draw a three-generation genealogy of your family. Now, use a single characteristic—such as blue or brown eye color, presence or absence of dimples, presence of absence of a cleft in the cheek, or presence or absence of freckles—and draw the pedigree. Can you determine whether the characteristic you chose is inherited as a dominant or recessive pattern?

2. In Labrador retrievers, coat color is determined by two interacting genes (with alleles). The B allele (dominant) produces a black coat. The b allele produces a brown coat. So what color is a Bb dog? A bb dog? A second gene, E, must be present for any color to form at all; a dog that is homozygous recessive for e (that is, has ee) is white. What is the color of a BBee dog? And a bbEe dog? If you have a lab (I do—both the scientific and canine kinds!), draw its pedigree and try to figure out its genetics.

Lecture Three
Genes and Chromosomes

Scope: Animals and plants are made up of many cells. These units of biological structure and function have the fundamental characteristics of life, including chemical complexity, the ability to regulate what enters and leaves the cell, the ability to grow, and the capacity to reproduce. Even tiny bacteria are cells; viruses, which must infect cells to function, are not. The cell nucleus contains its genes, and cloning experiments show that a specialized cell has all of the genes for the organism of which it is a part. Within the nucleus, genes are carried on structures called chromosomes. These are duplicated before a cell reproduces. There are two chromosome sets—and two gene sets, as Mendel proposed—in every somatic (body) cell nucleus, but there is only one set in the sperm or egg cell, again, as Mendel proposed. With many more genes than chromosomes, each chromosome must carry many genes. For instance, the 24,000 genes in humans are carried on 23 chromosomes. In mammals, including humans, the X and Y chromosomes are a special pair. The X has many genes not present on the Y, which has the gene for maleness. Genes present on the X chromosome will always be expressed in males.

Outline

I. Opening story. You do not have to be a geneticist to figure out genetics.

A. In the Jewish religion, all males must be circumcised as a symbol of their covenant with God.

B. In the Babylonian Talmud, an ancient commentary on the Bible written about A.D. 500, the rabbis had a dilemma: Boys were bleeding to death during circumcision.

C. The rabbis noted that the boys who bled to death were all the sons of certain mothers. It did not matter who the father was. They exempted further sons from these mothers from circumcision. They made their judgment

based on observations of genetic pedigrees.

D. We now know that the boys were dying because of the disease hemophilia, which is inherited through the X chromosome of the mother.

II. Cells carry genes.

A. Complex organisms are made up of cells.

1. Two kinds of pneumonia with similar symptoms can define what a cell is. The agent causing the first kind of pneumonia can be observed under a common microscope as a rod-shaped object, about 1-millionth of a meter (1 thousandth of a millimeter) across. This infectious agent, a bacterium, can grow and reproduce in a laboratory dish if supplied with simple nutrients such as sugars. It is made up of thousands of different types of chemical substances.

2. The second type of pneumonia is caused by a much smaller agent, 1% the size of a bacterium. This infectious agent, a virus, cannot grow and reproduce on its own in the lab; it needs to infect living tissue. It is chemically simple.

3. Bacteria are cells. Viruses are not cells.

B. Cells are the basic building blocks of biology.

1. All living things are made up of cells.

2. Cells exhibit the characteristics of life: They are complex, have many chemical reactions, can determine what comes in and out, can grow, and can reproduce. While nonliving things can do some of these things, only cells can do all of them.

3. Cells are the unit of biological continuity. Complex organisms reproduce by cells (e.g., sperm and egg). So these cells must contain genes.

C. There are two types of cells.

1. Those without complex internal structures are called prokaryotic (e.g., bacteria).

2. Eukaryotic cells (e.g., liver cells) have complex internal structures for compartmentalization of functions. One of these structures is the cell nucleus.

III. The cell nucleus contains the genes.

A. Experiments define the genetic role of the nucleus.

1. Removal of the nucleus from a single-celled amoeba leads to cell death.

2. Swapping of the nucleus between species of a single-celled plant called *Acetabularia* leads to the new cell having the genetic characteristics of the nucleus it received.

3. In 1958, Frederick Steward took a specialized cell from a carrot plant, and by putting it into a special chemical environment he was able to coax the cell to act like a fertilized egg and grow into a whole plant. This cloning experiment showed that the nucleus of a specialized cell has all of the genetic determinants of the organism: It is totipotent.

B. Genes in the nucleus are carried on large, visible structures called chromosomes that are visible during cell reproduction (also called cell division).

IV. Chromosomes carry genes.

A. Chromosomes provide an explanation for Mendel's genetics laws.

1. There are two copies of every chromosome in every nonsex cell (also called somatic cells).

2. When cells divide, the chromosomes reproduce prior to separation so that each of the two new cells gets a full set of the original chromosomes.

3. There is only one copy of each chromosome pair in the sex cells (gametes).

B. Specific genes have been linked to specific chromosomes.

1. The full set of chromosomes with genes is the genome. Because there are thousands of genes (humans have 24,000) and few chromosomes (humans have 23 pairs) there must be many genes on each chromosome. The genome is like a library: The chromosomes are its volumes, and the genes are

paragraphs.

2. In mammals, including humans, the X and Y chromosomes are an exception to the rule that chromosomes come in complete pairs. The X has many genes, but the Y has only a few, most notably the gene that determines male sex. So any person with a Y chromosome is a male.

3. Males have only one copy of most genes carried on the X chromosome (they are not on the Y). So any recessive allele on the X chromosome will be expressed in males. In females, a dominant gene on the other X chromosome would mask the expression of the recessive allele.

4. The most common form of hemophilia (see opening story) and red-green colorblindness are caused by recessive alleles carried on the X chromosome. So these characteristics are much more common in males than in females.

5. For the majority of human characteristics whose genes are carried on nonsex chromosomes (also called autosomes), there is no preferential sex distribution.

6. Primary sex, the formation of sperm or eggs, is determined genetically by the presence of a gene called SRY on the Y chromosome.

7. Secondary sex, the appearance of male or female body parts such as developed breasts, body hair, muscular development, etc., is also genetically determined, but by a different set of genes. In this case, the genes involved are responsible for hormone signaling involving such familiar hormones as testosterone (males) and estrogen (females).

Essential Reading:

Benjamin Pierce, *Genetics: A Conceptual Approach* (New York: W. H. Freeman and Co., 2005), chap. 4.

David Sadava, Craig Heller, Gordon Orians, William Purves, and David Hillis, *Life: The Science of Biology*, 8th ed. (Sunderland, MA: Sinauer Associates; New York: W. H. Freeman and Co., 2008), chap. 10.

Supplemental Reading:

Martin Brookes, *Fly: The Unsung Hero of 20th-Century Science* (New York: HarperCollins, 2001).

Matt Ridley, *Genome: The Autobiography of a Species in 23 Chapters* (New York: HarperCollins, 2006).

Questions to Consider:

1. The cloning of Dolly the sheep involved transferring the nucleus of a specialized sheep cell (the donor) into a sheep egg cell (recipient) whose nucleus had been removed. Following chemical stimulation, the egg acted like it had been fertilized and divided to form an embryo. The embryo was surgically implanted into a foster mother (surrogate) and grew to a lamb, Dolly. Was Dolly genetically identical to the donor, recipient, or surrogate? What does this result indicate about the genetic capacity of the nucleus?

2. If cells are the unit of life, where did the first cells on Earth come from? Two scientific ideas are that they came from another place in the solar system (life may have existed on other planets and/or their moons) or that cellular life evolved from chemicals on Earth. Which of these ideas is most intellectually and emotionally satisfying to you?

Lecture Four
The Search for the Gene—DNA

Scope: The link between smoking and lung cancer, long proposed from clinical case histories and population studies, only became incontrovertibly strong after a substance from cigarette smoke was shown to cause genetic damage on DNA. To act as the genetic material, DNA must contain information, must duplicate (replicate) accurately and change (mutate) occasionally, and must be able to be expressed. There was early circumstantial evidence from its location in the nucleus and amount that DNA might be the genetic material. But this was only demonstrated conclusively by dynamic experiments. First, scientists isolated DNA from one genetic strain of bacteria, introduced it into a second, different strain, and showed that the second strain was genetically transformed into the first one. Later, this was repeated on many other kinds of cells from many different organisms. DNA was even shown to be the genetic material of viruses, being injected into host cells to convert them into virus manufacturers.

Outline

I. Opening story.

A. The smoking gun is in DNA.

1. In 1948, Ernst Wynder, a medical student at Washington University in St. Louis, attended an autopsy of a man who died of lung cancer. He noticed that the lungs were black, and when he saw the man's medical record he knew why: The man was a long-term cigarette smoker of two packs a day.

2. Curious, Wynder looked at many more such cases over the next two years. Using a case-control approach, he found that comparing 649 cases (lung cancer) with 600 controls (no lung cancer), the rate of smoking was 40 times higher in the cases. This is the science of epidemiology.

3. At the same time, British scientist Richard Doll was doing a cohort study. He followed doctors who smoked and those who did not over many years and found that far more smokers developed lung cancer.

4. However, these studies were controversial. For instance, people who had lung cancer might have been exposed to more air pollution or have a poor diet.

5. Finally, in the 1990s an incontrovertible link was found between smoking and lung cancer. As we will see later in the course, cancer is in part caused by genetic damage. A gene called p53 gets mutated (permanently changed). In cancerous lung tissues, these changes occur only at certain locations in the gene, called "hot spots."

6. When tobacco smoke was analyzed by sophisticated chemical machines to separate out its many products, one, called benzpyrene (BP) stood out. In the lung, BP is converted to a highly active form called BPDE, and it is this that binds to and damages genes. When scientists looked at the p53 gene in the lungs of smokers, they found BPDE binding to it and causing damage—right at the hot spots for genetic change. This closed the circle of evidence.

B. The p53 gene, like all other genes in most organisms, is made of DNA.

II. DNA fits several requirements for the genetic material.

A. It must be able to contain lots of information. Organisms have many genes (humans have 24,000), and these can be and are encoded by DNA. Few other molecules have the ability to carry so much information.

B. It must be able to replicate (a scientific word for duplicate) in an error-free fashion. A human has 60 trillion cells. These came from a single cell, the fertilized egg. Imagine the consequences if replication had a lot of errors that were then perpetuated. As we will see, DNA replication is impressively accurate.

C. It must be able to be expressed as the phenotype. As we have seen, genes encode the capability for phenotype: A gene in peas gives them the capacity to form spherical seeds, for example. It turns out that the substances ultimately responsible for phenotype, such as pea seed shape, are proteins. Cells have machinery to express DNA, usually as proteins in the phenotype. Of course, this expression can be modified by the environment.

D. It must be able to change. This might seem odd, since we have said that replication must be error free and that damage, as in lung cancer, is harmful. But variation is the spice of life and the raw material for evolution by natural selection. So some errors in genes are a good thing over the long term of generations. As we will see, DNA has ability to mutate.

III. There was circumstantial evidence for DNA as the genetic material.

A. It is in the right place.

1. A Swiss physician, Friedrich Miescher, had isolated DNA from nuclei of white blood cells in 1868, calling it "nuclein."

2. Later evidence showed that the nucleus has the genes.

B. It was there in the right amounts.

1. A dye was developed that quantitatively stained DNA red. When cell nuclei were examined, each species had its own unique content, all somatic cells of that species had the same amounts, and the sex cells had half as much of the DNA as somatic cells. All of this would be expected if DNA was the gene.

No other molecule in the nucleus (e.g., proteins) had these quantitative characteristics.

2. But circumstantial evidence is not cause and effect in science. We need experiments to prove the hypothesis that DNA is the gene.

IV. Experiments proved that DNA is the genetic material.

A. Frederick Griffith got sidetracked as he investigated a vaccine for bacterial pneumonia in 1928. There were no antibiotics yet.

1. There were two genetic (pure breeding) strains of the bacteria: S, smooth coat (virulent, caused pneumonia in mice); and R, rough coat (not virulent).

2. He wanted to see if heat-killed S could be used as vaccine to immunize against pneumonia. Heat-killed S did not cause pneumonia. But it did not work as a vaccine in mice.

3. So he tried combining it with a booster of live R cells (these do not cause pneumonia). Once again, there was no vaccine effect. But to his surprise, the mice injected with heat-killed S and live R got pneumonia! And he found live S cells in their bodies. The dead S cells had genetically transformed the live R cells into live S cells!

4. Now it became straightforward to find out what the chemical nature of this genetic transformation principle was. It took 15 years. In 1944, Oswald Avery and colleagues identified it as DNA.

5. DNA transformation in all cells has been possible since the 1970s and is now widely used in laboratories.

B. DNA is the genetic material of viruses as well.

1. Just like eukaryotic cells, prokaryotic bacterial cells get attacked by viruses. When a bacteriophage (virus) attaches to a bacterial cell, it injects something into the cell that takes over the cell's chemistry, directing it to be a virus manufacturer. Half an hour later, the cell bursts, releasing several hundred viruses that then go hunting for more cells to infect. Clearly, the genetic material is injected to take over the cell.

2. As noted earlier, viruses are very simple. Bacteriophage have only DNA and proteins. In 1952, Alfred Hershey and Martha Chase set out to show that it was the DNA, not the proteins, that got injected into the bacteria and must be the genetic material.

3. They looked for a way to label proteins and DNA differently. They used two different radioactive atoms, one that was specific for DNA and the other for proteins.

4. Bacteriophage with labeled DNA or protein were allowed to attach to the bacteria and inject genetic material into the cells. Then suspension of cells was agitated in a blender to shake the phage off the cells. The scientists looked inside the cells for some injected phage material. Only labeled DNA entered the cells; labeled protein did not. This showed that the DNA must be the genetic material of viruses as well as cells.

Essential Reading:

Michael Cain, Hans Damman, Robert Lue, and Carol Yoon, *Discover Biology,* 3rd ed. (New York: W. W. Norton, 2007), chap. 12.

David Sadava, Craig Heller, Gordon Orians, William Purves, and David Hillis, *Life: The Science of Biology*, 8th ed. (Sunderland, MA: Sinauer Associates; New York: W. H. Freeman and Co., 2008), chap. 11.

Supplemental Reading:

Anthony Griffiths, Susan Wessler, Richard Lewontin, William Gelbart, David Suzuki, and Jeffrey Miller, *An Introduction to Genetic Analysis*, 8th ed. (New York: W. H. Freeman and Co., 2007).

Benjamin Lewin, *Genes VIII* (Upper Saddle River, NJ: Pearson Prentice-Hall, 2005).

Questions to Consider:

1. There have been many claims of using DNA to transform complex organisms. Early on in the 1950s, scientists reported that injecting white ducks with DNA from black ducks gradually turned the latter ducks black. Others reported that feeding naive animals DNA from the same species of animal that had learned a task made the naive animals smarter. You can even find DNA tablets at stores in the vitamin section. These experiments have never

been reliably repeated. Why don't they work?

2. Until 1970, it was very hard to reliably transform cells using DNA. This phenomenon was confined to just a few species, including bacteria that had been used by Griffith and Avery. Then, some scientists discovered that the reason that cells would not take up added DNA was that the surfaces of cells are negatively charged, and so is DNA. Things with the same charge repel one another, so DNA would be repelled from the cells' surface. To get around this, the scientists added a salt to the cells and DNA. This neutralized the charges, and cells readily took up any DNA they were given. This simple method revolutionized genetics, as it made possible all of modern biotechnology. The scientists did not patent this method, and so neither they nor their university have received royalties for their discovery. Jonas Salk, who made the first polio vaccine, likewise did not patent it, saying that he made his discovery for all people and not for himself. Today, scientists take out patents on many of their discoveries. Do you think they should?

Lecture Five
DNA Structure and Replication

Scope: DNA is a remarkably stable molecule. It has been isolated from fossils and even from a man who was encased in ice for 5,000 years. DNA was first purified from nuclei of white blood cells in 1868. When the nucleus was established as the location of the genetic material, attention focused on DNA, and this was intensified by finding DNAs with highly variable sequences of the four bases, A, T, G, and C. Watson and Crick evaluated two lines of evidence from chemistry (in DNA, the proportions of A=T and G=C) and physics (the crystal of DNA is in a helical form) and came up with a double-helical model with A fitting opposite T and G opposite C. This model fits the requirements for genetic material in terms of information content and mutability. DNA is replicated accurately by semiconservative replication where each parental strand serves as a template for a new strand.

Outline

I. Opening story.

A. The Ice Man and his DNA.

1. In 1991, hikers in the Alps between Austria and Italy spotted a corpse frozen in the ice, with some tools and a dagger nearby. Four days later, the body, now known as Ice Man, was removed and brought to the University of Innsbruck in Austria, where scientists found that it was 5,000 years old. An arrow had killed him, and he froze into the ice.

2. Remarkably, his DNA was preserved inside his bones. When its information was analyzed, it was similar in its information content to the DNAs of Europeans currently living in the European alpine regions. That DNA lasts so long is testimony to its durability. Its structure is testimony to something else.

B. DNA as secular icon.

1. An icon is an image or symbolic representation, often with sacred significance. The 20th century saw the emergence of two great icons that symbolized science in the public mind: the atom, with its whizzing electrons, and DNA—deoxyribonucleic acid—with its double helix.

2. People talk about DNA all the time, as a shorthand. An ad for a financial services company tells customers that it "understands the DNA of business." A perfume called DNA is the "essence of life." A media software system is the "DNA server." A clothing line called DNA Divewear advertises "genetic design-extreme behavior." And of course, a biotechnology company has the stock market symbol DNA.

3. DNA is a subject for artists as well. Salvador Dali used the double helix. A portrait of Nobel laureate Sir John Sulston has parts of his DNA inside bacteria cells. Brazilian artist Eduardo Kac used the chemical building blocks of DNA to translate verses from the Bible; viewers can change the DNA and the verses it represents at will. Playgrounds have DNA-shaped slides. And of course, many DNA-based sculptures adorn the lobbies of laboratories and pharmaceutical companies.

II. Early studies of DNA suggested its structure.

A. Friedrich Miescher was the first to isolate DNA.

1. The son of a doctor, Miescher graduated from medical school at the University of Basel. Because of partial deafness, he did not pursue clinical medicine but instead chose a career in research. He began working on the chemistry of the cell nucleus with the German biochemist Ernest Felix Hoppe-Seyler. Using pus from dressings at the nearby hospital, in 1868 Miescher isolated nuclei from white blood cells and then extracted a complex substance rich in phosphorus and nitrogen that he called "nuclein." A student later called it "nucleic acid." It turned out to be DNA. Hoppe-Seyler was so surprised by this novel substance that he personally repeated all Miescher's work before letting him publish it.

2. Nucleic acid was a curiosity until it was realized that the genetic material resides in the nucleus. Now it became a candidate for the genetic material, along with protein, which is also abundant in the nucleus. In the 1890s, Albrecht Kossel found that nucleic acid was actually a polymer: a long chain of "beads" or monomer units of chemicals. In nucleic acid, these monomers are the nitrogenous bases adenine (A), thymine (T), guanine (G), and cytosine (C).

B. Chemists analyzed the makeup of DNA.

1. In New York, Phoebus Levene spent the first 30 years of the 20th century studying DNA. He found that the four bases (A, T, G, and C) were not linked to each other in a direct chain. Rather, they were attached to a backbone of sugar and phosphate. The sugar is deoxyribose; hence the name *d*eoxyribo*n*ucleic *a*cid (DNA).

2. Levene proposed that DNA was a very, very, very, long chain consisting of repeating units of ATGC: something like ATGCATGCATGCATGC, etc. If he was correct, and he thought he was, this meant that DNA lacked the variable information content to be the genetic material. Levene was a towering intellectual figure, and so most people believed him. Attention focused on proteins as the gene.

3. During the late 1940s, another great biochemist took a stab at DNA. Erwin Chargaff, at Columbia University, found that DNA was not just a repeating four-unit structure but was highly variable: ATCGTTCAATACGATGACTT, etc. The information content was there. This came at the same time as the identification of DNA as the genetic material in experiments on bacteria and viruses.

4. Chargaff made another discovery whose significance came to be fully appreciated soon after: For each species he examined, there was a constant ratio of the proportions of bases, and that A=T and G=C. So for one species, it might have 20% A and 20% T, and 30% G and 30% C. Another species might have 18% A and 18% T, and 32% G and 32% C.

5. Now the problem was fitting the observations together.

III. The structure of DNA is a double helix.

A. Physical chemists found that DNA was a helix.

1. Physical chemists study the arrangement of atoms in three dimensions when atoms combine to make molecules. For example, water (H_2O), with three atoms, can form regularly shaped molecules in liquid and ice crystals. Larger molecules such as DNA have millions of atoms. Is there any regularity to their arrangement?

2. When X rays are shone at a large molecule in solid form, the rays get bounced around before they pass through the molecule. By looking at the pattern that comes out the other side, chemists can figure out if there is a regular pattern to the arrangement of atoms in the molecule. Rosalind Franklin and Maurice Wilkins at King's College, London did X-ray analyses of DNA and found a regularity in the shape of a helix.

B. Watson and Crick solved the structure of DNA.

1. Nearby, in Cambridge, the young American geneticist James Watson and somewhat older British physicist Francis Crick looked at the X-ray evidence of a helix, and the notion that A=T and G=C, and came up with an inspiring idea: Maybe the structure of DNA was a double helix with two chains of bases opposite one another.

2. They made chemical models of the four bases and found that A just fit into T and G just fit into C; no other fits worked. Crick's wife, an artist, drew the double helix, and they published their structure in the

spring of 1953. It was immediately accepted for its elegance as well as its science.

IV. The DNA double helix fulfills the requirements for the genetic material.

A. DNA has huge potential information content.

1. The polymer of DNA is very long, ranging in size from 1 million base pairs in some bacteria to 100 million base pairs in humans.

2. These large molecules containing genetic information are what biologists now call chromosomes.

B. DNA can be accurately replicated.

1. Watson and Crick said in their 1953 article that their model suggested a way that the genetic material could replicate. This was confirmed a few years later in living cells and involved several steps:

2. The two parental DNA strands separate, exposing their bases.

a. New bases come in a pair with the exposed bases. Thus an AT base pair on the parental strand gets exposed GA and CT. These then get paired up with appropriate bases GA-*CT* and *GA*-CT, and voila: There are two new strands with the same sequence as the parent.

b. Because each new double-stranded DNA molecule has one strand from the parent and one newly built one, this is referred to as semiconservative DNA replication. Contrast this with conservative replication, which, like Xerox copying, has one brand new double-stranded molecule and the old double-stranded parent still intact.

C. DNA can be mutated. Errors do occur in replication (one in a million). They are mostly repaired by a proofreading mechanism, but sometimes are not. Errors change the base sequence in the new DNA, and so cause a change in information content of the new cell that is formed. This is mutation, the raw material for evolution.

D. DNA is expressed as the phenotype. This is the topic of the next lecture.

Essential Reading:

Francis Crick, *What Mad Pursuit* (New York: Basic Books, 1988).

Benjamin Pierce, *Genetics: A Conceptual Approach* (New York: W. H. Freeman, 2007), chap. 10.

James D. Watson, *The Double Helix* (New York: Athenaeum, 1968).

Supplemental Reading:

Anthony Griffiths, Susan Wessler, Richard Lewontin, William Gelbart, David Suzuki, and Jeffrey Miller, *An Introduction to Genetic Analysis*, 8th ed. (New York: W. H. Freeman and Co., 2007).

Benjamin Lewin, *Genes VIII* (Upper Saddle River, NJ: Pearson Prentice-Hall, 2005).

Questions to Consider:

1. Trace the changes in thinking about whether DNA was the genetic material from Miescher's time until Watson and Crick. This is an excellent example of the scientific process, where an idea is put out to the community and gets accepted, rejected, or modified with time.

2. It is important at this stage to understand base pairing and DNA replication. One strand of DNA has the sequence AGCTTCTGGATCTTTAGTCAGTGTAC. Write out the other strand that is paired with this one. Now, outline the steps by which this DNA is replicated: Draw the two separated strands and add new bases (in a different color, say red). Having one old and the new strand in the newly made DNAs is called semiconservative duplication. (Conservative duplication, which does not happen, would be like a copy machine, where there is an old piece and a new piece of paper.)

Lecture Six
DNA Expression in Proteins

Scope: Proteins are polymers (long chains of chemical "beads") composed of monomers called amino acids. With 20 amino acids, there are many possible orders of them in proteins up to 1000 units long. Because the amino acids vary in their chemical properties, proteins will also have these properties. Spider silk proteins are an example. One type of silk protein is in flat interlocking sheets that provide strength. Another type of silk is

made up of a protein that stretches and is more flexible. Proteins can fold to expose surfaces not only for structure, but also for functioning as enzymes: catalysts essential to speeding up chemical transformations in living things. Protein is the major expression of a gene, called the phenotype. The relationship between DNA and its protein expression was worked out with mutant strains of simple organisms such as bacteria.

Outline

I. Opening story.

A. Spider silk is an amazing protein.

1. A spider web has three roles: It is home, it is where the spider mates, and it is where it captures food. The web must be strong and stretch, yet not break, nor wobble so much as to get out of control. Webs are thinner than hair, and indeed strong: stronger than steel and more elastic than nylon.

2. Spider webs are made of a protein called silk. The silk proteins are giant molecules—macromolecules composed of millions of atoms of hydrogen, carbon, oxygen, nitrogen, sulfur, and phosphorus hooked together in a specific way. Silk is made in glands at the rear of the animal and spun out as fibers.

B. Proteins are polymers made up of monomers—beads on a chain. Like the nucleotides that make up DNA, amino acids make up proteins. Only instead of four monomers, there are 20 different amino acids. Each protein has its own composition of these 20, and spider silk has a unique collection in a specific order that determines its structure and function.

1. There are two kinds of silk.

a. The strong fibers have strands of protein that fold into flat sheets, with ratchets that fit them together (like Lego blocks) so they won't slip apart.

b. The flexible fibers have strands of protein that allow them to curl around and slip by one another.

2. These structures are determined by spider genes in their cell nuclei.

II. What are proteins?

A. Proteins are polymers composed of monomers.

1. Twenty amino acids are linked together in a specific order for each protein. The amino acids are important in human nutrition. They contain amino groups (NH_2), which humans can only get from eating other creatures. Some organisms have the genetic capacity to get nitrogen from the air (79% N_2) and turn it into amino groups. Humans cannot breathe in steak!

2. Humans lack the genetic capacity to make 8 of the 20 amino acids; these must be eaten as part of proteins in the diet. This has great implications in human nutrition, to which we will return in Lecture Twenty-Three.

B. The specific order of the amino acids in a protein (they range from 10–1,000 amino acids long) determines how the protein will fold in three dimensions.

1. This is a chemically spontaneous process in the cell once the protein is made.

2. Protein shape is very sensitive to the environment: Heat changes it, often irreversibly (boiling an egg).

3. Protein shape presents specific surfaces to its surroundings (see spider silk, above). These can fit other substances like a lock and key (or baseball and glove).

III. Proteins have essential functions in the organism.

A. Some proteins are structural.

1. The shape of proteins allows for structural specificity. See spider silk for examples. Other structural proteins include: hair (keratin), muscle (actin and myosin), antibodies (immunoglobulin), and connective tissue (collagen).

2. These all have a unique composition of the 20 amino acids in a unique order.

3. They are only made by certain tissues at certain times (you don't have hair growing on your eyeball).

B. Some proteins act as enzymes.

1. The shape of proteins allows for a surface where chemical reactions can occur. Example: Consider

DNA replication. In humans, a single DNA molecule (chromosome) has millions of base pairs. These must be replicated in just a few hours during cell division. If you put DNA into a text tube (or cell) and give it the building blocks for new DNA, the various steps of replication will eventually happen just because the building blocks will randomly bump into the DNA as it happens to unwind. But this will take thousands of years!

2. To speed things up, we might heat the cell to make the DNA unwind faster and the building blocks move faster. This might work, but we know that living tissues generally cannot tolerate heat. The reason is that their proteins will change shape (like the cooked egg).

3. Another way to speed things up is to provide a "workbench" to grab and line up the DNA and building blocks. This needs a specific surface, and proteins provide this. They are catalysts called enzymes.

4. The enzyme DNA polymerase binds to DNA and its building blocks in such a way as to speed up the polymerization many millionfold.

5. There is an enzyme for virtually every one of the thousands of chemical transformations in a cell (e.g., digestion of food).

6. Proteins are the phenotypic expression of genes.

IV. Mutations in model organisms relate genes to proteins.

A. The strategy of mutations proves cause and effect.

1. In biology, mere correlation does not prove causation. If we want to show that A causes B, we need to have a way of experimentally disrupting A and then seeing that B does not occur to prove our hypothesis.

2. Genetics is a way to do this. If there is an organism that has a genetic mutation that disrupts the gene for A, and B does not occur, then our hypothesis is correct.

B. Simple organisms can be genetically manipulated and tested in the laboratory.

1. Bacteria such as the gut bacterium *Escherichia coli* and molds such as the bread mold have only one set of genes (all mutations show up), and it is easy to grow them in the lab. They are especially useful for showing nutritional mutations.

2. Unlike humans, bacteria can make all 20 amino acids. Sometimes, a genetic mutation occurs so that the strain of bacteria no longer makes the amino acid. Now it's mutated, like us, and like us needs that amino acid to grow. Scientists can compare the chemistry of these mutant bacteria with those normal ones that can make the amino acid. And when they do, they find that an enzyme is missing: one of the enzymes that catalyzes the transformation of chemicals to that amino acid. So the normal gene must be expressed as that enzyme.

3. A bacterium arose that genetically could not make DNA. (It just sat there and did not reproduce.) The missing enzyme in this case was DNA polymerase.

4. These observations, repeated many times for many organisms, led to the one gene–one enzyme concept.

5. And since almost all genes are proteins, it's "one gene–one protein." The mystery of genotype and phenotype is solved: The gene is DNA, and the phenotype is protein.

Essential Reading:

Anthony Griffiths, Susan Wessler, Richard Lewontin, William Gelbart, David Suzuki, and Jeffrey Miller, *An Introduction to Genetic Analysis*, 8th ed. (New York: W. H. Freeman, 2007).

David Sadava, Craig Heller, Gordon Orians, William Purves, and David Hillis, *Life: The Science of Biology*, 8th ed. (Sunderland, MA: Sinauer Associates; New York: W. H. Freeman and Co., 2008), chap. 12.

Supplemental Reading:

Horace Freeland Judson, *The Eighth Day of Creation: Makers of the Revolution in Biology* (Woodbury, NY: Cold Spring Harbor Laboratory Press, 1996).

Benjamin Lewin, *Genes VIII*, (Upper Saddle River, NJ: Pearson Prentice-Hall, 2005).

Questions to Consider:

1. Spider silk is as strong as Kevlar, the strongest synthetic material known. What uses can you suggest for the silk (for humans)?

2. With 20 amino acids and proteins up to 1000 amino acids long, the possibilities for different amino acid sequences and protein shapes seem virtually endless (20^{1000} is a very big number). Yet there are far fewer protein sequences—and hence, shapes—in the living world. Thinking about evolution by natural selection, why might this be so?

Lecture Seven
Genes, Enzymes, and Metabolism

Scope: Genetically inherited diseases such as phenylketonuria and alkaptonuria led to the one gene–one enzyme hypothesis in humans. Biochemistry is the expression of the phenotype, and enzymes are the actors that determine conversions in biochemical pathways. Metabolism is the sum total of the biochemical conversions in a cell, tissue, or organism. Cells obey the same physical laws of thermodynamics as the rest of the inanimate universe. Anabolism is using energy in chemical conversions to make energy-rich substances, and catabolism is releasing energy for use in other conversions. In both cases, energy is not created or destroyed, only changed. In biology, where each step in a biochemical pathway is determined by the presence of an enzyme encoded by a gene, the biochemical capabilities of an organism are genetically determined.

Outline

I. Opening story: Phenylketonuria is a human genetic disease.

A. Dr. Asbjorn Folling solved a medical mystery.

1. A Norwegian mother had watched her children, a 6-year-old daughter and a 4-year-old son, get progressively sicker over time. By mid-1934, both were profoundly mentally retarded. Her family physician told her that he could do nothing to help the children and advised her to see Dr. Asbjorn Folling, a medical specialist in both chemistry and mental retardation.

2. Folling thought that the children's symptoms might be related to blood sugar, so he tested their urine by adding a solution of ferric chloride, which turns from brown to purple when there are a lot of ketones from sugar. To his surprise, the solution turned green!

3. Folling first had to rule out that whatever was causing this color change came not from the children's own body chemistry but from something they consumed. So he asked the mother to stop giving them any of the special medicines they were taking for a week and retested. Once again, the green color appeared.

4. Folling now used his chemistry knowledge to extract the mystery substance from the kids' urine and identified it as phenylpyruvic acid. He named the disorder with the rather cumbersome Latin "imbecillitas phenylpyruvica." Later, it was called phenylketonuria (PKU).

5. Over the next decade, Folling saw other mentally retarded people who excreted (put into urine) phenylpyruvic acid and had PKU. Over half of them were siblings, like the original pair. In all four families, both parents were mentally normal and did not excrete phenylpyruvic acid. Knowing Mendelian genetics, he realized that PKU must be inherited as a recessive.

B. Other human genetic diseases were known.

1. In 1896, English physician Archibald Garrod saw patients with a rare disorder known as alkaptonuria, where the urine turns black when exposed to air. Because the disease seemed to occur most often in children of first-cousin marriages, he concluded that alkaptonuria is a genetic disease caused by a recessive allele.

2. Garrod took this one important step further. Enzymes had just been discovered as essential biological catalysts, and he proposed that the error in alkaptonuria was due to a lack of an enzyme that converts a molecular product of protein breakdown. Now knowing about the identity of genes and proteins, he came up with the one gene–one protein hypothesis. He called alkaptonuria an "inborn error" of biochemistry.

3. "I believe that no two individuals," wrote Garrod, "are exactly alike chemically any more than structurally." This was an amazing and true prediction.

II. Protein is the phenotype, and biochemistry is the expression of the phenotype.

A. A biochemical pathway describes the sequential conversions of substances in the body.

1. Each step in a pathway is catalyzed by a specific enzyme, encoded by a gene.

2. A single pathway involves the enzymes that are deficient in PKU and alkaptonuria. That pathway involves the conversions of an amino acid, phenylalanine.

3. Because phenylalanine is an amino acid and part of proteins, and because phenylalanine cannot be made by humans, we need to take it in our food. Proteins containing phenylalanine (e.g., in corn) are digested to amino acids. These are absorbed into the blood and transported to the liver.

4. In the liver, some of the phenylalanine is converted to other substances. Each step in this pathway is catalyzed by an enzyme. The enzyme step in converting phenylalanine to tyrosine is called phenylalanine hydroxylase. The product of the conversion is tyrosine, which is phenylalanine with a hydroxyl group.

5. To summarize: The substrate is phenylalanine, the product is tyrosine, and the enzyme is phenylalanine hydroxylase.

6. Other genetic disorders along the pathway include, besides alkaptonuria, the more familiar albinism.

B. Each enzyme is coded for by a gene. A gene can be mutated such that the expressed enzyme is not functional. When people are heterozygous (one normal and one mutant allele), the normal allele determines the good enzyme, and enough of it is made to provide normal function. So the parents who saw Dr. Folling were not mentally retarded and did not excrete phenylpyruvic acid into their urine.

III. Metabolism is the totality of biochemistry.

A. Metabolism is the sum total of all of the chemical transformations that occur in a biological entity: a cell, tissue, or organism. It is the expression of the phenotype.

1. Much of metabolism involves energy. Thermodynamics is the study of energy in physics and chemistry. There are two laws of thermodynamics.

a. The first law states that energy is not created or destroyed, only changed. This means that in any metabolic conversion, energy gets changed. For example, we eat sugar. The sugar gets converted to carbon dioxide (CO_2), which we breathe out. Sugar has a lot of energy stored in its chemical bonds. CO_2 has much less energy. So where does the energy go in this conversion? It gets released as heat and gets transferred to a conversion in the body that needs energy, like making fat!

b. The second law states that in any conversion, disorder increases and usable energy is lost (recall the old Woody Allen film in which the child gets depressed because the sun is "running down" and will be out of usable energy … in billions of years). What this means in biology is that it takes an input of energy to make complex substances from simple ones (or for a teenager to clean his or her room).

2. Biology is part of the physical universe. Life obeys these laws, just as the sun does. Scientists do not believe that there is a vital force that is different in life than elsewhere in the universe. This unity of nature is called "mechanism."

B. Anabolism is the term for building up, and catabolism is the term for breaking down in biochemistry.

IV. There are several rules for metabolism.

A. Metabolism occurs in small steps to release energy in small, usable packets. For example, consider sugar being converted to CO_2. This releases a lot of energy, about 50 times more than is needed for any single anabolic conversion in the cell. So if it happened in one step, most of the energy released would be lost as heat. Instead, it happens in over 40 steps, and energy is released in about a dozen of them.

B. Because each metabolic conversion is determined by a gene, the pathways present in an organism are genetically determined. For example, bacteria have genes that code for enzymes that can take the carbon, hydrogen, and oxygen atoms in simple sugar and rearrange them to make ascorbic acid, vitamin C. We can't do this because we lack the enzymes. Other bacteria have the ability to take cellulose from wood or paper and convert it to sugar so they can use that for energy. Again, we lack the gene (and enzyme) to do this. Of course,

there are many things we can do that bacteria can't. The point is, metabolism means phenotype, and this is determined by genes.

Essential Reading:

Jeremy M. Berg, John Tymoczko, and Lubert Stryer, *Biochemistry*, 6th ed. (New York: W. H. Freeman, 2006).

Michael Cain, Hans Damman, Robert Lue, and Carol Yoon, *Discover Biology*, 3rd ed. (New York: W. W. Norton, 2007), chap. 7.

Supplemental Reading:

Katherine Denniston and Joseph Topping, *Introduction to General, Organic and Biochemistry*, 4th ed. (New York: McGraw-Hill, 2003).

Thomas Devlin, *Textbook of Biochemistry with Clinical Correlations*, 6th ed. (Hoboken, NJ: Wiley-Liss, 2006).

Questions to Consider:

1. Think about your daily activities. Which are catabolic and which are anabolic? If you could get the genes for it, what biochemical pathway that some other organism has would you want?

2. PKU and alkaptonuria both occur because of enzyme deficiencies that result in accumulation of certain toxic substances. Both diseases are in the same pathway, which metabolizes the amino acid phenylalanine. Can you suggest a nutritional treatment to prevent the symptoms of these diseases? What would be the problems with nutritional treatment?

Lecture Eight
From DNA to Protein

Scope: The gene (DNA) resides in the cell nucleus, but its expression (protein) occurs outside the nucleus at a molecular "workbench" called the ribosome. A gene is transcribed into messenger RNA (mRNA), which is a copy of the gene that is sent to the ribosome. A specific sequence of DNA bases near the gene called the promoter determines whether a gene will be transcribed and therefore expressed in the phenotype. Substances such as hormones that cause cells to specialize may act as promoters to enhance transcription of certain genes. The gene and its mRNA contain a sequence of nucleotide bases. This sequence is translated at the ribosome into a sequence of amino acids to make a protein. The intermediary between the mRNA and amino acids is transfer RNA (tRNA), which carries the amino acid to the ribosome and binds to the appropriate sequence on mRNA. This sequence is determined by the genetic code. So the order of amino acids in a protein is determined by the order of codons in mRNA, which is determined by the order of nucleotides in DNA. A mutation is a change in the base sequence of DNA. A single base-pair change can lead to an amino acid change, which leads to a change in the function of a protein in the phenotype.

Outline

I. Opening story: toxic revenge on gene expression.

A. Journalist Georgi Markov was waiting for a bus in London.

1. A man carrying an umbrella (not an unusual sight) brushed up against him, and Markov felt a pinprick in the leg as the man's umbrella poked him. Within a few hours he felt weak, and two days later he was dead.

2. The police found a tiny pellet where the umbrella had poked him, and chemical analysis revealed it was coated with ricin, a poison extracted from the seeds of the castor bean plant. I remember taking castor oil as a child, with my mother telling me it would clean out my stomach. Fortunately, ricin is a protein that does not dissolve in the oil; otherwise, I might have had an excuse to avoid the oil's awful taste.

3. Since the Markov assassination in 1978, ricin has been in the news. It was probably used by Iraq in its war with Iran in the 1980s and was the partial basis of possible weapons of mass destruction that the government of Iraq allegedly stockpiled prior to 2003. In 2002, ricin was found in caves in Afghanistan abandoned by Al-Qaeda. In 2004, traces of it were found in a mailroom at the U.S. Senate building, prompting evacuation. Less than 1/10,000 of an ounce can kill a person.

B. Ricin inhibits gene expression.

1. As Markov found out the hard way, ricin is very toxic to people. The plant makes it as a storage protein in its seeds. As they germinate, the ricin is broken down (catabolized) to amino acids for the growing plant embryo to use.

2. Unfortunately, the ricin protein has enzyme activity. It catalyzes the modification and breakage of an essential molecule in the ribosome, the part of the cell that is used to make proteins. Inactivation of the ribosome essentially inactivates gene expression, and the cell dies.

II. Translating the information in the gene (DNA) to its expression (protein).

A. We have seen that a gene is expressed as a protein. This poses a chemical problem: These are very different molecules.

1. Gene: DNA with a sequence of nucleotide bases (A, T, G, C).

2. Expression: protein with a sequence of amino acids (20).

B. The locations of the gene and its expression are different.

1. Gene: in the nucleus.

2. Protein: made in the cytoplasm (outside the nucleus).

III. Information signals for protein synthesis.

A. DNA sends a copy of its instructions to the ribosome.

1. There are many genes on a chromosome: For example, a human chromosome has thousands of genes.

2. A cell only expresses certain genes as proteins (you don't make hemoglobin in your hair, and don't make hair in your red blood cells).

3. Determining which genes to express is the central issue of cell differentiation.

4. A copy of the gene is sent to the ribosome (similar to a copy of architectural plans sent to the job site). That copy is mRNA (m = messenger). It is made by an enzyme, RNA polymerase, and is a base-paired copy:

If a region of DNA is AAGTATGTTAGCCGT
TTCATACAATCGGCA,

then if the bottom strand is copied to RNA, it will be (RNA has U, not T): AAGUAUGUUAGCCGU.

This is called gene transcription.

B. A signaling sequence on DNA attracts the RNA polymerase for copying.

1. The signal is called a promoter sequence.

2. The RNA polymerase is guided to the promoter by a host of other substances that bind to it: These are the factors that cause a cell to specialize.

3. Example: Hemoglobin is made by developing red blood cells. A hormone is made to stimulate this. The hormone enters the cells and goes to the nucleus, where it acts at the promoter for hemoglobin, directing RNA polymerase to copy that gene (transcribe it).

C. The mRNA goes to the ribosome, where it sits and waits for the appropriate amino acid.

1. Amino acids are each brought to the ribosome by a different RNA called tRNA (t = transfer).

2. The tRNA can bind by base pairing to the mRNA. This is done by triplets of bases. For example, for the amino acid lysine, where there is AAA in mRNA, the appropriate lysine-carrying tRNA has UUU (U bonds with A in base pairing).

D. The genetic code is the key to translating the DNA information to amino acid information.

1. The code is in three base letters: AAA in mRNA means lysine tRNA will bind and be put at that spot in the growing protein chain.

2. It is very important to relate this to the gene (DNA). If the mRNA has AAA, then it must have come from a DNA that is

AAA
TTT

3. So the order of bases in the gene determines the order of amino acids in the protein.
4. There are 64 codons of 3 bases each (4 × 4 × 4), and almost all code for amino acids.
5. The "meaning" of each codon was determined by clever test tube experiments using synthetic mRNAs.
6. The code is virtually universal for all of life on Earth: This is vital for understanding genetics, evolution, and biotechnology. We have a common language.

E. It takes a minute for a cell to make a protein with 500 amino acids. There are hundreds of ribosomes in a cell. Ricin blocks the ribosome in eukaryotic cells. Antibiotics such as tetracycline and neomycin block the ribosome in prokaryotic cells (bacteria) by binding to proteins. Genetically determined antibiotic resistance is a phenotype often caused by a gene mutation resulting in an altered ribosomal protein.

IV. The genetic code explains mutation.

A. Consider phenylketonuria, a genetically inherited disease caused by a defective enzyme, phenylalanine hydroxylase.

1. In its common genetic variant:

Normal: Mutant:

Protein is 451 amino acids Protein is 451 amino acids(but not functional)

Amino acid 408 is arginine Amino acid 408 is tryptophan

Codon in DNA at 408 is Codon in DNA at 408 is

CGG TGG

GCC ACC

mRNA from bottom mRNA from bottom

CGG UGG

Amino acid is arg Amino acid is trp

2. Note that this is one base-pair change in a gene that is thousands of base pairs long. The human genome has 2 billion base pairs, and this one change leads to a protein with the incorrect shape to do its job.

B. There is a new way to define a gene and a mutation.

1. Genetic mutation can now be defined chemically as a change in DNA base.
2. Genetic capacity can now be defined as the presence of a DNA sequence that codes for a protein with a specific function.

Essential Reading:

Horace Freeland Judson, *The Eighth Day of Creation: Makers of the Revolution in Biology* (Woodbury, NY: Cold Spring Harbor Laboratory Press, 1996).

Harvey Lodish, Arnold Berk, Paul Matsudaira, Chris Kaiser, Monty Krieger, Matthew P. Scott, Lawrence Zipursky, and James Darnell, *Molecular Cell Biology*, 5th ed. (New York: W. H. Freeman, 2005), chaps. 4 and 11.

Supplemental Reading:

Benjamin Lewin, *Genes VIII* (Upper Saddle River, NJ: Pearson Prentice-Hall, 2005).

David Sadava, Craig Heller, Gordon Orians, William Purves, and David Hillis, *Life: The Science of Biology*, 8th ed. (Sunderland, MA: Sinauer Associates; New York: W. H. Freeman and Co., 2008).

Questions to Consider:

1. The ribosome is the workbench for protein synthesis. The molecules that make up the ribosome in bacteria differ from those in eukaryotes, although their roles are the same in making a protein. This is an example of convergent evolution. Antibiotics such as streptomycin, neomycin, and tetracycline block protein synthesis at only the prokaryotic ribosome (Otherwise they would kill the patient, like ricin!). Can you explain how these antibiotics might work and why they are specific?

2. What are the implications of a common genetic code for all organisms for the origin and evolution of life?

Lecture Nine

Genomes

Scope: The Human Genome Project grew out of a desire to understand radiation damage to Japanese survivors of atom bomb explosions. DNA sequencing methods can determine the complete base sequence of 800 base-pair fragments. However, most chromosomes are much larger, so a way to order the fragments was needed. One way was to pinpoint specific short DNA marker sequences at intervals throughout the genome, then sequence the fragments and order them; this was developed by the publicly funded Human Genome Project. The other method was to sequence first, and then use computers to order the fragments. Both sequences were completed in 2003. Only 2% of the entire human genome is its 24,000 genes that get expressed as the phenotype. Over half of it is noncoding repeated sequences, and most of the other half appears to be noninformational. There are several ways that these relatively few genes can end up making a greater diversity of proteins. Many other genomes have been sequenced. A minimal genome of about 400 genes has been described for a prokaryote, and synthetic biologists are trying to make this genome in the lab, possibly creating life.

Outline

I. Opening story: Genome sequencing uses new technologies.

A. The impetus for genome sequencing arose from radiation damage.

1. From the early 1900s, when genetic mutations were first studied in fruit flies, scientists found that various chemicals could increase the mutation rate—they were mutagens. Recall that DNA can also spontaneously mutate due to errors when it is replicated. Among the mutagens was ionizing radiation. This was found the hard way by people working in uranium mines, and radium "painters" for dials on luminous watches developed mutations that caused cancer. By the 1930s, this type of mutagenesis was studied under controlled conditions in the lab, with a clear dose-mutation relationship.

2. At the end of World War II, the U.S. exploded atomic bombs on the Japanese cities of Hiroshima and Nagasaki. Hundreds of thousands were killed and many more exposed to radiation as fallout. These people and their descendants have been intensively studied for any increases in mutations, both in somatic cells (leading to cancer) and in sex cells (leading to genetic diseases in the next generation).

3. This has been a paradigm for studies of environmental mutagens.

B. By 1980, methods were developed to sequence DNA, about 800 base pairs at a time.

1. Up to then, the scientists studying genetic damage in the Japanese survivor group looked at genetic damage by its effects on the phenotype (e.g., cancer, inherited abnormalities). Now they realized that the best way to analyze genetic damage was to actually determine the DNA sequence and look for differences between people exposed to radiation and those not exposed.

2. In 1984, Nobel laureate Renato Dulbecco suggested that the entire human genome be sequenced. The U. S. Department of Energy, which oversaw the radiation damage project, was the first sponsor.

II. The human genome was sequenced in two ways.

A. The initial challenge was to get DNA signposts.

1. The problem: We can sequence DNA fragments that are 800 base pairs long. But each chromosome in humans is about 100 million base pairs long. So we cut the chromosome into 800 base-pair pieces (Can you figure out how many pieces that makes?), sequence each one, and then line them up. This seems straightforward.

2. But the problem is, how do we line up the sequenced fragments? If every word in this printed Guidebook was cut out and all the words put on the floor, could you line them up to make the sentences I have written?

B. Two methods revealed the signposts.

1. The first way was sponsored by the government. It set out to identify short sequences of one to several base pairs that would "mark" each segment of DNA. This would create a "marker map" along the chromosome. So if a sequenced fragment had that marker, it must be at a known location on the

chromosome. It turns out that human DNA has short, often repeated sequences at intervals along each chromosome. It took over 10 years to find them and another 2 years to sequence the fragments. It was painstaking work by thousands of scientists led by Francis Collins.

2. The second way was to break up the chromosome into fragments, sequence them, and then have a computer look for markers and arrange the fragments. A key was to do staggered breaks. For example, consider the sentence:

THIS COURSE IS GOOD.
It can be broken into four-letter fragments:
THIS COUR SEIS GOOD.
Or, if the break is internal:
TH ISCO URSE ISGOI OD, etc.

3. This approach was undertaken by a scientist-entrepreneur, Craig Venter, and relied on the development of bioinformatics, the use of very sophisticated computer programs to analyze a mountain of DNA sequence data. These programs were developed in the 1990s. It took less than a year to sequence the genome in this way.

III. There are several types of information from DNA sequences.

A. Open reading frames are regions of genes that code for proteins.

B. Amino acid sequences of proteins can be deduced from sequences by the genetic code.

C. Gene control sequences such as promoters can give information on regulation of gene expression.

IV. Nonhuman genomes have been sequenced.

A. Comparative genomics provides important information.

1. This information relates to gene functions: If a gene is also present in nonhuman genome and its function is known there, this may point to its function in humans or other organisms.

2. A gene sequence may have a new function in a different organism. This allows for evolution by natural selection and is an important argument for it.

B. The nonhuman genomes sequences have a wide range.

1. The first genome sequenced (in 1995) was the bacterium that causes meningitis: *Haemophilus influenzae*, which has 1,830,137 base pairs and 1743 genes.

2. Many new genes have been discovered in prokaryotic genomes, for example, genes for virulence in bacteria that cause typhus (*Rickettsia*); genes for cell surface attachment in the bacterium that causes TB (*Mycobacterium*); and genes for attachment to plants in nitrogen-fixing bacteria that colonize plants (*Rhizobium*) and are important in ecology and agriculture.

a. The yeast genome was the first simple eukaryote: 12 million base pairs, 6000 genes. It is a model organism for eukaryotes. It has the basic bacteria set for metabolism but also genes for building cell compartments and targeting proteins to them.

b. The nematode worm genome (1000-celled organism) has 97 million base pairs and 19,000 genes. It is a model for complex eukaryotes, with different tissues. It has genes for cell differentiation and signaling between cells.

c. The rice genome has 430 million base pairs and 35,000 genes. Note that a lot of the genome is repeated sequences.

V. The human genome has been sequenced.

A. The two groups finished a draft at the same time in the spring of 2000 and the final draft in the winter of 2003.

B. The human genome has several characteristics.

1. Of the 3.2 billion base pairs, less than 2% are coding, with about 24,000 genes.

2. The average gene size is 27,000 base pairs, including control regions.

3. Over half of the genome has short, repeated sequences that do not code for proteins.
4. About 99% of the genome is the same in all people. There are 2 million single nucleotide polymorphisms (single base-pair changes) that differ in at least 1% of people.
5. The functions of some protein-coding genes are not yet known.
6. Other organisms have sequences very similar to human genes: The worm has hundreds of these, as do fruit flies.
7. The first genomes sequenced from identifiable people were those of James Watson, codiscoverer of the DNA double helix, and Craig Venter, pioneer in shotgun sequencing.

VI. The problem of so few human genes can be explained in three ways.

A. How can humans have 20% more genes than a 1000-celled worm or fewer genes than rice?
1. Genes are interrupted by noncoding sequences called introns. These are removed after the initial mRNA transcription. So a gene is really:
coding 1——intron 1——coding 2——intron 2——coding 3
Many genes have dozens of introns that must be removed as the initial RNA is cut and spliced.
2. This can be done in alternate ways: Remove intron 1 and intron 2: Product is coding 123. Remove intron 1, coding 2, and intron 2: Product is coding 13. These get translated to different products. In this way, the average human gene gets translated to about five proteins.

B. After a protein is made, it is modified, and this gives it new functions. For example, sugars put on a protein that sticks out of the cell surface allow for cell-cell recognition and adhesion. These modifications occur far more often in human cells. So there is great variety there.

C. Micro-RNAs are short transcripts from the genome that remain in the nucleus but are not translated. They appear to be involved in gene regulation, and there are more of them in humans than other genomes.

VII. The next frontiers of biology are the minimal genome and synthetic biology.

A. The prokaryote *Mycoplasma genitalium* has only 482 genes.

B. Scientists have inactivated these genes one by one and asked whether the cell still survives: 100 genes are dispensable. So the minimal genome for a cell is 382 genes.

C. These genes code for proteins involved in basic cell structures (e.g., cell membrane, ribosome) and functions (e.g., enzymes for energy, tRNAs).

D. Craig Venter and colleagues are trying to make each gene in the lab and put them together, creating life. This is called "synthetic biology."

E. There are many potential uses of this technology, for example, custom-made bacteria that can perform any role we assign them: ethanol for fuel, environmental cleanup, making plastics, biological warfare agents, etc.

Essential Reading:

Kevin Davies, *Cracking the Genome: Inside the Race to Unlock Human DNA* (New York: The Free Press, 2001).

David Sadava, Craig Heller, Gordon Orians, William Purves, and David Hillis, *Life: The Science of Biology*, 8th ed. (Sunderland, MA: Sinauer Associates; New York: W. H. Freeman and Co., 2008), chaps. 13, 14, and 18.

Supplemental Reading:

Benjamin Lewin, *Genes VIII* (Upper Saddle River, NJ: Pearson Prentice-Hall, 2005).

James D. Watson, Jan Witkowski, Richard Myers, and Amy Caudy, *Recombinant DNA: Genes and Genomics*, 3rd ed. (New York: W. H. Freeman, 2007).

Questions to Consider:

1. The Human Genome Project included a component studying the ethical and social implications of its findings. If you were in charge of this taxpayer-supported effort, what questions would you ask?

2. If the average human gene can code for five different proteins by alternate splicing out of introns when mRNA is made, and each protein can be modified at least five ways, with 24,000 genes, how many different proteins can be

made? Does this ease your mind at the prospect of having far fewer genes than a rice plant?

Lecture Ten
Manipulating Genes—Recombinant DNA

Scope: The discovery of how bacteria protect themselves from a virus invasion led to a revolution in gene manipulation. Bacteria make an enzyme that recognizes a specific DNA sequence on the invader and cuts it at that sequence. There are many such restriction endonucleases in nature. They can be taken from cells that are brought into the lab, where they are used to cut any DNA in the test tube. The resulting DNA fragment can be spliced to another DNA molecule, creating recombinant DNA. If the DNA recipient is a chromosome, it can be introduced into its host cell, and now that cell has a new gene. Genes from any source can thereby be put into any cell. This circumvents the need for conventional reproductive processes to genetically modify a cell organism.

Outline

I. Opening story: A basic research study leads to a revolution.

A. Werner Arber studied bacterial viruses (phage).

1. Born in Switzerland in 1929, Arber graduated from one of the world's great universities, the Federal Institute of Technology, in Zurich. As a graduate student at the University of Geneva in the 1950s, he studied with a physics professor who had been converted from a pure physicist to a biophysicist and was interested in genetics. Arber, too, caught the "biology bug." The structure of DNA had just been described, and looking at genes was all the rage in science.

2. Arber's thesis was on the phenomenon of bacteriophage restriction. He was trying to find out why a specific genetic type of bacterial virus successfully infects only one strain of host bacteria. Other bacteria were inhospitable.

3. Recall from our earlier lecture how such viruses infect: They actually inject their DNA into the host cell, and the virus DNA takes over the cell, turning it into a virus factory. The cell soon dies, releasing hundreds of viruses.

4. Arber's professors must have been impressed with him, for they hired him as a junior professor at the university in 1960. By 1962, he and his graduate student, Daisy Dussoix, had found that bacteria evade viral infection by chopping up the invading virus DNA into fragments. Successfully infective viruses don't get their DNA chopped.

5. Arber now proposed a hypothesis for what he called virus restriction.

a. The host bacteria make an enzyme that recognizes specific DNA sequences on viral DNA and chops it up at these sequences.

b. The bacteria also have an enzyme that modifies their own DNA to make it resistant to chopping.

c. Virus strains that are successful in infection must have mutations in their DNA that make them resistant to the chopping enzyme.

B. Arber's hypothesis was soon confirmed by the isolation of its components.

1. Arber and a colleague found that successful virus strains had genetic mutations that made them resistant to getting their DNA cut.

2. In the U.S., at Johns Hopkins University, Hamilton Smith and his team isolated and described the chopping enzyme from bacteria. Because it cut DNA only at a certain sequence, it was called a restriction endonuclease (or restriction enzyme).

3. Arber also characterized the system that modifies the host bacteria's DNA to make it resistant to the restriction enzyme.

C. The revolution begins.

1. Scientists soon described several other restriction enzymes for other invading viruses. Each recognized a specific DNA sequence on the virus and cut there.

2. They realized that these enzymes could be removed from cells, and if they put them in a test tube along with a large DNA, scientists could cut that DNA at the specific locations. This essentially led to a way to

"map" large DNAs. At Johns Hopkins, Daniel Nathans and his team used Smith's enzyme to cut up DNA of a tumor virus and map it in this way.

3. Meanwhile, on the West Coast, Stanley Cohen at Stanford and Herbert Boyer at the University of California San Francisco saw Nathans's work and wondered whether they could take it one step further and put the DNA fragments back together after they were cut.

4. A colleague of theirs had described a joining enzyme that can do this. If their proposal worked, Cohen and Boyer realized that any two pieces of DNA could be cut and joined in this way. They tried their experiment with DNA from two different strains of bacteria. One bacterium had a chromosome with a gene for resistance to antibiotic A, and the other to antibiotic B. Cohen and Boyer cut them both with the restriction enzyme, got fragments containing the two genes, and glued them together. Indeed, testing showed that the new chromosome carried genes for both A and B resistance. When they put the DNA into a bacterium that was not resistant to the antibiotics, the bacteria were genetically transformed to be doubly resistant. Since these antibiotic resistant cells were genetically identical, they were clones. It was 1973. They had created genetically functional recombinant DNA (recDNA).

5. This was a revolutionary discovery. It meant that genes from any sources could be swapped and spliced. We were no longer limited by the normal processes, such as fertilization, to mix genes in a single cell. This new method was far broader than fertilization: Soon, human genes were put into bacterial chromosomes, something that could never happen in nature (we don't mate with bacteria).

D. Arber, Smith, and Nathans won Nobel Prizes. The two California universities patented the method to make recombinant DNA and have reaped millions of dollars in royalties. Boyer soon founded Genentech, the first major biotechnology company.

II. There was initially great concern about recombinant DNA.

A. Concerns were raised in several aspects:

1. The bacteria used were *Escherichia coli*, which commonly occur in the human intestine. What if the recDNA got out and into human DNA?

2. Since any gene could be out into the cells, what if a disease-causing (e.g., cancer) gene was under study and it got out?

3. What about just putting any DNA that was unidentified into bacteria (called shotgun cloning)?

B. In February 1975, a conference was held at Asilomar, California, that brought together scientists, ethicists, physicians, and lawyers. The meeting urged a moratorium on certain experiments and safety precautions on others. Oversight by government agencies and institutional boards that continues to this day was set up. It was an important event in government regulation of research. In retrospect the concerns were overblown, and untoward events have not occurred. Experiments that required severe precautions in 1975 are done in high school science labs today.

III. The methods of recombinant DNA are widely applicable.

A. There are two goals for making recombinant DNA.

1. The first goal is to study a particular DNA sequence, for example, to find out about promoter sequences within a gene. In this case, the recDNA is inserted into a cell, and when that cell divides, the recDNA is amplified along with it.

2. The second goal is to take the recDNA and insert it into a cell and have the cell express the gene(s) carried by the recDNA. This can be done for two reasons.

a. As an experiment to see what happens. Adding a new gene to a cell can be used to investigate cause-and-effect relationships. For example, suppose we think that a certain hormone (like adrenalin) must bind to a recognition protein on the surface of a muscle cell before acting on the cell. We can get the gene for that protein, insert it into any cell that normally does not respond to the hormone (e. g., a skin cell), set it up so that the gene is expressed in the new cell, and see if the skin cell now binds the hormone.

b. To have the recipient cell make a product. For example, a hormone such as insulin is needed by

many diabetics. The only way to get the hormone used to be by extracting it from the pancreases of animals. Now, bacteria can be used to make the protein.

B. Many organisms are used as a host for recDNA. In theory any cell can be used, but in practice we use cells that we can manipulate in the lab and that are well characterized.

1. Bacteria were the first cells used. But they are prokaryotic, and may not do all of the added things (e.g., add sugars) that make a protein active in eukaryotic cells.

2. Yeast is a simple, single-celled eukaryote. Its genome is sequenced.

3. Mammalian cells, even human cells, can grow and reproduce in the laboratory.

4. RecDNA can be added to the egg of an animal, and then the animal reproduces. This makes a transgenic organism.

5. Plant cells are easily grown in the lab. If plant cells are grown in an appropriate environment, they make a new, transgenic plant.

C. There are many restriction enzymes that cut DNA at specific sequences. For example, one cuts at

xxxxGAATTCxxxx

xxxxCTTAAGxxxx.

Whenever this sequence is in DNA, that DNA will be cut by that enzyme. This specificity is useful, as the scientist may know the sequence of a target DNA and can choose where to cut it.

D. There are two typical ways to get DNA into host cells.

1. DNA can enter cells naked if salts are used to neutralize the charges in the DNA and the cell.

a. The problem with this is that the recDNA often stays outside the nucleus and is not part of a chromosome.

b. So when the cell gets ready to divide and replicates its DNA, the recDNA will not be replicated and will not get to the new cell. Soon, it gets diluted out, and few cells have it. This is like buying a stock so that you have 10% of the company, and then the company issues more and more shares to other investors, so your actual ownership ends up at 1%.

2. A chromosome from the host organism can be isolated and put into the test tube. Then the chromosome is cut with the restriction enzyme and the new DNA spliced into it. This recDNA is then put into the host cell, and since it is part of normal chromosome, the host cell thinks that it "belongs," and the recDNA stays there and is replicated.

a. The carrier chromosome is called a vector.

b. Vectors are usually very small chromosomes so that they are easily manipulated in the lab and also contain, preferably, only one restriction enzyme recognition sequence, so they are cut only once for insertion.

c. Because vectors are chromosomes, they have genes, and these genes can be useful. Once they are expressed in the host cell, the vector genes may be "markers" for the presence of the vector and its new gene hitchhiker.

d. Vector marker genes include antibiotic resistance, so any cell carrying the vector has that property and can be selected in a mixture of cells.

e. In addition to small chromosomes, vectors include viruses that normally infect the host cells. These viruses are disabled by lab manipulations so that they inject their DNA into the cells but do not reproduce and kill the host.

IV. To summarize:

A.

B.

C.

Essential Reading:

Susan Barnum, *Biotechnology: An Introduction* (Belmont, CA: Thomson Brooks-Cole, 2005), chap. 3.

William Thieman and Michael Palladino, *Introduction to Biotechnology* (San Francisco: Benjamin-Cummings, 2004), chap. 3.

Supplemental Reading:

George Acquah, *Understanding Biotechnology* (Upper Saddle River, NJ: Pearson Prentice-Hall, 2004).

John Smith, *Biotechnology*, 4th ed. (Cambridge: Cambridge University Press, 2004).

Questions to Consider:

1. The initial reaction to the creation of recombinant DNA in the 1970s ranged from positive excitement to panic. This led to the Asilomar conference. In retrospect, the concerns about the dangers of this new technology were largely groundless. Can you think of any recent advances in science and technology that provoke similar concerns? Is it a good idea for society to be cautious, or are the concerns usually exaggerated?

2. The creation of genetically new organisms by recDNA manipulation led to a new industry. To protect their inventions, scientists can patent their discoveries. This led to the patenting of living organisms. Is this a good idea?

Lecture Eleven
Isolating Genes and DNA

Scope: Finding the protein that is defective in muscular dystrophy presented a challenge because the protein is present in muscle in tiny amounts. So scientists first isolated the mutant gene responsible for the disease and then used the gene's DNA sequence to identify the protein. An important technique in finding genes is nucleic acid hybridization, in which single strands of DNA (or RNA) from different sources are incubated together to see if they will bind by base pairing (A to T and G to C). If that occurs, the sequences must be closely related. An entire genome can be chopped into fragments, and these displayed by cloning into bacteria (a gene library) or directly on a microarray (a DNA chip). In both cases, the fragment containing a certain DNA sequence can be "fished out" by using that sequence as a hybridization probe into the library or array. This has been useful not only in isolating genes but also in finding out patterns of gene expression. The latter may be important in new ways for medical diagnosis and prognosis. The ability to actually make DNA in the laboratory has freed scientists from looking for useful mutations in nature and opens up the possibility of making genes whose protein have customized properties.

Outline

I. Opening story: Muscular dystrophy was a genetic challenge.

A. In traditional genetics, scientists went from phenotype to gene.

1. As we have seen, genes were first identified as phenotypes, the genes were inferred, and then the genes were isolated.

2. For example, PKU was described in terms of genetics (phenotype), and then the missing phenylalanine hydroxylase enzyme was isolated, and this led to the isolation of the actual DNA sequence coding for it.

B. Muscular dystrophy was a challenge to this paradigm.

1. The most common form of muscular dystrophy was first described by French neurologist Guillaume Amand Duchenne in 1861. He noted a progressive muscle wasting in boys, first with the pelvic and calf muscles before age 5, and then losing the ability to walk by age 12. The weakness spreads to all the so-called voluntary muscles and the breathing muscles. Death is usually before age 30, due to respiratory failure.

2. The disease is genetically determined, linked on the X chromosome (like hemophilia). So it is most common in males, with 1 in 3500 births.

3. Treatment is symptomatic (braces, wheelchairs, and ventilators). There is no good treatment and no cure. So it is urgent to find out what exactly is wrong with the muscles (the precise phenotype).

4. For decades, scientists looked for a protein difference between the normal and dystrophic muscle. They knew that the muscle fibers tended to fall apart and that they got replaced by fatty or other nonmuscle tissues. But with the methods they had, they could not find one. It was like looking for a needle in a

haystack.

II. In reverse genetics, the gene is isolated before the protein.

A. Methods for analyzing DNA in chromosomes were developed during the 1980s.

1. By the 1980s, scientists were ready to take another approach. Why not try to isolate the gene first and then go after the defective protein?

2. At Harvard Medical School, Louis Kunkel and his colleagues saw a boy with Duchenne muscular dystrophy in the clinic. When they examined his chromosomes under the microscope, they saw that he was missing a small piece of the X chromosome. Figuring that the missing piece was the gene that is defective (in this case, missing) in muscular dystrophy, they compared the DNAs of X chromosomes of the affected boy with normal boys. They found that a piece of DNA was missing, and in 1985 they isolated the gene.

3. The gene was huge; in fact it turns out to be the largest in the human genome, at over 2.5 million base pairs. It makes an mRNA in developing muscle cells that is 14,600 bases long. Why is all that extra DNA in the gene? It is introns (the gene has 78 of them and so is in 79 pieces) as well as a complex of 9 promoters.

4. The protein coded for by the gene was indeed hard to find: It accounts for 0.002% of muscle protein. Appropriately called dystrophin, it connects the inside of the muscle cells to the outside, running through the cell membrane. Without dystrophin, muscles tend to get injured by the shear forces of muscle contraction. Ultimately, this kills the muscle cell. This identified the primary phenotype—after the gene was isolated.

5. The discovery of the mutant allele and then protein responsible for this disease has led to molecular medicine, a targeted approach to treatment.

a. A protein similar to dystrophin called urotrophin is already made by the body, and drugs are being developed to increase its synthesis so it can replace defective dystrophin.

b. Gene therapy, adding the normal allele for dystrophin to muscle, is also being developed.

B. Nucleic acid hybridization is a key method to detect genes.

1. How can scientists compare two DNAs such as the normal X chromosome and the one that was deleted in the boy with muscular dystrophy?

2. A method to compare nucleic acids was developed by Sol Spiegelman in the 1960s. It is called nucleic acid hybridization.

3. Suppose a DNA has the sequence AAAGGGCCCTTT

TTTCCCGGGAAA.

We have another DNA and want to know if it matches this sequence. One way to find out is to separate the DNA strands of the target DNA (a) AAAGGGCCCTTT and
(b) TTTCCCGGGAAA and put these on filters. Now separate the strands of the unknown DNA and let them settle onto the filters. The unknown DNA will stick to the filters only if it has the complementary base pairs (A to T and G to C), that is, if it sticks to filter (a), then the unknown DNA must have the sequence TTTCCCGGGAAA. The double-stranded DNA that results comes from two different sources and is called hybrid DNA.

4. Now let's apply this to the situation with the X chromosomes from the muscular dystrophy patient and the normal boy. We try to hybridize the DNAs with the normal boy's DNA as the target. What will happen is that most of the two DNAs are very similar and will hybridize, but there will be a part of the X chromosome of normal DNA that does not have a partner on the patient chromosome. That is where the gene for dystrophin is located.

III. DNA libraries and chips are used for rapid screening of a genome.

A. Genome libraries can be made by cloning.

1. Often only part of a gene has been isolated (for example, the deletion of the X chromosome in a patient with muscular dystrophy may be smaller than the huge gene).

2. If we could make the isolated gene fragment into a hybridization probe, and if we could display the entire genome in fragments, the probe would bind to the fragment with the gene, and the gene could be "fished out" with this "hook."

3. A genome library does this: The entire DNA of an organism is cut into small (gene-sized) pieces that are cloned into bacteria using the recombinant DNA method we described in the last lecture. So we have many colonies of bacteria, each having a different part of the genome.

4. Now a mass DNA extraction of all the colonies is done, and the gene fragment is hybridized to all the colonies. The one it sticks to must have the whole gene, and so this bacteria colony can be grown and the whole gene isolated.

B. A DNA microarray is the ultimate library.

1. A major technological innovation in electronics in the 20th century was the silicon chip: a thin micro-wafer with circuits stamped on it.

2. A DNA microarray is a small, glass microscope slide to which single-stranded DNAs are attached as targets for hybridization. The entire process is miniaturized so that the DNA chip can have thousands of genes on it.

3. If the unknown DNA that is to be hybridized is colored, the spot location where it hybridizes will be colored, so detection is easy.

4. DNA microarrays are being used to find genes and mutations, and in diagnosis.

5. RNAs can also be hybridized, so that if all the human genes are present as DNAs on a chip, one can extract RNAs from a cell and ask which genes are expressed in a tissue at a certain time.

6. For example, in breast cancer, it is important that the physician get an idea of what stage the tumor is in. If it is at an early stage, maybe the tumor is localized and surgery may "get it all," as surgeons sometimes (unwisely) say. Some early tumors may have already spread (metastasized), but the doctors don't know it. It would be good if the tumor had some sort of marker that would show whether it was the spreading type. If the physician had this information, only those patients who need chemotherapy and radiation after surgery would get it.

a. But there has been no good way to tell the "spreading" from the "non-spreading" type of tumor, so the physicians play it safe and give most women the extra treatments.

b. Enter gene chips. Dr. Laura van 't Veer at the Netherlands Cancer Agency took biopsies from both types of breast cancer, extracted their RNAs, and hybridized them to microarrays with thousands of human genes. She and her colleagues found that some genes were expressed more in the tumors that were found to spread, and others in tumors that did not. So she created a "gene expression signature." This may be very useful in targeting further aggressive treatments to only those who need them.

IV. Synthetic DNA can be used to make new genes and mutations.

A. Chemists have developed methods to make DNA in the lab, base by base.

1. Knowing the genetic code makes it possible to create mutations.

2. This is a major advance in the use of genetics to ask "what if" questions. Previously, we needed to either look in nature for a mutation that arose spontaneously or treat the organism with a mutagen like X rays and look for the mutant phenotype (e.g., recall the prokaryotes that could not make an amino acid because they had a defect in a gene coding for an essential enzyme). With custom DNA in the lab, we can hypothesize that a certain gene is essential for the synthesis of an amino and then mutagenize it deliberately and see if an organism carrying the mutant now cannot make the amino acid.

B. These lab techniques can be used to create useful mutations.

1. For example, scientists found a naturally sweet protein that was too sweet. It stuck to the taste buds for hours.

2. The scientists determined that the three-dimensional fit for the sweet molecule was too tight, so they proposed changing the protein so that its amino acid sequence at the binding region was changed for less tight binding. Looking up the genetic code, they altered the DNA for the protein, cloned it into bacteria

and—voila—sweet, but not lasting.

C. These lab techniques are being used in synthetic biology to create "life" and custom-made organisms.

Essential Reading:

Benjamin Pierce, *Genetics: A Conceptual Approach* (New York: W. H. Freeman, 2005), chap. 19.

William Thieman and Michael Palladino, *Introduction to Biotechnology* (San Francisco: Benjamin-Cummings, 2004), chap. 3.

Supplemental Reading:

Ricki Lewis, *Human Genetic*s, 7th ed. (New York: McGraw-Hill, 2006).

John Smith, *Biotechnology*, 4th ed. (Cambridge: Cambridge University Press, 2004).

Questions to Consider:

1. This ability to make any DNA we want in the lab has implications on discussions of genetic resources. One argument for conserving nature has been the human-centered one: There are genes out there that we may be able to use for our purposes (making new drugs to cure cancer, as in the film *Medicine Man*). But if we can make any gene or mutation that we want, this reason becomes less compelling. Do you agree?

2. Compare the "old way" of doing genetic analysis by phenotype first and then isolating the gene to the "new way" of isolating the gene first and then the protein phenotype.

Lecture Twelve
Biotechnology—Genetic Engineering

Scope: Biotechnology is the use of microbes (such as bacteria), plants, and animals to make products useful to humans. Agriculture was probably the first biotechnology, when people domesticated plants and used them as crops for food and fiber. Plant materials were used to make beer and bread, followed by yogurt, cheese, wine, etc. Microbes are grown for their products such as antibiotics and amino acids, an activity which became the fermentation industry. The discovery of genetic engineering by recombinant DNA, as well as knowledge from basic research of the conditions that maximize gene expression, led to the creation of a new biotechnology industry based on recombinant DNA. Proteins that are hard to get in usable amounts in nature, such as the clot-busting drug tPA and human insulin, are now made by genetically modified cells. Even whole animals and plants can be turned into factories for useful proteins.

Outline

I. Opening story: Biotechnology to the rescue.

A. A stroke victim was saved by a genetically engineered drug.

1. As he drove home from work, Bill first felt his face twitch, and then his speech became slurred. He was the latest of several million people a year in the U.S. who have a stroke, in which a blood clot blocks an artery leading from the heart to the brain. This deprives brain cells of oxygen, normally carried by the red blood cells, and irreversible damage can occur quickly.

2. Normally, blood clots "go away" after some time. There is a clot-dissolving system in the blood that gets activated by cells near the clot. Of course, the time factor is important, because we want a clot to be there for a while to stop the flow of blood at an injury. Unfortunately for Bill, a stroke (as well as a heart attack) is not a situation where you want a clot around for long. Every minute that blood flow is blocked to the brain is harmful.

3. Lucky for Bill, he was near a hospital and drove his car into the ER parking lot, leaning on the horn. The staff ran out, got him into the ER, and immediately injected a drug right onto the clot. It quickly dissolved, and blood flow was restored with minimal permanent brain damage. Bill was home the next day.

4. The drug the ER staff injected was tissue plasminogen activator (tPA, also called PLAT). It is the protein that actually initiates the clot-dissolving process. In the normal course of events, this takes time (for good reason, as just stated). Adding it to the blood right at Bill's clot just sped up the process.

5. Thankfully for the typical blood clotting processes, tPA is made in small amounts and only when needed. So it is virtually impossible to get enough of it to store on the shelf of the ER, ready for injection when needed. Enter recombinant DNA.

6. First, DNA was extracted from human cells and the gene for tPA was isolated by the methods described in the last lecture. Then this DNA, with an appropriate promoter that stimulates active gene expression, was inserted into hamster cells growing in the laboratory by genetic engineering methods described two lectures ago. These cells containing the recombinant DNA churned out tPA in amounts far in excess of what could ever be extracted from blood. The tPA protein was purified and put into a vial, ready for a patient like Bill.

7. This scenario—of gene to useful protein by genetic engineering—has now been played out for dozens of products. It is part of a revolution in biotechnology.

B. Biotechnology is not new.

1. Biotechnology—the manipulation of microbes, plants, and animals to make produces useful to people—began long ago.

a. Agriculture—harvesting, planting, and cultivation of plants and animals for food and fiber—was probably the first biotech. It may have begun about 10,000 years ago in what is now Iraq, when Sumerians learned that barley plants that grew near their settlements made seeds that could be mashed up and used to make beer, and later, bread. So they grew the seeds in plots near their settlement—the first farms. In ancient Egypt, the hieroglyphic symbol for food was a pitcher of beer and a loaf of bread.

b. The process of making alcohol from mashed-up seeds (or grapes) was actually carried out by yeast cells and called fermentation (from Latin *fervere*, to boil; bubbles form because of the release of carbon dioxide).

c. Fermentation under different conditions and by different organisms produced other products as well: wine, spirits, cheese, yogurt, vinegar, etc.

2. By the mid-20th century, the industry called biotechnology was using microorganisms in huge vats, to make antibiotics, oils, amino acids, and enzymes for the food industry.

II. Modern biotechnology uses recombinant DNA.

A. The invention of recDNA methods during the 1970s opened up two new horizons for the manipulation of organisms in biotechnology.

1. First, the existing organisms could be genetically manipulated to be more efficient to make their products by deliberate mutations, addition of active promoters for gene expression, etc.

2. Second, new genes could be inserted into productive organisms to turn them into factories for products that would ordinarily be inaccessible. This is the story of tPA.

B. The key to production of a protein product is the promoter.

1. Recall that the promoter is a DNA sequence adjacent to a gene that attracts RNA polymerase, the enzyme that makes mRNA and is responsible for gene expression.

2. Biologists learned that promoters are specific for place (cell type) and time.

3. So a promoter DNA would be added to the gene DNA appropriate to the cell type and environmental conditions. The vector DNA to which this was added is called an expression vector.

III. Human insulin was the first major product of DNA biotechnology.

A. The problem: Insulin (from the Latin word for island—the type of tissue in the pancreas gland that makes it) is a protein that acts as a hormone to stimulate uptake of blood sugar into tissues. People with type 1 diabetes (from the Greek word meaning "passing through," referring to excessive urine production) don't make insulin and so need to inject it.

B. Previously, insulin came from slaughtered animals. But animal insulin is a bit different in amino acid sequence than the human one, and some people reacted against it, thus the need for human insulin. There was

no supply of human pancreases.

C. So a team of scientists developed a way to make a lot of it by recombinant DNA.

1. At City of Hope Medical Center in California, Keiichi Itakura synthesized the insulin gene (which has 51 amino acids).

2. His colleague, Art Riggs, put the gene into an expression vector next to a high-expressing promoter.

3. The recDNA vector was put into bacteria.

4. The bacteria were grown in a large vat by the (then new) biotech company Genentech (founded by Herbert Boyer).

5. The insulin was extracted from the vat, sent to a drug company, and then sent to physicians. This is the source of all insulin now used to treat diabetics.

IV. Other medically useful substances are produced by biotechnology.

A. Some proteins replace ones that are missing in genetic diseases. For example, in hemophilia (Talmud story in Lecture Three, etc.), a blood-clotting protein is missing. When the patient gets injured, it can be replaced by the clotting protein made by biotechnology.

B. Some proteins are used as drugs to treat diseases. tPA is one example. Another example is erythropoietin (EPO). This protein is made by the kidneys, enters the bloodstream, and goes to bone marrow, where it stimulates the production of red blood cells.

1. People with kidney failure are treated with dialysis. This is very helpful, but it removes EPO, so the patients get very anemic (low red blood cells). They need transfusions, unless they get EPO. People with functional kidneys make very little EPO, so getting EPO for the patients from other people's blood would be very difficult.

2. So the EPO gene was isolated and EPO made by recDNA. It is now used not just for dialysis but also for people undergoing cancer chemotherapy (which often destroys bone marrow cells).

3. EPO became the first biotech drug of abuse: Athletes found it could increase their red blood cells by about 10%, and this gave them an edge in competition (e.g., the Tour de France).

V. Plants and animals can be genetically engineered to make products.

A. Pharming results in transgenic animals.

1. Dairy animals such as sheep, goats, and cows produce a lot of milk. Biologists have found that the expression of genes for major protein in milk is under the control of a lactoglobulin promoter that is expressed in the mammary gland.

2. Idea: Take any gene you want expressed and put it in a DNA vector that has the lactoglobulin promoter. Then put this recDNA vector into a goat egg-cell nucleus. Allow the egg to develop to an embryo and insert the embryo into a surrogate mother. Result: a transgenic sheep that makes a desired protein in its milk.

3. Example: Some people lack adequate human growth hormone (hGH). This is a protein made in the pituitary gland in the brain. Supply is extremely limited as a result. In Argentina, scientists put the gene for hGH into cows with high expressing lactoglobulin promoter. Ten cows can make enough hGH to supply the annual world demand.

B. Plants can be genetically engineered to make products.

1. As we will see, genetic engineering of plants is easier than animals because any cell can be used (you do not need an egg). In addition, plants grow a lot and produce a lot of protein in their leaves and fruits.

a. Plants can be engineered to make proteins that are extracted. For example, tobacco plants (the farmers are always looking for new uses) have been modified by recDNA to make tPA in their leaves. This is a huge source.

b. Plants can be engineered to make proteins in fruits that are eaten. For example, a "plantibody" project involves a gene for a vaccine against bacterial meningitis (see Lecture Nineteen on the immune system). This gene has been put into banana plants and is expressed in the fruits. This would

be useful as it circumvents the need for vaccine administration by health professionals.

2. Plants can be engineered to make enzymes that create new biochemical pathways and new products.

a. For example, a major component of detergents is lauric acid (look at a label). This molecule is made in a biochemical pathway in tropical plants such as the coconut and the palm tree. Palm kernel oil is a major source, and this must be imported from the tropical countries to those in more temperate climates. Recently, scientists pinpointed a key enzyme responsible for the biosynthesis of lauric acid in a tropical plant, cloned its DNA into an expression vector, and put the vector into a rapeseed plant (canola) that is grown widely in temperate climates. The transgenic plant makes oil that has 60% lauric acid (up from 0%).

b. Tobacco mosaic virus infects, and reproduces in, but does not kill, tobacco plants. In a novel approach, the viral genome can be replaced in part with a new gene, in this case one for a vaccine, and instead of a lot of viral coat protein, the virus makes the vaccine protein in large quantities in tobacco leaves. This may be a new use for this widely grown crop.

Essential Reading:

Susan Barnum, *Biotechnology: An Introduction* (Belmont, CA: Thomson Brooks-Cole, 2005), chaps. 1, 6, and 7.

Maarten Chrispeels and David Sadava, *Plants, Genes and Crop Biotechnology* (Sudbury, MA: Jones and Bartlett, 2003), chap. 19.

Supplemental Reading:

Cynthia Robbins-Roth, *From Alchemy to IPO: The Business of Biotechnology* (Cambridge, MA: Perseus Group, 2001).

John Smith, *Biotechnology*, 4th ed. (Cambridge: Cambridge University Press, 2004).

Questions to Consider:

1. Why did drugs such as insulin, tPA, and hGH have to be made by biotechnology? What were the other ways to get these drugs, and why were they inadequate? Can you think of other useful proteins that need to be made this way?

2. Some people are opposed to genetically modified organisms in the food supply. How would these people react to finding out that drugs for human consumption such as insulin are made by genetically modified organisms (in this case, a bacterium making a human protein)?

Lecture Thirteen
Biotechnology and the Environment

Scope: Some organisms have genes whose products can be used to sense or break down environmental pollutants. Other organisms can be genetically engineered to do so. A plant has even been made into a biosensor for land mines using a promoter from bacteria that is sensitive to the explosive and a gene from a jellyfish that makes a protein that glows. Biological control uses organisms to consume other organisms that we regard as pests. Many bacteria species have genes for using what humans think of as wastes for their energy and growth, and this forms the basis of such familiar processes as composting and wastewater treatment. Other bacteria have genes for enzymes that degrade pollutants such as oil and are used in cleaning up oil spills in a process called bioremediation. Still other bacteria can help extract minerals in mining. All of these species can be improved by genetic engineering. Creation of pollutant-digesting bacteria was the first example of patenting of a genetically modified organism.

Outline

I. Opening story: Biotechnology can be used to detect landmines.

A. Landmines are a sad legacy of recent history.

1. Mines have wreaked havoc both during and after recent wars. They are cheap, plastic casings filled with an explosive such as TNT (trinitrotoluene) that are triggered to explode when disturbed. The UN estimates that over 100 million unexploded mines lie on or below ground in many countries, from Angola to Cambodia. This makes plowing the fields for farming or other uses a risky business.

2. Encased in plastic, the mines avoid metal detectors. The best way to find them is the most dangerous: A

person gingerly walks around, poking the ground with a long stick and jumping out of the way of any explosion.

B. Neal Stewart at the University of Tennessee is trying to use plants to remedy this ecologically adverse situation—this is an example of bioremediation. He is making plants that will tell us where the mines are.

1. The idea is to make the plant a biosensor, making a detectable protein wherever the landmines are. This requires two genetic components.

a. A gene whose protein product can be visualized. Certain deep-sea jellyfish glow in the dark of the depths because they make a protein that fluoresces in weak black (ultraviolet) light. In the 1990s, the DNA coding for this green fluorescent protein (GFP) was isolated and biotechnologists put it into vectors and used it as a marker for the presence of recombinant DNA in many organisms. Glowing organisms were the result (glowing fish were sold for a while).

b. A promoter that turns on a gene in the presence of TNT. Microbiologists found a bacterium, *Pseudomonas putida*, that could digest TNT and use its breakdown products. The genes coding for the enzymes responsible for this metabolism have promoters that become active for gene expression in the presence of a very low amount of TNT.

2. Stewart has put the two together: a TNT-sensitive promoter and the DNA coding for GFP. He added them to a vector and made transgenic plants. When they are around TNT, they glow under ultraviolet light. A biosensor is born.

3. A major challenge is detecting the glowing plants. Stewart and his colleagues propose remote sensing and seeding from airplanes.

II. Organisms have long been used to solve environmental problems.

A. Biological control uses the pests of pests.

1. An old biology poem says:

> Big things have little things Upon their backs to bite ‘em And lesser things, still lesser things And so ad infinitum.

2. This is the food chain in ecology. Knowledge of this chain has allowed biologists to use it to introduce pests that eat pests.

a. For example, in my first job as a student biologist, I described the tiny insects that preyed on the bud moth caterpillar, a worm that eats up the leaves of apple trees. We introduced this insect to reduce populations of the harmful worm, thereby reducing the need for pesticides. This is called biological control.

b. This method is now widespread and works with traditional ways of pest control in a process called integrated pest management.

B. We can use bacteria as nature’s recyclers. Bacteria have the genetic capacity to thrive on all sorts of nutrients, including what we refer to as wastes.

1. Composting uses bacteria that break down carbon-rich stores such as cellulose in wood chips, paper, straw, and hay, as well as nitrogen-rich sources, such as protein wastes, coffee grounds, and fruit and vegetable scraps.

2. Wastewater treatment uses a variety of bacteria to act on human wastes, paper products, and household chemicals.

III. Bacteria and plants have genes for environmental cleanup.

A. In addition to being nature’s recyclers, bacteria can break down many human-made pollutants. They have been discovered simply by mixing soil or water, or some other source of bacteria, with the pollutant and seeing what survives, and thrives.

1. The first bacteria developed for cleaning up oil spills were developed by Ananda Chakrabarty in the 1970s. He contaminated soils separately with various components of crude oil and isolated bacteria that survived. Then he mated the bacteria sequentially to get a single stain that could break down multiple

substances (this was before genetic engineering, which is the way it is done today). Importantly, he applied for and was awarded a patent for this "super-bug." This landmark decision led to a flood of patents for genetically modified organisms.

a. In 1989, the oil tanker *Exxon Valdez* ran aground near the Alaskan shore, releasing 11 million gallons of crude oil over 1000 miles of shoreline. Physical methods, such as skimming the water and spraying the rocky shore, were used first. This dispersed about two-thirds of the oil. Bacteria did the rest: Nitrogen and phosphorus salts were sprayed on the shoreline and rocks to stimulate the growth of *Pseudomonas* bacteria already there; other strains were added. Soon, the bacteria became active and worked quickly to degrade the oil. The process is ongoing.

b. The government of Kuwait is using bioremediation to try to clean up the 250 million gallons of oil that was spilled onto the desert in the Gulf War of 1991. This may be the single largest bioremediation project.

2. Extremophiles have many useful genes.

a. These are microbes that live in very hot or cold or pressured or salty environments. They are part of a separate group of organisms called archaea because they resemble the organisms thought to be the first on Earth. In archaea genomes, half of the genes do not resemble genes in bacteria or eukaryotes.

b. *Deinococcus radiodurans* is a microbe that lives in the most dangerous environment of all, high levels of radiation. Normally, radiation kills cells by damaging their DNA and overwhelming the cell's ability to repair the damage. It gets around this by having the most efficient radiation DNA repair system in nature. The genes from other extremophiles are being engineered into *D. radiodurans*. No matter what the environment, it keeps coming back. No wonder it is called "Conan the Bacterium"! It is being used to clean up the most toxic sites.

B. Plants can be genetically engineered for environmental cleanup.

1. The bacterial genes that allow environmental cleanup can be put into plants. Thus there are transgenic plants that can break down oil, convert solid mercury and arsenic into harmless substances, etc.

2. Why use plants when the microbes are available? The issue is getting the bacteria out of the soil when they are done. This is very energy-intensive. Plants can just be harvested by conventional agriculture.

IV. Biomining uses bacteria to help extract metals.

A. Mining for minerals is an old human activity that has changed little for millennia.

1. The earth is dug up, ores are taken out, and minerals such as copper and gold are extracted from the ores by harsh methods such as chemicals or heat.

2. It is environmentally damaging.

B. The bacterium *Thiobacillus ferrooxidans* uses copper sulfide (instead of carbohydrates, for example) to get energy.

1. This process releases the copper. Ores are sprayed with sulfuric acid and, if necessary, the bacteria. They grow and leach out the copper. This accounts for 25% ($1 billion) of all copper mined in the world.

2. The bacteria release a lot of heat as they break down the ore, and this slows down the process. Scientists are now using recombinant DNA technology to transfer the genes for copper release into an extremophile bacterium from a hot spring to make the process more efficient.

Essential Reading:

George Acquah, *Understanding Biotechnology* (Upper Saddle River, NJ: Pearson Prentice-Hall, 2004), chap. 18.

William Thieman and Michael Palladino, *Introduction to Biotechnology* (San Francisco: Benjamin-Cummings, 2004), chap. 9.

Supplemental Reading:

John Smith, *Biotechnology*, 4th ed. (Cambridge: Cambridge University Press, 2004).

Sharon Walker. *Biotechnology Demystified* (New York: McGraw-Hill, 2007).

Questions to Consider:

1. Have you visited a water treatment plant? Your water company's Web site or plant brochure can give you an idea of the role of bacterial genes in this process.

2. Ecologists have been surprised by the rapidity of recovery of ecosystems after the disastrous oil spills in Santa Barbara (in 1969) and Alaska (in 1989). How important are bacteria in the recovery processes? Would genetic engineering help?

Lecture Fourteen
Manipulating DNA by PCR and Other Methods

Scope: The book and film *Jurassic Park* brought DNA to public attention. Although the story is fictional, it was based on the science of DNA. Three major ways to manipulate DNA are described in this lecture. The first, the polymerase chain reaction, allows any DNA sequence—even in tiny amounts in a single cell—to be amplified in the test tube, obviating the need for recombinant DNA cloning. This method uses an enzyme from a heat-loving bacterium that lives in hot springs. The second method, DNA sequencing, uses chemical modifications to determine the sequence of any short DNA. This can be automated, and computers analyze the sequence for its biochemical and genetic meaning. The third important technique, RNAi (i = inhibition), is a way to specifically block the expression of a single gene. It came from accidental observations in petunia plants and is now the subject of intensive research for drug development to treat diseases involving gene expression.

Outline

I. Opening story: DNA manipulation in *Jurassic Park*.

A. Science fiction can be based in science facts.

1. In 1990, author Michael Crichton, trained as a physician and scientist, published *Jurassic Park*, a best-selling novel that brought the letters "DNA" into public consciousness as never before.

2. Dr. Raul Cano, a microbiologist, had announced that he had extracted some intact DNA from a bee preserved in amber, the preserved resin of a tree that lived about 40 million years ago. Crichton then extended this idea into fiction.

3. In the novel, a mosquito was trapped in amber after sucking up some dinosaur blood during the Jurassic period of geological time. Dinosaur DNA fragments were extracted from the mosquito and spliced together with DNA from current reptiles, birds, and frogs to fill out complete genomes of dinosaurs. These were then inserted into special egg-like structures, and the dinosaurs of the book—and the 1993 film—were cloned.

B. Crichton's book sparked interest in the science involved, even to the extent of sparking debates about dinosaur cloning.

1. Although the presence of undegraded DNA is deemed unlikely by most scientists, the techniques of DNA manipulation continue to be developed.

2. These powerful lab methods have come from increasing knowledge of DNA in nature.

II. The polymerase chain reaction can amplify any known DNA sequence.

A. In 1956, Arthur Kornberg described the first enzyme that can catalyze DNA replication, a DNA polymerase.

1. He showed this in an experiment. First, the two strands of DNA had to be separated to expose their bases for base pairing with the new strands. Recall that this process is called semiconservative replication. When Kornberg added the enzyme extracted from cells, along with nucleotide building blocks (A, T, G, and C), to DNA in the test tube, the DNA was duplicated. For this tour de force of biochemistry he was awarded the Nobel Prize three years later.

2. The discovery of DNA polymerases (there are several) opened up the possibility of continuously amplifying DNA in the test tube. This polymerase chain reaction (PCR) could be useful in many types of genetic studies. In Kornberg's experiments, DNA was replicated once. Why not let it go on for more

rounds … 2, 4, 8, 16, 32 ... ?

3. The problem was that separating the two strands of DNA takes energy, since the opposite bases (A with T and G with C) fit together exactly. Weak chemical interactions called hydrogen bonds must be broken. This separation is usually done with heat, up to over 80°C (175°F). So in the DNA doubling experiment, after the DNA doubled, the two new DNA molecules had to be separated into their four component strands by heat before being cooled down for polymerase action.

4. Heat destroys the three-dimensional structures of most proteins irreversibly (like boiling an egg), and since Kornberg's bacterial DNA polymerase is a protein, it would be irreversibly destroyed. So, new DNA polymerase would have to be added to the mixture for each round of replication. Not only is this a hassle, it is prohibitively expensive.

5. Meanwhile, microbiologist Thomas Brock had been studying the first extremophilic bacteria that live in the hot springs in Yellowstone National Park. Appropriately named *Thermus aquaticus* ("hot water"), these prokaryotes thrive in water above 70°C that would kill most other organisms. Brock and his colleagues wrote research papers that described how these bacteria survive by having heat-resistant biochemical machinery. Brock and his student, Hudson Freeze, described the first DNA polymerase enzyme that survives heat.

6. At a biotechnology company in San Francisco, a group of scientists led by Kary Mullis came up with the idea in 1983 of using *Thermus aquaticus* DNA polymerase in PCR. Heating and cooling cycles could process without the need to add new DNA polymerase. They published it in 1985, and it was an immediate hit. The Cetus Corporation sold the patent rights for $300 million to a larger corporation. They gave Mullis a $10,000 bonus. In 1993, he won the Nobel Prize.

B. PCR is a major technique in basic and applied biology.

1. The most important advantages of PCR to amplify DNA over cloning by recombinant DNA is that PCR is fast. Typically, it takes just a few hours to amplify a DNA sequence a millionfold.

2. Another advantage of PCR is that is it extremely sensitive. The DNA of just a single cell can be amplified for analysis or use. As we will see in the next few lectures, this makes PCR valuable in forensics and diagnosis.

III. DNA can be sequenced and analyzed.

A. In 1968, Robert Holley was awarded the Nobel Prize for leading a team that determined the sequence of the first nucleic acid. It took his team five years (1959–1964) to get the 80-nucleotide sequence of a transfer RNA. Today, this is done by machine and takes a minute.

1. The widely used method for DNA sequencing, 800 base pairs at a time, was developed by Frederick Sanger in 1977.

2. DNA sequencing is similar to PCR—with a twist. The two strands of the DNA to be sequenced are separated, and DNA polymerase is added, along with the four bases, A, T, G, and C. DNA replication begins. Say our DNA has the sequence:

TTGTGCATTAAACT …

Replication will add: AA … and continue.

3. But now comes the twist: Included in the mix is normal C but also a modified C (C*), which terminates replication at that point. Now, the next base in the parent DNA is a G. So the next base to be added to the growing new chain is C, making it AAC. But instead of normal C, DNA polymerase, which doesn't know the difference, might add C*, making the new strand AAC*.

4. The fate of these two new strands is now different. If normal C is added, replication continues: AACACGTAATTTGA … But if C* was added, the strand stops right there and ends up much shorter: AAC*.

5. At the end of replication, the new DNA strands are separated. Suppose C* has a dye attached so that it shines red under laser light. The various DNA strands are separated by size and detected by laser light. The only one that shines red is AAC*. So we know that the third base is C!

6. In separate reactions, modified T, G, and A are used in the replication, each base with a different

colored dye: T (green), G (blue), and A (yellow). So in our example, when there are fragments that end in A, they will shine yellow; these will be 1, 2, 4, 8, and 14 bases long. So there is an A at positions 1, 2, 4, 8, and 14!

7. The whole process is now automated. The scientist puts the DNA in one end and gets the sequence on a computer at the other.

B. There are powerful methods to examine sequences.

1. Getting the DNA sequence isn't enough. We want to know what it means. What is the sequence of a protein-coding region, and what is the amino acid sequence of the protein? What are the sequences of gene control regions, like promoters? Where are the noncoding introns?

2. We know the genetic code, so with a DNA sequence we can surmise the coding sequences. But this too is now done by computer. The DNA sequence is entered in a DNA search program (there are several available), and the information comes out.

3. But there is more: The program also checks to see if the DNA sequence has ever been seen before in nature. All new sequences are sent to a central database (and there are millions of sequences out there now). If the sequence has been seen, you now know what the protein might be.

IV. RNAi is used to inhibit gene expression.

A. In the early 1990s, scientists in the Netherlands discovered that the expression of several genes for flower color in petunia plants could be turned off simultaneously. They called this "gene silencing."

B. Scientists soon found this phenomenon in other organisms. In 1998, Andrew Fire and Craig Mello found out how it works in a tiny worm. (This was the same worm whose genome was sequenced, as we described in Lecture Nine.)

1. Gene silencing occurs when the cell makes a double-stranded RNA, with one of the strands complementary to the mRNA for the gene. To be clear: If the target mRNA that wants to be translated to protein has the sequence:

AAAUGAAGUU,

the anti-RNA will be UUUACUUCAA,
and its other strand AAAUGAAGUU.
After being made, this double-stranded RNA gets bound up with an "escort" enzyme complex that guides it to the ribosome and peels off the "anti" strand. Once the "anti" strand binds to a target mRNA, the expression of that mRNA is blocked.

2. Fire and Mello's explanation soon was found to be true for other animals as well. Gene silencing is a way for cells to turn off genes in humans, flies, and petunias.

3. This science has led to technology: If most eukaryotes have the machinery for RNAi, we can introduce RNA against any gene and use it to silence that gene, and only that gene.

a. This is powerful manipulation for asking "what if" questions: If we hypothesize that the expression of gene A causes phenotype B, then we can turn off gene A by RNAi and see if B does not occur.

b. This extends to diseases caused be gene expression. Suppose cancer is caused by a gene being inappropriately expressed; we can use RNAi as a drug to turn the gene off, and shut off the cancer as well. No wonder Fire and Mello won the Nobel Prize.

c. In my own research on lung cancer, I have shown that if I add RNAi directed against the mRNA for a protein that removes life-saving drugs from tumor cells, the protein is no longer made and the cells become once again sensitive to tolerable doses of chemotherapy.

Essential Reading:

Harvey Lodish, Arnold Berk, Paul Matsudaira, Chris Kaiser, Monty Krieger, Matthew P. Scott, Lawrence Zipursky, and James Darnell, *Molecular Cell Biology*, 5th ed. (New York: W. H. Freeman, 2005), chap. 9.

David Sadava, Craig Heller, Gordon Orians, William Purves, and David Hillis, *Life: The Science of Biology*, 8th ed. (Sunderland, MA: Sinauer Associates; New York: W. H. Freeman and Co., 2008), chaps. 11 and 16.

Supplemental Reading:

Kerry Mullis, *Dancing Naked in the Mind Field* (New York: Vantage Books, 2000).

William Thieman and Michael Palladino, *Introduction to Biotechnology* (San Francisco: Benjamin-Cummings, 2004).

Questions to Consider:

1. Two ways to manipulate DNA are technological advances that came from basic scientific research. Can you trace the flow from basic to applied for PCR and for RNAi? What does this tell you about the need for public support of basic research?

2. Recently, DNA fragments that are 400,000 years old were isolated from specks of permafrost in Siberia. The DNA is both from plants (28 species) and from animals (mammoth and bison). Does this bring the *Jurassic Park* scenario any closer?

Lecture Fifteen
DNA in Identification—Forensics

Scope: Genetic identification of people has been done by analysis of the expression of genes, such as blood typing. There are few genes for blood type, so many people share the same genes. As a result, this method can only eliminate the identity of a person but not positively identify an individual. Short tandem repeats (STR) are DNA sequences a few base pairs long that are repeated side by side in an inherited pattern. There are thousands of STRs scattered throughout the human genome. They are polymorphic, meaning that there are rare and common alleles (in this case, repeat number). With many STRs and alleles, the probability that two people are alike is extremely low, and so this can be a positive identification. A tiny amount of tissue, even a single cell, can be the starting point for DNA identification. It has many uses, including in criminal investigations, in historical analysis, and in the case of disasters where there are human remains, among other places.

Outline

I. Opening story: Baby 81 was identified by DNA.

A. The tsunami of December 26, 2004, left thousands of dead bodies in its wake.

1. A baby, Abilass Jeyarajah, was torn from his mother's arms when the tsunami hit Sri Lanka. Amazingly, he survived and was brought to a local hospital while his parents frantically looked for him.

2. Because he was the 81st infant brought in that day, he was called "Baby 81."

3. A few days later, his parents came to the hospital, where they heard there were unclaimed babies, both dead and alive. Joyfully, they were reunited with Abilass.

4. However, in the previous two days, other couples had come to the hospital searching for their missing babies, and eight had claimed Baby 81 was theirs. The question ended up in court.

5. In court in February 2005, Judge M. P. Mohaiden had not just wisdom to rely on to find out the true parents: He had the evidence from genetics and DNA.

B. The genetic evidence in DNA clearly identified Baby 81's parents.

1. As described in Lecture Nine, we humans are over 99% identical in our 3.2 billion base pairs of DNA. That leaves a lot of room for variation.

2. A major source of variation is in repeated sequences, with individuals having different numbers of repeats in an inherited pattern.

3. When Baby 81's DNA was examined for these repeats, he had a pattern that was shared by his true parents but not by any of the other eight couples.

II. Genetics by phenotype can be used to identify individuals.

A. Phenotypes that reflect alleles have been used for identification.

1. In genetics, it is possible to predict genotypes and phenotypes from inheritance patterns.

2. Think back to Mendel: If short pea plants are recessive to tall, we can predict that two short parents will produce short offspring and not tall ones. That is, tall plants would not ordinarily come from short parents.

3. Now consider human genetics and identifying Baby 81. One phenotype that is clearly inherited and not subject to environmental variation is blood type. This is due to proteins expressed on the surface of red

blood cells. Which protein is expressed is determined by the genetic alleles present. There are three alleles: A and B are codominant, and O is recessive.

4. For example, a person with type A blood either inherited an A allele from each parent or A from one parent and O from the other. A person with type AB blood inherited A from one parent and B from the other, etc.

5. Consider Baby 81. We don't have the real data, but suppose Baby 81 was type AB. That would mean that his parents would have to pass on A and B, and neither of them could be type O. Blood type analysis could eliminate parents in some cases. But a real problem is that there are many people in Sri Lanka with A or B alleles. There has to be a better method, and there is.

B. The HLA (human leukocyte antigen) system has more alleles.

1. This genetic system codes for proteins on the surfaces of cells, including white blood cells. It is used in transplants, since the cell surface is recognized by the immune system if it is different.

2. There are many more alleles: HLA-A has 23, HLA-B has 47, HLA-C has 8, and HLA-D has 23. A person might inherit A11 B16 C3 D11 from one parent and A9 B12 C3 D20 from the other parent.

3. With more alleles and four genetic systems, it is more likely that people will be different and parents and children can be matched.

4. Problem: You need well-fixed tissues or blood; the HLA proteins are not always present on all cells; there is a lot of mixture of the genes in gamete production.

III. Genetic analysis of DNA variants is the best identification.

A. To match Baby 81 with his parents, DNA fingerprinting was done. It works as follows.

B. The human genome contains short sequences, 2–10 base pairs long, that are repeated in tandem: The STR might be TCAT; and the sequence might be TCATTCATTCATTCAT, repeated four times.

1. There are about 10,000 different STRs in the human genome. The repeat number is inherited.

2. Of the many STRs, 13 scattered throughout the genome are used in DNA identification. These are short sequences repeated up to five times, and they have common and less common alleles (repeat numbers). We call a gene with common and less common alleles polymorphic. A population survey must be done to find out the frequencies of these alleles in a population before we do identification.

3. Suppose we are dealing with two of the STR loci and they have alleles (repeat numbers) that I will call A, B, and C.

4. For STR 1, A is 1 in 100 (0.01), B is 1 in 5 (0.2), and C is 4 in 5 (0.8).

5. For STR 2, A is 1 is 10 (0.1), B is 1 in 2 (0.5), and C is 2 in 5 (0.4).

6. Here is the key argument, and it comes from Mendel and probability. For a person to be carrying alleles A and B of STR 1, the probability is the product of the two probabilities, or 0.01×0.2, which is 0.002. For a person to be carrying A and B of STR 2, the probability is 0.1×0.5, which is 0.05.

7. Now comes the important number: For a person to have A and B from both STR 1 and STR 2, the probability is once again the product, which is 0.002×0.05, or 0.0001, which is 1 in 10,000.

8. With 13 gene STR systems, the probability of two people having the same genetic markers is vanishingly low. That is why DNA matches are used in identification.

C. DNA profiling is done by cutting DNA with restriction enzymes and sizing the region that has the STR for the number of repeats.

1. For this, about 1 ug (microgram; 1 millionth of a gram) of DNA is needed. A single cell has 1 millionth of that—about 1 trillionth gram of DNA.

2. To get information from this, PCR is used and the millionfold increase in that cell's DNA makes it ready to analyze. A single cell of hair, skin, or blood is enough to get going.

IV. There are interesting examples of DNA identification.

A. Sir Alec Jeffreys developed DNA fingerprinting.

1. In the early 1980s, Professor Alec Jeffreys at the University of Leicester, UK, was studying genetic

differences between individuals. He was looking at the genes for the muscle protein myoglobin and compared DNA from seals with humans. To his surprise, there were common, short, repeated sequences in many animals. When he examined the STRs in a human family, the parents and children had them. But he noticed that the children's sequences were a composite of the parents, indicating that they were inherited. Realizing he had a way to identify people by DNA, he published his findings in spring 1985. The genetic floodgates opened.

2. The first case involved immigration. A family from Ghana had immigrated to the UK. When one of the four sons visited Ghana, he was detained on return by British immigration officials because he had a forged passport. They refused readmission, claiming he was not the son from the UK but a cousin from Ghana sneaking in. Jeffreys did DNA analyses of the mother and the three undisputed sons (the father was missing), as well as the son in dispute, and this showed that he was definitely her son.

3. Soon, Jeffreys was called to a criminal case in Leicestershire. Two girls had been raped and murdered in the same area in similar circumstances, two years apart. A man in jail had confessed to the second crime, but claimed innocence of the first one. Jeffreys did DNA analyses of the victims, the suspect, and the semen found on the victims. It showed that the same man had probably committed both crimes, as police suspected—but it was not the man in custody who had confessed!

a. The police asked all men in the area to give a blood sample for DNA, and 5000 men did so. Over 90% of them could be eliminated by blood typing (HLA on the semen). When DNA analyses were done on those samples remaining, there was still no match. Then, a woman overheard a man saying that he had given two blood samples, one for himself and one in the name of a friend. That friend, when truly tested, turned out to the killer.

b. DNA is now used widely in forensic cases.

B. An interesting use of DNA identification is in historical analysis.

1. In July 1918, with the Russian Communist Revolution raging, the last Romanov Emperor, Tsar Nicholas II, his wife, and three of their children were killed in a town in the Ural Mountains and buried in a shallow, unmarked grave. Seventy-three years later, in a new Russia, two amateur historians found what they thought was the grave. The sizes of the skeletons were consistent with the family, and gold dental fillings certainly suggested that they were rich. But the skeletons were too damaged for further identification.

2. Fortunately, the bones had cells with DNA that could be analyzed. STR alleles in the bones were compared with those of survivors from the Romanov family. DNA from a great-granddaughter of the Tsar's sister and a great-grandson of his aunt, as well as the body of the Tsar's brother (buried in 1899), were tested and showed the same alleles as the dead family. This proved that they were the Romanovs. They were reburied in a state funeral.

C. DNA identification has many other uses.

1. DNA was used to identify the victims in the World Trade Center terrorist attack.

2. The military genotypes the DNA of all personnel to aid in identification in battle. There are over 3 million DNA types stored in the U.S. military database.

3. DNA databases, with stored STR genetics, are being built up all over the world. For example, in the UK, people arrested for serious crimes are DNA typed, and there are now over 3.5 million people in the database. This has led to a great increase in "cold hits," where criminals are identified for arrest only on the basis of DNA left at the crime scene. In the U.S., the FBI has about 3 million DNAs stored and typed of people convicted—first of sex crimes but now of all serious crimes.

Essential Reading:

Ricki Lewis, *Human Genetics*, 7th ed. (New York: McGraw-Hill, 2006), chap. 14.

David Sadava, Craig Heller, Gordon Orians, William Purves, and David Hillis, *Life: The Science of Biology,* 8th ed. (Sunderland, MA: Sinauer Associates; New York: W. H. Freeman and Co., 2008), chap. 16.

Supplemental Reading:

Robert Massie, *The Romanovs: The Final Chapter* (New York: Random House, 1995).

Philip Reilly, *Abraham Lincoln's DNA and Other Adventures in Genetics* (Woodbury, NY: Cold Spring Harbor Laboratory Press, 2000).

Questions to Consider:

1. In 1994, Californians voted to have mandatory DNA identification testing not just of people convicted of felony crimes but of those arrested as well. This will create a similar database to that in the UK, which has been doing this for some time and gets many "cold hits" of suspects police never would have sought from DNA left at crimes. Are there privacy issues about the government holding DNA samples?

2. Trace the evolution of genetics in the courtroom from using phenotypes such as blood types to exclude people to using DNA to identify people. What are the genetic-statistical arguments used? Are they convincing?

Lecture Sixteen
DNA and Evolution

Scope: Charles Darwin looked at nature and the environment and proposed two ideas that unified biology. First, he related organisms by descent with modifications from a common ancestor. Second, he proposed that these modifications become a permanent part of organisms by natural selection. In his view, confirmed not only by his extensive observations but by much data since then, many organisms of a species are born, and they differ slightly in terms of genetics. Those genetic characteristics best adapted to the environment at the time are passed on (selected) to the next generation. Both spontaneous and induced DNA mutations provide genetic variations. Organisms carrying advantageous protein phenotypes are selected for reproduction. Genetic bottlenecks, in which a few organisms are responsible for a large population, can lead to a special set of genes in a population. Some organisms evolve by DNA changes that are neutral to selection. These changes can serve as a molecular clock to determine relatedness of organisms.

Outline

I. Opening story: Darwin proposed natural selection to explain evolution.

A. Charles Darwin was born on the same day as Abraham Lincoln, February 12, 1809.

1. The son of a society doctor and a mother from the Wedgwood family, Darwin was sent to medical school in Edinburgh but dropped out and transferred to Cambridge to study for the ministry.

a. He studied botany under Professor John Stevens Henslow and was fascinated by natural history.

b. When Darwin graduated in 1831, Henslow recommended him to Capt. Robert Fitzroy, who was looking for an amiable companion and naturalist for his ship the *Beagle*, which was about to leave on a surveying voyage first to South America and then around the world.

2. Before they left, Fitzroy gave Darwin a copy of a recently published book on geology that explained the very slow changes in rocks over time.

B. Darwin changed biological science.

1. The *Beagle* left Plymouth Harbor on December 27, 1831, and returned to Falmouth after a round-the-world voyage on October 2, 1836.

2. While Fitzroy did his job, Darwin did his—and changed biology forever. Darwin made careful observations of both the organisms he saw and the environment. He saw animals and plants that appeared specifically adapted to their environments. This was not new; people had seen this since ancient times and attributed it to special creation.

3. Darwin noticed resemblances between organisms in different places.

a. Organisms on the Galapagos Islands in the middle of the Pacific Ocean were similar to ones of that type on the coast of South America (Chile).

b. Organisms in the temperate regions of South America more closely resembled those in the tropics of South America than their temperate counterparts in Europe.

c. Fossil organisms in South America resembled living organisms he saw on that continent more than fossils and living organisms in Europe.

d. How could this be if organisms were specially created for each environment? Wouldn't all the organisms living in all the temperate climates be the same?

4. Darwin proposed his first idea: The organisms in South America had a common ancient ancestor (that had traveled to the Galapagos). He called this "descent with modification." His geology book and observations confirmed the idea that the Earth was very old.

5. Darwin then proposed his explanation for descent with modification (evolution): natural selection.

a. Many members of a species (type of organism) are born.

b. There are inherited variations among these organisms. These occur randomly.

c. The changing environment selects those organisms with the best adapted variations for survival and reproduction. In this way, organisms change over generations of time.

6. Note that genetic changes are random, but selection is directed to the environment at that time (and not any future time, when the environment may change). Evolution is not progress; it is not linear, but branched.

II. There is a lot of evidence for evolution by natural selection.

A. Agriculture: In the Book of Ruth in the Hebrew Bible, Naomi sent her daughter Ruth to lie with the rich man Boaz on the barley threshing floor.

1. Before barley became a crop, its stalks would shatter when harvested. This is good for the plant (seeds on the ground) but not the farmer (seeds on the ground!). Better to pick up the stalk and then thresh it to separate out the seeds.

2. Humans harvested barley plants that did not shatter and threshed them. They ate most of the seeds and planted some for next year's crop. Thus, they selected for the characteristic of not shattering.

3. Much of agriculture has been more conscious selection. For example, a wild mustard plant (*Brassica*) grows like a weed with flat leaves. Selection led to different vegetables from genetic variation of the same species:

For terminal growth: cabbage For leaves: kale

For lateral growth: Brussels sprouts For stems and flowers: broccoli

For stem: kohlrabi For flowers in clusters: cauliflower

B. In England in 1842, the bison moths were 2% dark color, 98% white. In 1898, they were 95% dark. Dark color is due to a dominant allele, so the frequencies of the alleles were not related to whether they were dominant or recessive. What happened for this change through time (evolution)?

1. Black was a dominant allele but was selected against because the dark-colored moths on the light tree trunks would get picked off by birds.

2. Industrial plants during the 1800s in that region spewed out black smoke that coated the tree trunks; now the light moths were picked off by birds. So the dark moths were selected for reproduction.

3. In the 1950s, this hypothesis was confirmed experimentally.

C. On the Galapagos, Darwin famously saw finches with big beaks (that eat big and small seeds) and small beaks (that eat only small seeds).

1. In 1977, there was a severe drought, and fewer seeds were produced. Both big and small seeds were affected. With fewer seeds, big-beaked birds were favored because they could eat any size, but the smaller ones could eat only small ones—and these ran out. The population of finches fell from 1200 to 200.

2. Over the next decade, there was evolution to big-beaked birds, because when there were few seeds, there were equal numbers of big and small seeds, and the big-beaked birds had a selective advantage because they could eat both kinds.

D. Fossils show evolutionary change over long periods.

1. As explorers went to new lands and dug for canals and mines, they saw rocks in layers, with the top rocks the most recent. There were bones and plant impressions in these rocks, the remains of ancient organisms. The deepest (oldest) rocks had fossils that least resembled modern organisms.

2. When these fossils were examined, in many cases there was a progression over time: evolutionary change.

E. Homologous structures are evidence for common ancestry.

1. Examination of anatomy of current organisms shows homologous structures that have been adapted for different functions: For example, the front limbs of vertebrates all have a large bone, then smaller ones, and tiny ones at the end. In humans, these are the arms and fingers. But the same pattern is adapted in birds (for flying), seals (for swimming), and sheep (for running).

2. The biochemical unity of life (same DNA, genetic code, etc.) also can be explained by common ancestry.

III. Protein changes due to DNA mutations explain natural selection.

A. Mutations are the raw material of evolution, and they are nondirected.

1. Spontaneous mutations are due to errors in DNA replication.

a. They are rare: The usual rate is 1/100,000. But with millions of germ line cells there is a good chance of it occurring: Germ line cells produce eggs/sperm.

b. Duplications are an important mutation source: One or more copies of a gene are made by DNA polymerase "stuttering," as was mentioned earlier. Extra copes of a gene mean that one copy can mutate and not cause adverse effects because there is still a good copy.

2. Induced mutations are caused by environment—most are natural factors (e.g., ultraviolet radiation from sunlight damages DNA). The damage can be repaired but not always, and this can end up as a mutation. Most mutagens are natural, such as substances in our diet.

B. The phenotype protein is selected.

1. Antibiotic resistance in bacteria is inherited.

a. All bacteria have an enzyme that breaks down certain waste products.

b. Penicillin, an antibiotic made by molds, binds to and inhibits another bacterial enzyme that makes the cell wall.

2. Some bacteria have a mutation in the gene coding for the first enzyme such that it breaks down penicillin.

a. This makes these cells antibiotic resistant.

b. In a normal bacterial population, this mutation is rare. But in the presence of penicillin, the few bacteria carrying it are selected for.

IV. Some DNA changes in evolution are not selected.

A. Some DNA mutations do not lead to phenotypic changes.

1. Hemoglobin—There are hundreds of DNA changes that:

a. do not lead to a different amino acid (the genetic code is redundant) and

b. lead to an amino acid change that is not selected for because the reproductive fitness does not change (e.g., at position 7 there is a change of:

A to G
T C

with amino acid change but no effect).

2. Many DNA mutations in genes are of this type.

B. Some DNA changes lead to evolution but not by natural selection.

1. There used to be tens of thousands of elephant seals in the North Atlantic.

a. Hunting reduced the population to about 20 in 1890.

b. The species was conserved, and now there are 30,000. They are genetically almost the same because they all came from a few ancestors. So the genes that were in those seals in 1890 are present in all of their descendents. This is an example of a population bottleneck. It is one way that evolution can occur without selection.

2. In 1968, population geneticist Motoo Kimura proposed that evolution can occur by neutral mutations that are not subject to natural selection but accumulate in a population of organisms.

a. This can occur in small populations (see the seals, above).

b. The rate of accumulation of mutations in this case is equal to the mutation rate of the allele (or vice versa). For example, for a protein, one can trace relationships from fossils and give an age when the last common ancestor lived (e.g., for insects vs. vertebrates this was 600 million years ago). Then look at a protein (e.g., cytochrome c) that both insects and vertebrates have and at gene changes. Then calculate changes per year: It turns out to be one per 20 million years. Now, if there are two organisms and we want to see when they last had a common ancestor for this gene, we look at the differences between them and then calculate based on 1/20 million years. This is called "molecular phylogeny."

Essential Reading:

Benjamin Pierce, *Genetics: A Conceptual Approach* (New York: W. H. Freeman, 2005), chap. 23.

David Sadava, Craig Heller, Gordon Orians, William Purves, and David Hillis, *Life: The Science of Biology*, 8th ed. (Sunderland, MA: Sinauer Associates; New York: W. H. Freeman and Co., 2008), chaps. 23–26.

Supplemental Reading:

Charles Darwin, *The Origin of Species* (New York: Random House, 1979).

Richard Dawkins, *The Selfish Gene: 30th Anniversary Edition* (Oxford: Oxford University Press, 2006).

Douglas Futuyma, *Evolution* (Sunderland, MA: Sinauer Associates, 2005).

Questions to Consider:

1. Compare evolution (change in allele frequencies in a population through time) by natural selection, a genetic bottleneck, and neutral mutations. Which do you think accounts for most evolutionary changes?

2. Mendel published his experiments on genetics (1866) after Darwin published his book *The Origin of Species* (1858). In fact, Mendel read Darwin, but there is no evidence of the reverse. How do you think Darwin would have used Mendel's gene concepts to support the theory of evolution by natural selection?

Lecture Seventeen
DNA and Human Evolution

Scope: The mechanisms that explain evolution in the rest of the living world apply to humans as well. In sickle cell disease, a harmful genetic variant (allele) has been subjected to natural selection because it affords protection against a more serious disease, malaria. There are many examples of population bottlenecks in humans, which lead to distinctive frequencies of certain alleles. Molecular clocks can be used to trace human origins through DNA markers on the Y chromosome (males) and mitochondrial DNA (females). The origins and spread of human populations can also be traced in this way. Comparisons of the human genome with the recently sequenced chimp genome reveal some hints of the evolution of humans. Some gene differences as well as inserted sequences may be important. This is underscored by the findings of evolutionary developmental biology, in which a relatively small set of genes appear to trigger key events in the fascinating processes that occur in the embryos of many complex animals.

Outline

I. Opening story: Sickle cell disease is an example of natural selection in humans.

A. People often think they are "above nature"—in control of their destiny. Sickle cell disease proves that natural selection acts on humans, as on any other species.

1. Sickle cell disease is a disorder affecting the structure of red blood cells. It is inherited as an autosomal recessive, meaning that of the 1100 children born in the U.S. every year with the disease, in most cases their parents were healthy carriers for the harmful allele.

2. Normally, red blood cells, which have the red oxygen-carrying pigment, hemoglobin, are donut-shaped and flexible so they can pass through narrow blood capillaries. They have a lifetime of about 120 days and

are replaced by new cells that are made in bone marrow.

3. People with this disease have red blood cells shaped like sickles. They are brittle and tend to block narrow capillaries, starving the tissues involved of life-giving oxygen. So patients tend to have lung, spleen, and kidney damage, and pain in the abdomen, legs, and chest. The abnormal shape targets the cells for early destruction (after about 16 days), so patients have a low blood count (anemia).

4. There is no cure for this chronic disease. Treatments include pain medications, antibiotics (the spleen damage makes them especially vulnerable to transfusions), and blood transfusions. A new drug, hydroxyurea, appears to prevent sickling and has had some success.

5. Sickle cell disease was the first human genetic disease whose molecular nature was described. In 1948, Linus Pauling and Harvey Itano pinpointed the abnormal phenotype on the protein hemoglobin. A DNA base-pair change in amino acid coding position 6 of 141 in the gene coding for the protein portion of hemoglobin:

A to T
T A

leads to a single amino acid change in the protein. This results in abnormal hemoglobin folding, which leads to bad binding of oxygen and sickle-shaped red blood cells. Knowing this precise phenotypic description led to the development of hydroxyurea.

B. The population distribution of sickle cell disease is unusual.

1. The sickle allele is not universally distributed among humans. In fact, it is relatively common in some populations, but rare in others. The disease apparently originated in Africa, and is still relatively common there: In the U.S., 300 million have the disease, and about 1100 babies are born with it annually. In Nigeria, 120 million have the disease, and 80,000 babies are born with it annually. The transatlantic slave trade brought people carrying the allele for sickle cell disease to the U.S., where it occurs mostly in African Americans.

2. The appearance of such an allele in a certain populations prompted geneticists to ask how it happened and why it is maintained there. The first question is easily answered by a founder effect, where a spontaneous mutation occurred in some people that spread to their descendents. The second question is not so easy. Why is a clearly harmful mutation, which certainly lowers reproductive fitness, maintained at such a high level?

3. In 1954, geneticist Anthony Allison noticed a similarity in the geographic distributions of people with the sickle cell disease and of malaria, which continues to be a scourge of Africa and other tropical regions. The organism that causes malaria lives in human red blood cells during part of its life. Allison proposed that if cells were sickled the parasite might not reproduce. So people with this disease were resistant to malaria, a far more harmful one. In fact, it is the heterozygous carriers for the allele that are at the greatest advantage against malaria.

4. The term “balanced polymorphism” describes this situation.

II. Population bottlenecks lead to unusual allele frequencies in humans.

A. Hereditary asthma is frequent on an island.

1. In the middle of the South Atlantic Ocean is the island of Tristan de Cunha. The island was fist settled by a Scot, William Glass, who brought his family there in 1817. They were joined by a few settlers from another island, but after Glass’s death in 1856, most of the 120 or so descendents left for the Americas. The 300 or so people on the island today all came from 12 of the group in 1850s.

2. The group of 120 that had descended from Glass had many different alleles. The 12 who remained on the island and “begat” the subsequent population had their own subset of these alleles that were then passed on.

3. For instance, the current islanders have the highest rate of hereditary asthma in the world, due to a group of alleles that affect their respiratory system. This is an example of a population bottleneck. It is one way that evolution can occur without selection.

B. There are some "Jewish diseases."

1. The Jewish people have been subjected to many genocidal population reductions over the centuries. After each of these events, a small population remained, and this caused bottlenecks.

2. It is not surprising that there are a number of "Jewish genetic diseases," more common in them than in other groups of people. These include Gaucher disease, Tay-Sachs disease, dystautonomia, and Canavan disease.

III. Molecular clocks can trace evolutionary relationships.

A. In the previous lecture, we saw how alleles can accumulate in organisms without selection, just driven by the mutation rate, and that this can serve as a molecular clock. If we compare two organisms, we can estimate when they had a common ancestor and thereby gain insight into how closely related they are.

1. The human Y chromosome is inherited through males and has over 30 regions that are nonselected and mutate at a constant rate. If we compare males around the world with one another, we can determine when they had a common ancestor. My brothers and I probably have identical Y markers: Our common ancestor is recent. My third cousins and I have different Y markers: We have a common ancestor longer ago. And so on, to the most diverse humans. This analysis traces Y chromosome "Adam" to about 80,000 years ago. Note: "Adam" is not the first man or the only one living at that time, just the male whose DNA alleles end up in males today.

2. DNA in the mitochondrion, a part of the cell outside the nucleus that is passed on through the female, can also be used as a molecular clock. In this case, mitochondrial "Eve" lived about 150,000 years ago.

3. Molecular markers can be used to trace migrations of populations.

a. As the human population has spread over the Earth, those starting the new populations have specific genetic markers (e.g., skin color over time).

b. DNA markers (Y and mitochondrion) can be used to identify populations and, along with the clock, when they diverged.

c. This has set up the "genographic" and other projects to trace human origins and migrations by DNA markers.

B. Genealogy can be done by genetics.

1. In Jewish tradition, priests called Cohens (hence the common surname) are males descended from Aaron, brother of Moses, who lived 3300 years, or 100 generations, ago. A Cohen, Dr. Karl Skorecki, was attending synagogue services, and when a Cohen was called for, a visitor named Cohen with very different skin color, eye color, facial features, and hair color than Skorecki's stepped up. "Can we both be descended from the same man?" he wondered.

2. Analysis of Y chromosomes of Cohens from around the world provided an affirmative answer. Virtually all had a distinctive genetic marker on the Y chromosome.

3. The use of such DNA sequence markers to trace the origins of human groups has become a major effort. Web sites invite people to send in a tissue sample and have their DNA analyzed to find their family, ethnic group, or geographic origins.

IV. Genome sequences highlight human evolution.

A. Several phenotypic characteristics make us human.

1. Culture: the use of tools; genes have not yet been found for this.

2. Walking upright: A single gene, AHI1, appears to control placing one foot in front of the other. Humans have it; chimps have a mutated sequence of it.

3. A large brain: A chimp gene, MYH16, is mutated in humans. In chimps, this strong muscle protein in the jaw prevents skull growth to accommodate the brain. In humans, the mutation prevents the strong jaw and the brain can grow.

4. Thought: The human genome has repeated sequences for a number of genes expressed only in the brain. Chimps do not.

B. The chimp genome was sequenced in 2005; it is 2.9 billion base pairs (like the human genome).

1. According to both fossil and molecular clock data, humans and chimps had a common ancestor that diverged about 6 million years ago.

2. Comparison of the two genomes shows that in humans the average protein differs by two amino acids, there are 35 million single base-pair changes, and there are 5 million insertions and deletions that account for 70 million base pairs. The latter may be key (see below).

3. Human genes that show the most differences with the chimp include ones involved in brain chemistry, speech, and disease resistance.

4. This is only the beginning of this comparative effort.

C. The evolution of development is surprising.

1. Development is the process by which an organism goes from fertilized egg to birth. In humans, it involves expression of about one-third of the total genes.

2. The amazing aspect of development is that the basic tool kit for making a complex organism is similar throughout biology. A limited number of genes is involved in making a worm or fly or mouse or person.

3. The first hint of this came with the discovery that genes that control which segment of a fruit fly developed which structure are similar to the genes that control the development of the body plan in humans. This initiated a new field: evo-devo (evolutionary developmental biology).

a. Example: Over 90% of animals have eyes, but there are very different types. We have camera eyes, but insects have compound eyes (many eyes in a single structure). In 1915, a mutant fruit fly was discovered that had no eyes. This remained a lab curiosity until developmental geneticist Walter Gerhing found that the actual protein phenotype was a protein that stimulates transcription of eye genes.

b. When the DNA sequence for the fly eye factor was compared by computer to other genomes, a similar gene was found in the human genome. Gene swapping experiments by recombinant DNA showed that the fly gene could stimulate mouse eye formation in the embryo and vice versa. So the mechanism for control is the same throughout the animals.

4. This leads to rapid evolution: Just adding (or mutating) a promoter of a transcription protein can activate a gene that was not active before. Since the human and chimp genomes have lots of inserts and deletions, this may be the key difference.

Essential Reading:

Ricki Lewis, *Human Genetic*s, 7th ed. (New York: McGraw-Hill, 2006), chap. 16.

David Sadava, Craig Heller, Gordon Orians, William Purves, and David Hillis, *Life: The Science of Biology,* 8th ed. (Sunderland, MA: Sinauer Associates; New York: W. H. Freeman and Co., 2008), chaps. 20 and 23.

Supplemental Reading:

Elof Axel Carlson, *The Unfit: A History of a Bad Idea* (Woodbury, NY: Cold Spring Harbor Laboratory Press, 2001).

Linda Stone, Paul Lurquin, and Luca Cavalli-Sforza, *Genes, Culture and Human Evolution: A Synthesis* (Malden, MA: Blackwell, 2007).

Questions to Consider:

1. Can you think of human populations that have a specific set of alleles? How do you think these alleles ended up in the population?

2. In the molecular evolution story, "Adam" and "Eve" did not live at the same time. How can this be so?

Lecture Eighteen
Molecular Medicine—Genetic Screening

Scope: Genetic testing for mutant alleles in people is now possible at both the phenotype and gene levels. It can be done in newborn babies or adults, and even prenatally. Screening is done either on the whole populations (e.g., all newborns for phenylketonuria) or on populations in which an allele is prevalent (e.g., African Americans for sickle cell disease). The tissues used for screening range from blood serum to any cell for DNA testing. The tests must be both reliable (repeatable) and valid (an accurate reflection of the genetic

condition). Newborn screening for disorders such as phenylketonuria and hypothyroidism has been successful in terms of optimizing human potential and economic and social benefit. DNA testing can be effective only if the mutation targeted is known. The next frontier in genetic testing is the identification of alleles for altered susceptibility to drugs and for susceptibility to complex diseases.

Outline

I. Opening story: an early diagnosis.

A. Diagnosis of a genetic disease is now possible on early embryos.

1. John and Mary were concerned about their infant son, who was not gaining much weight. Jimmy was listless and seemed to have a cough all the time. Searching the internet over the weekend before a Monday appointment with the pediatrician, Mary came across cystic fibrosis as fitting Jimmy's symptoms. When she licked his neck and tasted salt, she was pretty sure. The doctor confirmed it.

2. Cystic fibrosis is due to a recessive allele and is one of the most common inherited diseases, with a carrier frequency of about 1 in 50 and a birth frequency of 1 in 2500. It is caused by a mutation of the CFTR gene, so that its product, a protein that allows chloride salt into and out of cells, does not work. As a result, mucus accumulates in the respiratory system and blocks the digestive secretions of the pancreas from reaching the intestine.

3. The disease appears during childhood, with persistent infections, digestive problems, diabetes, and later, infertility. Treatment of these symptoms has improved the lifespan for these patients: 40 years ago, few survived to their teens. Now, most survive until their mid-30s.

4. The most common mutation in the CFTR gene in Caucasians is a three-base-pair deletion of DNA at amino acid coding position 508. This results in a single amino acid missing in the large, 1480-amino-acid protein. This tiny deletion is enough to inactivate the protein. This is the mutation that Jimmy had. And because this is a recessively inherited disease, he had two mutant alleles—both John and Mary were carriers, with one normal and one mutant allele.

5. Two years later, with Jimmy's condition stable, John and Mary decided to have another child. This time they wanted a child without the disease. They knew the chances for a normal child were three in four. Fortunately, a new test is available to detect the cystic fibrosis allele.

6. First, Mary was given drugs to augment the number of eggs she would produce; instead of just one in the next monthly cycle, eight were matured. In the clinic, the eggs were removed, and John's sperm was used to fertilize them. Six eggs were successfully fertilized. Over the next three days they began to divide—the egg produced two cells, then four, then eight.

7. At this point, the thin membrane surrounding the cells of one of the embryos in the dish was punctured and a tiny straw used to suck out one of the eight cells. PCR was used to amplify the single cell's CFTR allele, and the presence of the missing three base pairs and normal CFTR allele were probed. The probe did not detect the mutant allele, only normal ones. The remaining seven cells were enough to go on to divide further, and after a week the embryo was implanted in Mary's womb. Thirty-seven weeks later, she gave birth to Susie, a child without cystic fibrosis.

B. This preimplantation genetic diagnosis, while not widespread, is the culmination of reproductive and genetic technologies.

II. Genetic tests identify alleles in people.

A. Screening involves testing people, often without symptoms, for the presence of an allele.

1. Screening should provide information that is useful to the patient (or guardian) for decision making and/or for the physician in treatment.

2. Screening can be done on different groups.

a. An age group when the allele gets expressed: newborns for cystic fibrosis.

b. Targeted populations: African Americans for sickle cell disease.

3. Screening can be done on tissues expressing the phenotype or any cell's DNA. For example:

a. Blood: serum for proteins, cells for DNA.

b. Mouth: cells washed out for DNA; this was done on Saddam Hussein when he was found.
c. Hair, semen: for forensic purposes.
d. Fetal tissues (embryo, as above; chorionic villi at 6 weeks; amniotic fluid at 12 weeks).

4. A test for an allele has two components.

a. Reliability: Is it reproducible? Do two radiologists looking at a mammogram or two pathologists looking at a PCR result come to the same conclusion?

b. Validity: Does the test actually indicate the allele? This in turn has two components.

i. Sensitivity: The proportion of people with the allele who test positive. Failure here is a false negative.

ii. Specificity: The proportion of people without the allele who test negative. Failure here is a false positive.

B. Phenotype screening at the molecular level involves looking for the abnormal protein.

1. In sickle cell disease, the change in hemoglobin involves an amino acid with a neutral charge replacing the normal amino acid with a negative charge. Although this is only one amino acid in hundreds, the slight difference in overall charge of the protein (less negative in the mutant) is easily detected in a blood sample. This takes an hour.

2. For phenylketonuria, screening is harder. As you may recall from Lecture Six, the protein missing in PKU is an enzyme present in tiny amounts. And it occurs only in the liver. Taking a piece of liver out to make the diagnosis is invasive. Instead, the test is done on a newborn's blood for the phenotypic effect of the missing enzyme.

a. Since the liver enzyme, phenylalanine hydroxylase, catalyzes the conversion of what I will call X (it's actually phenylalanine) into Y (tyrosine), when the enzyme is missing, X piles up in the liver and spills out into the blood.

b. A drop of blood is taken from the heel of the infant on the first day of life and put onto a blotter card. That drop is then used to the test for X (and therefore, PKU). This test, invented by Robert Guthrie in 1962, is now widely used.

c. A number of other tests for genetic abnormalities can be done on the same drop of blotted blood.

III. Testing for alleles by DNA is possible.

A. Sequencing DNA would be the ultimate test for a genetic abnormality. But this is very costly at present (although the costs continue to fall).

B. For now, other rapid tests are used.

1. DNA hybridization probe is feasible. In Lecture Fifteen, we described the use of nucleic acid hybridization, in which a single-stranded DNA probe is used to see if it binds, or does not bind, to the tested target DNA. Consider the example of cystic fibrosis in the opening story. In this case, DNA on the blotter paper of a newborn would be hybridized with DNA probes that match the known mutation in the two parents. If the probe found its match, the newborn would be considered to have cystic fibrosis. It would be important to show that a probe to the normal allele did not bind in this case; if it did, the infant would be a carrier.

2. Polymerase chain reaction can be used to detect the presence of an allele. PCR begins with a short strand known as a primer (like what you paint on a wall). The primer must bind by base pairing to a known part of the DNA to be amplified. Now, consider the DNA on the blood blotter. To get PCR going, a primer that binds to the normal, intact allele would be used. This primer would not bind to the deleted, mutant allele. So the amplification of the CFTR gene would depend on whether it was intact or not.

IV. Screening programs are widespread.

A. Newborn screening is medically and economically beneficial.

1. Consider PKU: Testing was begun in 1965 and is now mandatory for all newborns in most countries and all U.S. states.

a. Babies who test positive and are put on a special diet end up mentally normal. Untreated, a person with PKU has an IQ of about 55 and is severely retarded. The human costs are huge.

b. The economic analysis is that the test costs about $2. The frequency of newborns with PKU is about 1 in 14,000. This means it costs $28,000 in tests for every infant with PKU detected. Add to this the costs of being PKU in terms of diet and medical care—about $3,000 per year—and for 30 years the total cost is $118,000. This is far lower than 30 years of caring for a mentally retarded person.

2. Other diseases are screened in newborns, such as some rare genetic disorders and genetic hypothyroidism. The latter used to be hidden as a cause of mental retardation. Of 5 million infants screened a year, 1500 have it: They are treated with thyroid hormone, and retardation is prevented.

3. Cystic fibrosis DNA testing has begun in some places, as has sickle cell DNA testing. (Many states and countries are working toward testing newborns for over 20 genetic disorders.)

B. Screening adults and children in target populations has been successful.

1. African Americans have set up screening for sickle cell disease, and this has led people to treatment.

2. Jewish groups have screened for Tay-Sachs disease (most common in Jews), and the test reveals carriers. Since so many have been screened, marriages between carriers are now rare, and prenatal testing is done on the fetuses. Studies show that the mother usually chooses to terminate the pregnancy when the test is positive. It is now very rare for a Tay-Sachs child to be born of Jewish parents.

C. Pharmacogenomics and disease susceptibility testing are the next frontier.

1. When any outside agent enters the body, it passes through the liver, where a series of proteins transforms it to more water-soluble forms for excretion in the kidney. This is evolutionarily advantageous, as it gets rid of these substances.

a. This system is under genetic control: There are alleles coding for proteins that are more or less active.

b. Consider smoking. The poisons in cigarette smoke are changed by the liver system. In most cases, the changes make them more water-soluble but also more harmful to cells, causing DNA damage that can lead to cancer.

c. If a person carries a mutation in the liver system that causes less activity on the molecules from smoke, there will be fewer poisons produced. We all know long-term smokers who do not get cancer; many have the right alleles.

d. The same goes for drugs. They are also modified by the liver. Having the right alleles may make one more or less susceptible to the actions of drugs. This is why drug companies sponsored the private human genome sequencing effort. They wanted to identify these alleles.

2. Many diseases, such as cancer and heart disease, have complex causes and many gene products interacting. So there are many disease susceptibility genes that are being identified. Tests will be developed for them.

Essential Reading:

Jack Pasternak, *An Introduction to Human Molecular Genetics*, 2nd ed. (Bethesda, MD: Fitzgerald Science Press, 2005).

David Sadava, Craig Heller, Gordon Orians, William Purves, and David Hillis, *Life: The Science of Biology*, 8th ed. (Sunderland, MA: Sinauer Associates; New York: W. H. Freeman and Co., 2008), chap. 17.

Supplemental Reading:

Thomas Devlin, *Textbook of Biochemistry with Clinical Correlations*, 6th ed. (Hoboken, NJ: Wiley-Liss, 2006).

Susan Lindee, *Moments of Truth in Genetic Medicine* (Baltimore: Johns Hopkins, 2005).

Philip Reilly, *Is it in your Genes? How Genes Influence Common Disorders and Diseases that Affect You and Your Family* (Woodbury, NY: Cold Spring Harbor Laboratory Press, 2004).

Questions to Consider:

1. Were you or your children tested for genetic diseases as newborns? If so, what is the difference between the current testing program in your area and the one that was in place when you were born?

2. Should employers or health insurers be informed of the results of a genetic test? What about a genetic test for disease susceptibility?

Lecture Nineteen
Molecular Medicine—The Immune System

Scope: The immune system identifies substances that are not part of an individual's genetic makeup and reacts to them by eliminating them. For example, this happens when a person is exposed to a virus or a blood transfusion from an unrelated donor with different genes. White blood cells first identify the substance, called an antigen, as foreign to the body and then initiate a two-pronged response. White blood cells called T cells identify and attack any body cells harboring the antigen; B cells make proteins called antibodies that bind to and eliminate the antigen if it is in the bloodstream. There is tremendous diversity in antigens, so there must be diversity in the T and B cells that attack them. This diversity is generated by an unusual recombination of genes in the cells so that a few hundred genes can be randomly selected and combined to make millions of "super-genes" that code for proteins. Vaccines are inactive antigens that provoke an immune response that stands ready to rapidly eliminate the true antigens when they infect a person. Antibodies can be used as very sensitive detectors, as in the pregnancy test, or to bind to cancer cells to get the body to attack them.

Outline

I. Opening story: George Washington and smallpox.

A. Smallpox has played a major role in history.

1. "Finding smallpox to be spreading much, and fearing that no precaution can prevent it from running through the whole of our army, I have determined that the troops shall be inoculated. Should the disease rage with its usual virulence, we should have more to dread from it than from the sword of the enemy."

2. So wrote the commander in chief of the U.S. Continental army to his chief physician on January 6, 1777. During the previous year, the disease had ravaged General Horatio Gates's American Northern Army. Of his 10,000 troops, over half got smallpox, and his military campaign had to be suspended for a month.

3. Commander George Washington knew whereof he wrote in that letter. When he was a teenager, he had a mild case of smallpox.

a. As a survivor of this often-lethal disease, Washington developed a healthy respect for it.

b. Washington had heard about inoculation. In Boston, Cotton Mather, the famous minister, watched three of his children nearly die of smallpox. Then, a slave from Africa, Onesimus, told him that in his homeland a person would be protected if a little dried pus from an infected person was put into a cut on the skin. When some sailors inadvertently brought smallpox to Boston and the next epidemic hit in 1721, Mather convinced a physician, Zabdiel Boylston, to try inoculating some normal, undiseased people. The epidemic was stemmed.

c. The 1777 inoculation ordered by Washington was the first known inoculation of an army, and it worked well. Casualties from smallpox were greatly reduced.

d. Smallpox claimed 300 million lives worldwide in the 20th century. As late as 1967, there were 15 million new cases and 2 million deaths due to the disease.

e. A massive effort at inoculation led to the eradication of smallpox from the world in 1979.

B. A series of events in the immune system is involved in this story.

1. Smallpox is caused by a virus that infects only humans. For 30% of people exposed, lethal infections such as pneumonia soon follow, and death is the inevitable result.

2. Fortunately, after the virus entered Washington's body, some white blood cells engulfed it, digested it into pieces, and presented some of the viral protein fragments on their cell surface. Other white blood cells recognized this "I've got a protein" flag, setting in motion an army of cells that destroyed other cells harboring the virus. Still other white blood cells made antibodies, proteins in the blood that would bind up any virus outside of Washington's tissues. This two-pronged attack—cells to kill infected tissues and antibodies to bind up free viruses—swiftly reduced the infection to a mild one.

3. After the infection subsided, most of these two armies of white blood cells died. But a small contingent remained, ready to do battle in case a new infection occurred.

4. These "memory cells" were behind the inoculation of the Continental army: When dead smallpox viruses entered the body through a cut in the skin, they looked enough like the real, live thing that they provoked the immune system to send in the two armies. Of course, there was no real infection to fight. Most of the cells went home (died!), but once again, memory cells stayed around, at the ready for a new infection.

II. The immune system involves cells and molecules.

A. There are two purposes of the immune system.

1. It recognizes nonself cells, viruses, or chemicals from self ones.

2. It destroys the nonself agents before they can do harm.

B. Three types of cells are major players; they are all white blood cells.

1. Phagocytes surround and engulf nonself cells or viruses. They present the products of digestion on their surface.

2. T cells recognize the nonself products in the cell surface. Some of them signal other T cells to attack other cells harboring the invader.

3. B cells respond to the T cell signal by making antibodies.

C. Three types of molecules play important roles in the immune system.

1. Antigens are the nonself substances that provoke an immune response. They are any small chemical grouping that is foreign—that is, an arrangement of atoms that the particular individual does not have. A virus may have many of these groupings and provoke many immune responses. Many groupings on the viral surface may look just like self chemicals and not provoke the immune response at all.

2. Antibodies are proteins made by B cells that bind to antigens. Since antigens are chemical groupings, they have a specific shape (lock). Antibodies fit into this shape exactly (key). Each B cell can make one specific type of antibody that fits an antigen.

3. T-cell receptors are the proteins on the T cell surface that bind to antigen fragments presented by the phagocytes. Like antibodies, these receptors are highly specific for each antigen fragment, and there are millions of different ones.

III. The genetic control of the immune response is unusual.

A. There are millions of different B cells, each making a specific antibody. There are millions of different T cells, each making a different T-cell receptor. These cells are constantly being made in small numbers. If they are not used (no antigen for them), they die within days. If they are used (antigen present), they divide many times and form the cellular army to perform their roles.

B. Genetically, one might expect there to be millions of genes, one coding for each antibody or T-cell receptor protein. This cannot be so: We humans only have a total of 24,000 protein-coding genes.

1. The solution is to combine alleles, each of which codes for a part of the antibody molecule (or T-cell receptor).

2. For one type of antibody, there are many alleles for four regions in one of its chains: 300, 10, 5, and 2 = 30,000 possibilities for different chains (multiply the combinations), and 300, 1, 5, and 2 for the other chain = 3000 possibilities. So when the two protein chains that make up the final antibody are considered, there are 30,000 × 3,000 = 90 million different types of antibodies! This from only 625 genes.

3. These genetic shuffles are going on all the time to make T and B cells to respond to any possible chemical grouping that is not self. We can make antibodies to substances we have never seen and have not even been invented yet.

IV. The immune system can be manipulated in medicine.

A. Vaccines make use of memory cells.

1. After an infection, the T and B cells made in response to it die off, except for about 1% of them, which stay around as memory cells.

2. This is why people who have had the flu won't get it twice that year: Their memory cells expand rapidly to fight off any new infection.

3. A vaccine is an inactive or dead antigen. It provokes an immune response that leaves memory cells. These cells then fight off any infection by the real agent carrying that antigen.

4. For viruses or bacteria that have a rapid rate of mutation (due to errors in DNA replication), the mutated infectious agent might be different in its antigen, and so the vaccinated person's T and B memory cells do not recognize it. This is why there is a new flu vaccine every year.

B. Antibodies can be used for detection and therapy.

1. Because antibodies are so specific, they will bind to their target antigen in a mixture, like a magnet. Rosalyn Yalow and Solomon Berson invented the immunoassay—a way to tag an antibody so that the antigen it binds to would be tagged. This is the way that substances in tiny amounts such as hormones can be measured.

a. Pregnancy is indicated when a developing embryo makes a hormone called human chorionic gonadotropin, which is released into the blood and urine. This is detected rapidly.

b. The immunoassay also led to testing newborns for hypothyroidism and to eliminating their symptoms of retardation by giving them the missing hormone.

2. Antibodies have been developed that bind to antigens that are expressed on cancer cells. These antibodies can be injected into a patient, and the binding initiates cancer cell death. The new drugs Herceptin (for breast cancer), Erbitux (for colon cancer), and Rituxan (for lymphoma) are in this class.

Essential Reading:

Jeremy M. Berg, John Tymoczko, and Lubert Stryer, *Biochemistry*, 6th ed. (New York: W. H. Freeman, 2006), chap. 33.

Lauran Sompayrac, *How the Immune System Works* (Malden, MA: Blackwell Science, 2002).

Supplemental Reading:

Richard Goldsby, Thomas Kindt, and Barbara Osborne, *Kuby Immunology*, 6th ed. (New York: W. H. Freeman, 2006).

Benjamin Lewin, *Genes VIII* (Upper Saddle River, NJ: Pearson Prentice-Hall, 2005), chap. 26.

Questions to Consider:

1. There is concern about the development of smallpox as a biological weapon. What are the challenges in making a smallpox vaccine?

2. HIV infects the T cells that orchestrate the immune response. How does this cause the immune deficiency in AIDS?

Lecture Twenty
Molecular Medicine—Cancer

Scope: Rapidly increasing knowledge of the details of cancer at the molecular level has led to the development of new drugs like Gleevec, which is used to treat a form of leukemia. Tumors form because of inappropriate cell division. Cells normally have finely tuned genetically inherited internal and external controls of cell division. In addition, cells stick together in tissues. In cancer, these controls break down, and tumor cells divide, spread to other tissues, and even recruit their own blood supply. In some cases, an activated oncogene that stimulates cell division is brought into cells by a virus, and in other cases a person inherits a mutation in a tumor suppressor gene that cripples this gene that normally blocks cell division. In most cases, cancer is caused by a series of mutations in these genes. While some cancer-causing agents such as those in cigarette smoke are known, most are not. Cancer is treated by surgery, radiation, and chemotherapy.

Outline

I. Opening story: Molecular medicine treats cancer.

A. A new treatment is targeted for a leukemia.

1. Alan noticed that he tired easily and had shortness of breath, so he decided to see his physician.
2. Tests showed that Alan had over 200,000 white blood cells per milliliter (one-fifth of a teaspoon): 40 times higher than normal. In the next few days, a hematologist took a sample from Alan's bone marrow and looked at the immature white blood cells. She saw an abnormality in the chromosomes called the Philadelphia chromosome, and the diagnosis was made: chronic myelogenous leukemia.
3. In this type of leukemia, the DNA in two different chromosomes in an immature white blood cell is cut and spliced. Parts of two genes, ordinarily on separate chromosomes, come to be right beside one another. The cells carrying it are stimulated to divide rapidly.
4. Alan, now under the care of an oncologist, was given drugs designed to kill any reproducing cells. One drug bound to DNA, another interfered with making the nucleotide building blocks that make up new DNA, and a third inhibited the mechanism that partitions chromosomes to new cells during cell division. These drugs had bad side effects, but Alan stayed the course for six months. His white blood cell count went down to 80,000 cells per milliliter and then stalled. This was still 16 times higher than normal.
5. Alan's oncologist now tried a new approach—a specific drug. During the 1990s, the molecular biology of this leukemia was described in detail. The new gene made by chromosome shuffling was sequenced and its protein product studied. Chemists went into the lab and designed a drug that would specifically bind to and inactivate the new gene product in the tumor cells.

a. At the University of Oregon, Dr. Brian Druker coordinated a clinical trial, in which patients like Alan who were stalling in conventional treatment were given the drug.
b. It worked: Alan's white blood cell count went down to 5,400. There were no white blood cells with the abnormal chromosomes in his bone marrow. He was cured.

B. The development of this drug, Gleevec, is the prototype of the molecular approach to cancer treatment.

C. The aim is to find out precisely what is going wrong in a tumor cell and design rational treatments on this basis.

II. Cancer develops in stages.

A. Cancer cells are different from normal cells.

1. Cancer cells lose control over cell division. Normally, cells have two strategies to control when and where they will divide.

a. Internal controls: Normally, an internal "clock" triggers events in the cell division cycle. Each event is under separate control, and this determines whether and how fast the cell divides. Most mature, specialized cells don't divide (like the white blood cells in the bloodstream). Immature cells divide rapidly (like the ones in bone marrow). What keeps cells from dividing is a set of proteins called tumor suppressors that act as "brakes." These act at the control points, for instance to block DNA from replicating. Cancer cells have defective tumor suppressors ("bad brakes").
b. External controls: Hormone-like proteins called growth factors stimulate cells to divide by acting like a "gas pedal" on the cell cycle control points. Genes whose protein products stimulate cell division are called oncogenes. Cancer cells often make their own growth factors (self-stimulate) or have changes that make them hypersensitive to even tiny amounts of growth factors. They turn on the "gas pedal."

2. Cancer cells can spread to other organs. This is the most feared aspect of cancer, called metastasis. It makes cancer hard to treat (you can't operate to remove tumors that are all over the place) and leads to multiple organ failure.

a. Normal cells have a "glue" that is adherent and specific. Cancer cells lose this adhesion.
b. Cancer cells can detach from a growing tumor, chew up adjacent cells until they reach the bloodstream or lymph system, enter these vessels, travel to a new organ, and then stop there, growing a "satellite tumor" or metastasis. They even recruit nearby blood vessels to make branches to the tumor, ensuring it oxygen and nutrients from the blood. This is called angiogenesis.

B. These events—inactivation of tumor suppressors, activation of oncogenes, metastasis, and angiogenesis—

occur in sequence. Cancer is a multistep disease.

III. Cancer is a genetic disease, but not usually inherited.

A. The major events in the development of cancer all involve the expression of genes: oncogenes, tumor suppressor genes, metastasis genes, and angiogenesis genes.

1. Viruses can bring active oncogenes into cells.

a. Tumor viruses were first discovered by Peyton Rous in 1910, when he showed that a muscle (meat) tumor in chicken could be passed from bird to bird in the hen house, and that this passage was virus induced.

b. Tumor viruses have active oncogenes (e.g., for growth factors), so when they get into cells they stimulate cell division.

c. Only about 10% of human cancer is caused by viruses. The most important ones are hepatitis B virus, which causes liver cancer, and a papillomavirus (warts), which causes cervical cancer. In both cases, antiviral vaccines are being used to prevent infection and cancer.

2. Inherited mutations in tumor suppressor genes can allow cells to divide inappropriately.

a. About 10% of cancers are inherited. These include 10% of breast cancer and colon cancer, as well as some childhood tumors.

b. Compared to sporadic (noninherited) cancers, inherited tumors occur earlier in life (e.g., in breast cancer, in the 20s and 30s instead of after their 50s) and in numerous places (e.g., in breast cancer, in multiple locations and/or both breasts, instead of one tumor in one breast).

c. Arthur Knudson proposed a "two-hit" hypothesis for inherited cancer. For tumor suppressor genes, both copies must be mutated to make inactive proteins so that cell division is no longer blocked. People with inherited cancer have one bad allele already through inheritance; they just need a mutation in a cell with the "good" allele to have no good tumor suppressor protein made.

B. Most cancers are sporadic and need many mutations in their cells to form. This means that carcinogens (which cause cancer) are usually mutagens (they damage DNA).

1. Spontaneous mutations can lead to cancer. Many people get cancer with no risk factors for carcinogen exposure. Recall that errors in DNA replication can lead to mutations.

2. There are some known carcinogens. These include cigarette smoke (which damages specific tumor suppressor genes in lung cells) and ultraviolet light from the sun (damages skin cell DNA in many genes).

3. Most carcinogens are natural substances in the diet, and in most cases we now know what they are.

IV. There are three ways to treat cancer.

A. Surgery is the most common treatment.

1. Removal of a localized tumor can be curative.

2. Much of cancer surgery has become conservative of surrounding tissue, as another treatment is used to get rid of any surrounding tumor in the area.

B. Radiation is used on localized regions that cannot be removed in surgery.

1. This might be in a diffuse area, or if a tumor is right against a vital organ.

2. Radiation damages DNA. The amounts used are very large and focused on the tumor. Lifetime exposure of a person to radiation from all sources is 0.12 gray [a physical unit]; during the course of radiation treatment, the tumor gets 50 gray in a month. This is 400 times the lifetime dose!

C. Chemotherapy is used when tumors have spread over the body.

1. Typical chemotherapy drugs kill all dividing cells, including the tumor. So there are side effects where there are dividing cells in normal tissues (bone marrow, intestines, skin, etc.).

2. There is a wide array of drugs that block cell division. Some are natural products, others are synthetic.

3. Precise molecular descriptions of the chemical biology of cancer are leading to targeted chemotherapies in molecular medicine.

Essential Reading:

Vincent DeVita Jr., Samuel Hellman, and Steven Rosenberg, *Cancer: Principles and Practice of Oncology* (Philadelphia: Lippincott Williams and Wilkins, 2005).

Lauren Sompayrac, *How Cancer Works* (Sudbury, MA: Jones and Bartlett, 2004).

Robert Weinberg, *The Biology of Cancer* (New York: Garland Science, 2007).

Supplemental Reading:

Ricki Lewis, *Human Genetics*, 7th ed. (New York: McGraw-Hill, 2006), chap. 18.

Gerald Litwack, *Human Biochemistry and Disease* (New York: Academic Press, 2007).

Questions to Consider:

1. A vaccine was developed against human papillomavirus, which causes genital warts and cervical cancer. Since this is a sexually transmitted disease, who should get the vaccine? Why is there some controversy about its widespread use?

2. What is the difference between the molecular medicine approach to cancer treatment and the conventional approaches now used?

Lecture Twenty-One
Molecular Medicine—Gene Therapy

Scope: Gene therapy is the addition to humans for medical benefit of protein-coding DNA along with a promoter sequence for its expression. There have been many successful pilot experiments on animals, including the production of "marathon mice" with altered muscle composition and improved endurance. In humans, gene therapy can be done by removing cells from the body, adding the new gene, and reintroducing the cells (ex vivo). Or the gene can be given directly to the patient (in vivo). Approaches include adding a new gene to correct the effects of a defective one and adding a gene that enhances cell killing by the immune system or drugs. Disabled viruses are used as vectors. There have been some modest initial successes at gene therapy for genetic diseases and for cancer.

Outline

I. Opening story: Gene therapy for athletic performance?

A. Gene therapy produces a marathon mouse.

1. Sprinters are different from long-distance runners, and a lot of the reason is, not surprisingly, in their muscles. The skeletal muscles that move bones in your arms and legs have two types of fibers.

a. Slow-twitch fibers need abundant oxygen from the blood to work well. They contain a lot of mitochondria, which are the energy-producing factories of cells and generate a long-lasting supply of energy for muscle contraction. They don't get tired.

b. Fast-twitch fibers, on the other hand, don't need as much oxygen. They have fewer mitochondria and generate quick bursts of energy but fatigue easily.

c. Not surprisingly, sprinters have more fast- than slow-twitch fibers, and in distance runners the situation is reversed.

d. Training improves the blood supply and even changes the ratio of fibers somewhat. But a lot of the ratio appears to be genetically determined. You don't hear of a person winning both the marathon and the 100-meter dash.

e. Distance runners also seem to burn fat effectively; they eat a lot and don't get fat.

2. Enter the marathon mouse. No, this isn't a video game or cartoon character. It's a mouse that has been treated with gene therapy to change the ratio of its slow- to fast-twitch fibers.

a. Dr. Ron Evans, who created this mouse in his lab, found that a protein called PPAR-delta (if you need to know, the "PPAR" stands for "peroxisome proliferator-activated receptor") got fat tissue to break down stored fat for energy.

b. Evans then wondered what would happen if mice were genetically modified to express a high level of PPAR-delta in muscles, which also need energy. He predicted that they would burn fat for energy. So he did gene therapy: He put the gene for PPAR-delta beside a promoter that would be expressed

in skeletal muscle. This recombinant DNA was put into a vector, and the vector injected into fertilized mouse eggs.

c. As expected, the genetically modified mice that grew up did not gain extra weight on a high-fat diet, as the PPAR-delta stimulated fat breakdown. What was unexpected was that the composition of the skeletal muscles changed from an even ratio of slow- to fast-twitch, to a marked increase in slow-twitch fibers. There were more mitochondria to burn the fat.

d. The performance of the mice on an exercise wheel reflected the muscle fiber changes. The genetically altered mice ran almost twice as long and twice as far on a wheel as normal mice.

B. This serves as a good example of gene therapy, the addition of genes to humans or animals for medical purposes.

II. There are several strategies for gene therapy.

A. Gene therapy is based on several assumptions.

1. We know the gene involved in a disorder.
2. We have a normal copy of that gene.
3. We know where and when the gene is normally expressed.
4. We are pretty sure what will happen when the normal gene is expressed appropriately.
5. Molecular medicine is giving us much of this knowledge.

B. Gene therapy must do the following:

1. Get the gene to the appropriate cells.
2. Get the gene expressed in the cells.
3. Get the gene integrated into the genome of the target cells.
4. Have no bad side effects.

C. The strategies for gene therapy are:

1. Gene augmentation is used for diseases in which a functional gene product is lost. The idea is to introduce extra copies of the normal allele so that the protein product is made and function is restored. An example might be muscular dystrophy, in which the muscles lack dystrophin, an organizing protein.

2. Targeted cell killing uses a gene that either produces a toxin that kills certain cells or stimulates the immune system to do so. This is useful in cancer, where it is called "suicide gene therapy."

3. Targeted mutation correction attempts to replace a bad allele with a good one. This is needed if the harmful allele makes a harmful gene product. This might be useful in cancer where a mutation in an oncogene has occurred in which the gene product, a growth factor, is always made. This stimulatory allele must be destroyed.

D. Method 3 above has been used successfully on animals but not humans. The human clinical experiments (trials) have used methods 1 and 2. No one has suggested germ line gene therapy on eggs or sperm. All of the potential uses are on nonsex (somatic) cells.

III. There are several methods for human gene therapy.

A. In ex vivo therapy, cells are removed from the patient, the new gene added in the lab, and the cells put back into the patient.

B. In in vivo therapy, the DNA is actually added directly into tissues, by injection or some other way.

C. There are several DNA vectors for human gene therapy.

1. These must be expression vectors with active promoters to express the gene of interest. They have the same requirements as vectors in recombinant DNA experiments on cells, plants, and animals: small size, a marker gene to show the vector got into the targeted tissue (harmless and easily detected gene product), and restriction enzyme sites for insertion of the DNA of interest. Most gene therapy vectors are viruses that are infective but have been genetically modified to not produce more viruses (they just get in and deliver their DNA).

2. Adenovirus: This is a large DNA virus that causes colds. The DNA remains outside of the cell's chromosomes, and so it will only affect the cell that gets the virus.
3. Adeno-associated virus: This small virus inserts its DNA into the host chromosome at a specific site (chromosome 19). It is not pathogenic (an advantage) but small in DNA carrying capacity (a disadvantage).
4. Retroviruses: These viruses use RNA as genetic material and reproduce in an unusual way. They incorporate their DNA into the host cells, but at any location, and so there is a chance they will get into the wrong place.

IV. Some gene therapy looks promising.

A. Very few successes have been reported in humans. The reasons are:
1. Therapy does not last long. Target cells often divide rapidly, and if the new gene does not get into most of them, the treated cells will be outnumbered.
2. Patients may mount an immune response to the vector. This has occurred unexpectedly, and all gene therapy was stopped for a time in 1999 when an 18-year-old boy died in this way.
3. Many diseases cannot be cured with a single gene therapy. Heart disease, diabetes, Alzheimer's, and arthritis are caused by defects in many genes interacting with the environment.

B. Here are some examples of gene therapy.
1. Severe combined immunodeficiency: A patient has neither cell-mediated nor humoral immunity due to a gene mutation for adenosine deaminase (ADA), which is normally expressed in white blood cells.
a. Prior treatment for this disease was to totally shield the patient from any germs in the environment (the "bubble boy") and give blood transfusions. Unfortunately, the bubble boy died of viral infection in a blood transfusion from his sister. Recall that there must be a genetic match in transfusions.
b. Then physicians tried giving the patients the missing protein, ADA, isolated from cows. This was moderately successful.
c. Finally, they tried ex vivo gene therapy, using gene augmentation to supplement the ADA injections: The patient's white blood cells were removed and the good ADA gene inserted into them via a retroviral vector. The cells were put back into the patient and made ADA, lowering the dose of external ADA necessary. But these blood cells were T cells, and they died after a few months. Now the doctors do this gene therapy on bone marrow cells that live for years. Still, the gene therapy is only a partial treatment.
2. Familial hypercholesterolemia: A patient cannot remove cholesterol from the blood because of a gene mutation for the receptor expressed in the liver.
a. Physicians tried gene therapy ex vivo, using gene augmentation to add the good gene for the receptor. A piece of liver was removed and the cells disaggregated in the lab. Then the gene was added by a vector and the cells grown to a larger number.
b. Upon reinfusion to the patient, the therapy cells fused with the liver, continued to divide, and made enough receptor to restore most function.
3. Ornithine transcarbamoylase deficiency: A patient cannot break down ammonia released by amino acids in the liver; the ammonia is very toxic. Most die within a year of birth, but milder cases can live for years by taking a special diet and ammonia-reducing drugs. Gene therapy was done in vivo by adenoviral vector with the good gene injected into an artery leading to the liver. The virus infected liver cells, and they expressed the enzyme, lowering ammonia. This was done in milder cases, but has not yet been done on the severe cases in infants.
4. Cancer is the subject of most gene therapy trials. There have been two approaches:
a. In vivo: gene augmentation by introduction of a virus expressing a tumor suppressor gene that is mutated in the tumor. This has been done in lung cancer for the tumor suppressor gene p53 that is mutated so that its protein is nonfunctional in the tumor. Generally, recipients have been very sick patients who failed other therapies. Scientists have shown that the vector gets into the tumor, is expressed, and the patients generally live longer.

b. Ex vivo: targeted cell killing by the removal of white blood T cells and addition of a T-cell receptor that binds to a tumor. The cells can home in on the tumor expressing the antigen to that receptor and stimulate an immune response to eliminate the tumor. This has been done for melanoma.

5. Muscle buildup is not just for athletes: In muscular dystrophy there is a lack of muscle repair and replacement by fat/fibrous tissue. Also, most older people have muscle wasting: Strength and mass decrease by up to one-third from ages 30 to 80. Finally, in some disabilities there is muscle wasting due to underuse: If we don't use it, the body does not repair it, and muscle cell death ensues.

a. At the University of Pennsylvania, Lee Sweeney is finding what controls muscle repair. He has found two genes involved and the proteins they code for: IGF-1, which stimulates nearby cells to promote muscle buildup; and myostatin, which stimulates muscle breakdown.

b. If the IGF-1 gene is added to mice, using an AAV vector for muscle expression, treated mice get bigger muscles.

c. If gene therapy adds a protein that inhibits myostatin, muscles get bigger as well.

d. Clinical trials are planned on patients with muscle disorders. There is concern that this may be abused by competitive athletes.

6. There have just been a handful of successes in gene therapy, which is proceeding with caution but promise.

Essential Reading:

Ricki Lewis, *Human Genetics*, 7th ed. (New York: McGraw-Hill, 2006), chap. 20.

Joseph Panno, *Gene Therapy: Treating Disease by Repairing Genes* (New York: Facts on File, 2005).

Supplemental Reading:

Susan Barnum, *Biotechnology: An Introduction* (Belmont, CA: Thomson Brooks-Cole, 2005), chap. 10.

Gavin Brooks, *Gene Therapy: Using DNA as a Drug* (London: Pharmaceutical Press, 2003).

Questions to Consider:

1. Do you think gene therapy to change the germ line, such as giving new genes to eggs, should be banned?

2. There have been many news stories of successful gene therapy in animal experiments but few in humans. Why?

Lecture Twenty-Two
Molecular Medicine—Cloning and Stem Cells

Scope: Cloning is the production of an organism genetically identical to another organism of that species or cell. Plants can be cloned easily from specialized cells. This shows that these cells have all the genes for all other cell types; that is, they are totipotent. Showing this is harder in complex animals. The nucleus of a specialized cell can be transplanted into an egg whose nucleus was removed, and the resulting egg can be induced to form a new organism, or clone. This was first done on frogs, then sheep (Dolly), and now many other animals. Reasons for animal reproductive cloning include preservation of valuable genotypes, preserving endangered species, and preservation of the genes of a pet. This technology makes human reproductive cloning feasible. Stem cells are constantly dividing cells that make a pool for specialized cell formation to replace cells that are lost or damaged. Stem cells can be pluripotent, such as those in bone marrow that form several kinds of blood cells, or totipotent, as in embryonic stem cells that can form all the kinds of cells in the body. There is potential in using specialized cells derived in the laboratory from stem cells to repair damaged tissues, ranging from heart to brain. Stem cell technology can be combined with cloning in the process of therapeutic cloning. This may make stem cells and tissues that will not be rejected by the recipient.

Outline

I. Opening story: stem cells from fat.

A. Many tissues have stem cells.

1. As a plastic surgeon in Los Angeles, Dr. Marc Hedrick's practice included liposuction, in which unwanted fat is removed from the body.

2. When he looked at some of the fat under the microscope, he saw not just fat cells, but some other types as well. He proposed that these specialized cells got there because fat tissue harbors some stem cells—unspecialized cells that constantly divide to form a pool of cells that then specialize when needed. He thought that the fat stem cells might form a population for cells whose origin in the embryo related to fat cells: muscle, bone, cartilage, blood vessel, and of course fat.

3. Hedrick and some colleagues set out to find these stem cells in discarded fat, and indeed found them. When implanted into animals, these cells will specialize into the tissue where they are located: So stem cells from fat can be put into damaged blood vessels and will take a hint from that environment and specialize into blood vessel cells, repairing the damage.

a. Fat stem cells are reaching the clinic. An advantage of using them is that a person's own fat can be used to get them so they will not be rejected as nonself when implanted into an organ. Hedrick has invented a way to get fat stem cells in about an hour in the operating room while a patient is there waiting for the implant. About 450 grams (a pound) of fat is enough to obtain 200 million stem cells, sufficient for therapy.

b. Recently, some women in Japan received fat stem cells to help repair breast tissue after mastectomy for cancer, and a child in Germany got them to help repair the skull after damage in an accident.

B. The biology behind this has come from studies on the molecular genetics of cells, focusing on a key biological question: How does an unspecialized cell become specialized? Put another way, how does a fertilized egg form the entire organism? The biological basis for the answer is that every somatic cell in the body has all of the genes necessary for the entire organism. That is, all cells are totipotent.

II. Cloning in plants confirmed totipotency.

A. In 1958, Frederick Steward, a plant biologist, showed directly that a specialized cell is totipotent.

1. Steward took specialized cells from a carrot (a certain cell type in the root that we eat) and put them into a chemical environment that mimicked the embryo. This provided signals for the root cells to first despecialize, then form a carrot embryo, and then form an entire carrot plant. The plant was a clone, genetically identical to the initial root cell.

2. This can be done for many different plant cells and organs. Specific chemical signals can turn one cell type into another.

III. Animals can be reproductively cloned.

A. Cloning animals cannot be done the same way as plants.

1. Some simple animals can regenerate organs from other cells (worms, hydra, starfish). This means that their specialized cells are easily manipulated and "plastic" in their genetic program.

2. For most complex animals, such as mammals, this kind of cloning cannot be done.

3. The proof of totipotency in animals came from studies not on the whole cell but the cell nucleus.

a. The idea was to surround a specialized cell's nucleus with the chemical environment of the fertilized egg. Scientists did it by replacing the egg nucleus with the specialized cell nucleus, thereby giving the latter an "egg" environment, with its proper signals to go on and become an embryo.

b. The first experiments were done on frogs: Their egg cells are large and easy to get from the water. The egg nucleus was removed and replaced with another nucleus, and then the egg was coaxed to try to form a tadpole and then a new frog. Donor nuclei from the embryo always worked; later stage embryo and tadpole nuclei sometimes worked; adult cell nuclei worked—but rarely.

c. The frog experiments showed that there is totipotency (all of the genes) in a specialized cell nucleus. But something prevents consistent cloning as the cell gets more specialized.

d. For 30 years, scientists tried to figure out what made an animal specialized cell nucleus largely unsuitable for cloning. In 1996, Ian Wilmut and colleagues found out. By starving the donor cells in the laboratory, they found that their nuclei would be much better for transplant into an enucleated egg and then cloning: Dolly the sheep was the result.

e. Dolly was cloned not just to give evidence for totipotency. The company Wilmut worked for was trying to produce transgenic animals that would make a human protein in their milk that treats cystic fibrosis. Cloning was the best way to propagate such an animal. This has now been done for human growth hormone made by cows, for example, as described in Lecture Twelve.

4. There are several reasons for reproductive cloning in animals.

a. Propagation of valuable animals (see above).

b. Preservation of endangered species that won't breed in zoos.

c. Preservation of a pet; the company doing this has done cats and dogs but is no longer in business because of lack of demand.

B. Human reproductive cloning is possible.

1. Human clones are born all the time: identical twins. The cloning of animals makes deliberate human reproductive cloning possible.

2. There are some reasons why human cloning might become available.

a. Many people have problems with normal reproductive mechanisms but want a child genetically related to them.

b. There may be a desire to perpetuate valuable genotypes. What if there was an old woman who never got cancer because she has certain mutations? These will die out with her, unless she is cloned. A person with unique characteristics (Einstein) might be cloned to preserve the genes that made him unique (whatever they are).

c. Perpetuation of a dying child.

3. But there are several concerns.

a. The process is not efficient. Dolly was 1 of 277 nuclear transplants; the others failed. The technology has improved, but it is not as good as test-tube baby technology.

b. In some species, clones have some problems. Defects in the immune system and disease susceptibility have been noted.

IV. Stem cell technologies are developing rapidly.

A. There is a need for new cells in medicine, for example:

1. In a heart attack, there is often permanent damage to the heart muscle.

2. In the brain, Parkinson's and other diseases result from a lack of functional cells.

3. In diabetes, the pancreas is damaged and its product (insulin) must be replaced.

4. In the musculoskeletal system, breaks and tears are often hard to repair.

5. Two problems in getting new cells are:

a. A source for these specialized cells. Organ transplants are hard to get.

b. The immune system ultimately rejects transplants as nonself.

B. Stem cell transplants are already performed.

1. Bone marrow gets damaged in cancer radiation therapy and chemotherapy as the dividing cells are killed. This makes the patient severely anemic and immunocompromised.

a. If the patient's bone marrow is removed before therapy and stored, it has enough pluripotent stem cells to form the new blood cells when returned to the patient after therapy.

b. If there is a blood cancer, the patient's bone marrow cannot be used. So a genetically related donor is sought.

c. There have been instances of parents conceiving a child to provide genetically appropriate bone marrow for transplant.

2. Bone marrow also contains stem cells that do not form blood but form the tissues around blood: blood vessels, muscle, and bone. These have been used as well and are pluripotent.

3. Pluripotent fat stem cells are just starting to be used (see opening story).

C. Embryonic stem cells are totipotent.

1. In the early embryo, cells are totipotent. The preimplantation genetic diagnosis described in Lecture

Seventeen shows this. One cell could be removed from an eight-celled embryo, and the rest of the cells took over.

2. At about the 10-day stage in a human embryo, there are several dozen of these undifferentiated, totipotent cells. These cells can be removed from the embryo. In 1998, James Thomson showed that they could be put into a lab dish and grown indefinitely.

a. In laboratory experiments, these stem cells can be induced to form many different cell types. In animals, these cells have cured brain, heart, muscle, and pancreas damage. This had led to great excitement about potential for humans.

b. It is proposed to use lab-grown stem cells as a supply. Not a lot of embryos would be needed. But these cells might be rejected, as they are nonself.

3. A solution to this rejection problem combined cloning and stem cells: therapeutic cloning.

4. Here is the scenario: I have a damaged heart because of a heart attack. Some of my skin cells are removed and grown for a few days in the lab, then sent to a cloning lab. There, a woman donates an egg cell after being induced to ovulate. Her egg cell nucleus is removed and replaced by mine. The egg is stimulated to divide and forms an early embryo. After 10 days, the embryonic stem cells (genetically mine) are removed, placed in a lab dish, and induced to form heart cells. These heart cells are then sent to me and implanted in my heart, where they repair it.

5. Another method produces embryonic stem cells without the embryo. Shinya Yamanaka at Kyoto University has found that mouse skin cells can be reprogrammed to act like embryonic stem cells if four genes are added by vectors for high expression. All four genes code for transcription factors. Their gene products cause cells to divide, prevent cell death, and cause them to be totipotent embryonic stem cells. Indeed, in mice these cells can be coaxed to form true embryos that develop into newborn mice. So a possible scenario might be for skin cells to be reprogrammed to make stem cells, which would then be reprogrammed to make heart cells in the example above.

6. In any case, like fat stem cells, this is a way to get around concerns about using cells from embryos: These cells never get to be part of an embryo.

Essential Reading:

Susan Barnum, *Biotechnology: An Introduction* (Belmont, CA: Thomson Brooks-Cole, 2005), chaps. 7 and 10.

David Sadava, Craig Heller, Gordon Orians, William Purves, and David Hillis, *Life: The Science of Biology*, 8th ed. (Sunderland, MA: Sinauer Associates; New York: W. H. Freeman and Co., 2008), chap. 19.

Supplemental Reading:

Michael Bellomo, *The Stem Cell Divide* (New York: AOMCOM Press, 2005).

Jay Gralla and Preston Gralla, *The Complete Idiot's Guide to Understanding Cloning* (New York: Penguin, 2004).

Christopher Thomas Scott, *Stem Cell Now: From the Experiment That Shook the World to the New Politics of Life* (New York: Penguin, 2005).

Questions to Consider:

1. Do you think that there are suitable alternatives to embryonic stem cells?

2. What are the moral and ethical arguments against reproductive cloning in humans?

Lecture Twenty-Three
Genetics and Agriculture

Scope: Agriculture is the earliest example of biotechnology. The development of agriculture was an important event in human history, as it allowed for a settled lifestyle. The challenge of agriculture is to feed a growing human population. Three crops—rice, wheat, and corn—directly provide about two-thirds of the human diet. While previously, people increased crop production by expanding the land under cultivation, now the best land is already taken. So the yield or production of crops on a given piece of land must be improved. This is done by intensively managing the soil, water, and pest ecosystems, as well as by altering the genetic capacities of the crops. The methods of crop plant genetics use the principles of Mendelian genetics and evolution by

natural selection. These methods include pure line selection, hybridization, and deliberate crosses. They have been very successful but have some limitations.

Outline

I. Opening story: the green revolution in genetics.

A. The genetics of wheat breeding culminated in Japan.

1. An article in a newspaper in Connecticut in 1794 described a new genetic variety of wheat that grew quickly, was resistant to a major mold disease, produced more grain, and was 25% shorter and stronger than other varieties.

2. For reasons unknown, farmers did not use this wheat when the U.S. Midwest was settled in the decades to come. But the Japanese developed a semi-dwarf variety, which obviously had the same height genes as the U.S. variety. By the 20th century, their semi-dwarf variety gave high yields.

3. Height and grain yield are complex phenotypes determined by numerous genes.

4. After World War II, the U.S. sent an occupation army into Japan, and among the first officers was an agricultural attaché. He was sent to find out how the Japanese had produced enough food, and once he went into the countryside and saw the short, high-yielding crops, he had his answer.

5. The Americans sent some of the unusual wheat seeds to colleagues in the Pacific Northwest, where they were crossed with local varieties and record yields were reported.

B. Wheat breeding in Mexico benefited the poor.

1. In 1944, a wheat research program had been set up in Mexico under the sponsorship of the Rockefeller Foundation. It was a joint U.S.-Mexico venture, with the goal of improving wheat production in Mexico as an aid to economic development. A young scientist from Iowa, Norman Borlaug, headed the plant-breeding effort.

2. In 1953, Borlaug received some of the semi-dwarf wheat. He set out to introduce genes that would adapt the wheat plants to the climates of Mexico in particular and the poor regions of the world in general.

a. Borlaug made his crosses and planted at two locations: One is on a cool, wet plateau near Mexico City; the other is in the hotter and drier state of Sonora.

b. Having two sites was deliberate: Borlaug wanted two crops a year; this would speed up experimentation. And the two sites with different climates would provide a wide adaptation to climates around the world.

3. By 1961, the crosses resulted in new wheat strains that were fast maturing so they could grow two crops a year. The crops were naturally resistant to a wide variety of pests, very high yielding, semi-dwarf, and adaptable to warm and cool climates.

a. In Mexico, wheat yields took off and had tripled within the next 12 years. Mexico no longer had to import wheat.

b. In 1964, Borlaug visited India, and upon returning to Mexico sent 100 kg (60 lb.) of seeds of the new wheat to his Indian colleagues. They were successful there and in Pakistan. Wheat yields skyrocketed. In 1968, an aid official described the astonishing events as a “Green Revolution.”

4. In 1970, Borlaug was awarded the Nobel Peace Prize. There is no Nobel Prize for agriculture, but it was appropriate that he be honored, because his work prevented massive starvation and the political instability that would have followed. By human impact, it was the most important genetics experiment of the 20th century.

II. Plants are humanity’s major source of food.

A. The race between population and food production is a human challenge.

1. Agriculture is the oldest example of biotechnology. It began about 8000 years ago when hunter-gatherers found plants that they could eat, possibly growing near their garbage dumps. They ate the seeds and sowed some of them nearby for the next year. This allowed settlement.

2. In 1999, the world population reached 6 billion, adding 75 million a year. The UN estimates that one person in seven is underfed. It could be much worse: The 20th century saw great increases in food

production.

3. Demography is the study of human populations. Population growth equals birth rate (additions) minus death rate (subtractions).

a. When agriculture began, there was increase in the birth rate. But the death rate was high because of infectious diseases, so overall population growth was low.

b. When education and knowledge improved, so did medicine and sanitation. The death rate was reduced and the growth rate went up: There was a population explosion.

c. Then the birth rate went down and overall growth stabilized.

d. The U.S. and Europe have gone through this "demographic transition"; less developed areas are at stages b and c above.

4. Current projections are that world population will level off at about 10 billion in about 2050. This places pressure on food production.

B. Crop plants are the major source of food.

1. Food is any substance that provides energy (fuels conscious and automatic actions) and nutrients (substances we cannot make by our genetic limitations).

2. Worldwide, direct consumption of plant materials accounts for 75% of the food humans eat. The rest is caught and farmed fish, and farmed animal products.

3. Three plants provide two-thirds of the human diet: rice, wheat, and corn. These plants are called staple foods, as they grow in certain areas and the cultures have used them as their primary food.

4. We are dependent on the genetic capacities of these plants for growth and seed production.

a. Growth: The plants have specific environmental requirements. Modifying the environment to fit these requirements is called farming!

b. Seeds: The seeds are "lunch boxes" for the plant embryo and contain stored carbohydrates, proteins, and vitamins. Unfortunately, the storage proteins have low contents of two of the eight amino acids that we humans cannot make. And the contents of certain vitamins of the grain are not adequate.

c. Both a and b above are under genetic control, and a major effort in crop plant breeding has been made to improve them for our needs.

III. Agriculture maximizes the genetic potential of crop plants.

A. Expanding the land under cultivation drove human history.

1. This was a reason for the rise of empires.

2. But the "easy" land (U.S. Midwest, etc.) is gone, and lands remaining are very dry ones or tropical forests.

3. Sometimes land is not farmed, but kept bare to keep prices up.

B. Increased crop yields is now the way to increase food production.

1. Example: Japan in early 20th century faced a rising population but not much more land available for rice crops. To be independent of outside sources, they made a big push to increase yields on the land they had, and yields increased threefold. This was done by ecological management of the land and genetic crossing to improve the potential of the plants.

2. The management of the crop environment takes several forms.

a. Soil: Plowing fields and conserving the soil from runoff are important. Improving the nutrients (for plants) in the soil is essential. This is why soil is fertilized.

b. Water: This is needed to dissolve soil nutrients and cool the plant. The problem has always been that there is too much or too little water at the wrong times. So huge dams and canals make water available and deep wells bring water to fields.

c. Pests: Crops are eaten by insects, molds, bacteria, viruses, animals, etc. These can be controlled by hand (e.g., weeding) or pesticides (e.g., fungicides). Pesticides must be used with care, lest they damage beneficial organisms or even humans.

IV. Plant breeding uses genetics to improve crop plants.

A. There are three methods of conventional plant breeding that seek to improve plants genetically so they will grow better and give better food.

B. Pure-line selection: This was the way that crops were domesticated.

1. Example: Wheat relatives that grow in the wild (the original wheat) have alleles for seed germination (sprouting) that provide for dormancy. Seeds shed by the plant will germinate at staggered intervals. Dormancy is advantageous to plants, but not to farmers. So when farmers were the selective agent, they planted seeds and harvested only the ones that grew right away. These plants were used for the next generation. After about 20 generations of selection for this multigene-determined phenotype, the plants were homozygous for lack of dormancy.

2. There are many such varieties selected for one genetic characteristic. Seed banks have tens of thousands of varieties stored so farmers can use them where and when they are needed.

3. Some varieties have been selected for nutritional characteristics.

a. For example, in the 1960s a group at Purdue University found a corn variety that made a seed protein that was higher than normal in lysine. Usually, corn proteins do not contain much of this amino acid, and so people who eat it are malnourished.

b. Lysine is one of the eight amino acids that humans cannot make because we lack the gene for the enzyme involved. So we get our lysine supply in our diet. People who eat mostly corn protein do not get enough lysine for their own protein synthesis and are malnourished.

c. The Purdue scientists found a high-lysine variety of corn. It was bred (see C, below) with other varieties by a team led by Surinder Vasal to give quality-protein maize, a strain that has higher yields, more protein, and better balanced protein than traditional varieties.

C. Hybridization: Crosses between pure lines give more than the individuals.

1. Example: In 1910, George Shull crossed two pure lines of corn. Line number one had a yield of 20 bushels per acre, and line 2, with different ecological characteristics, also had 20. The offspring had 80.

2. Hybrid corn is heterozygous for many genes.

3. Wheat and rice do not produce good hybrids; it is not clear why.

4. All corn in the U.S. is now hybrid. Farmers must buy the hybrid seeds from seed companies.

D. Deliberate crosses: The principles of Mendelian genetics and its successors are used to cross plants with desirable characteristics and get single ones into a recipient plant.

1. Example: Borlaug wanted to add a gene for resistance to the wheat rust mold into the high-yielding, semi-dwarf plants.

2. He took a variety that was genetically resistant but had none of the other desirable characteristics, and crossed it with the semi-dwarf variety that had all the desirable characteristics except resistance.

3. The offspring were a mix. He chose those that had all the good semi-dwarf characteristics and also were moderately resistant and crossed them to the very resistant plants.

4. After six generations he had "fixed" resistance alleles in the semi-dwarf plants.

E. These methods are widely used in a huge, worldwide effort to improve plants. But:

1. They are using many genes in selection and crosses, and other, hidden genes that are not desirable may be also transferred.

2. There are many genes in nature that cannot be crossed into plants because they are in different species. For example, soybeans make a more balanced protein than corn, but these genes could never mix naturally because beans do not mate with corn.

3. It is slow. Many generations of selection or crosses are needed to get new characteristics in crop plants.

4. The ecological thrust of agriculture remains to use genetics and technology to adapt the environment to the plant.

Essential Reading:

Maarten Chrispeels and David Sadava, *Plants, Genes and Crop Biotechnology* (Sudbury, MA: Jones and Bartlett, 2003),

chaps. 13 and 14.

Nina Federoff and Nancy Brown, *Mendel in the Kitchen* (Washington, DC: National Academies Press, 2004).

Supplemental Reading:

Anthony Griffiths, Susan Wessler, Richard Lewontin, William Gelbart, and David Suzuki, *An Introduction to Genetic Analysis*, 8th ed. (New York: W. H. Freeman and Co., 2007).

Benjamin Pierce, *Genetics: A Conceptual Approach* (New York: W. H. Freeman, 2005), chap. 22.

Questions to Consider:

1. Seed dormancy, a useful characteristic of plants, has been lost from crops through genetic selection. What other characteristics do you think have been changed by domestication of crops?

2. Organic foods are defined as crops that are grown in a certain way, without pesticides or chemical fertilizers. What are the benefits of the organic process for the farmer? What are the differences between organic produce and nonorganic produce? Why don't all farmers use organic practices?

Lecture Twenty-Four
Biotechnology and Agriculture

Scope: Plant biotechnology seeks to genetically modify plants for human use in agriculture. Totipotency of plant somatic cells is important because it allows any cell to be manipulated in the lab and then grown quickly to produce a plant. Specific plant expression vectors can introduce new genes into plant cells. Gene guns can shoot DNA-coated pellets into the cells. Genetic manipulation can overcome some of the drawbacks of traditional plant genetics: It transfers only single genes, any gene from any organism can be transferred, it is fast, and it can result in plants tailored to their environment. Several examples of genetically modified plants are in use, ranging from plants that make an insecticide, to herbicide-resistant plants, to rice with improved composition for human nutrition, to salt-tolerant crops. While there is great potential for this technology, some people are concerned. These concerns relate to a human aversion to manipulating nature, the safety of foods from genetically altered plants, and the danger of ecological accidents.

Outline

I. Opening story: the salt-tolerant tomato.

A. Most plants cannot grow in salty soils.

1. Normally, the small amounts of salts that are dissolved in soil water get removed from it by rainfall washing it down to lower levels in the ground. But in dry climates there isn't much rain, and as time goes on, salt builds up.

2. Salt is toxic to plants in two ways. First, it impairs the roots from taking up water, and second, it inhibits some of the enzymes involved in making proteins and in photosynthesis—the process by which the plant converts solar energy into stored energy in sugars.

a. Few plants can thrive in very salty soils, and certainly the major crops cannot.

b. Salt buildup is a global problem. An estimated 1% of farmland a year (65,000 acres a day!) is made unusable because of saltiness.

B. Biotechnology can make plants salt tolerant.

1. In Lecture Nine, we described the tiny mustard-like plant *Arabidopsis* and how it is a model for the genomes of major crops. In the 1990s, Eduardo Blumwald found that this plant has a gene that is expressed as a protein that takes salt from the soil and puts it in cell storage depots, called vacuoles, in leaves. This hides the harmful salt from doing damage to the plant.

a. The problem comes when the salt buildup in the soil is high. It simply takes too much energy for the plant to keep pumping the salt into the vacuoles, so the excess salt gets into the rest of the cells, and the plants wither and die.

b. Using genetic engineering, Blumwald added a very active promoter beside the salt pump gene so that its expression would be enhanced. Sure enough, the genetically modified *Arabidopsis* was able to not just withstand salty soil, but to thrive in it.

2. As we saw in Lectures Ten through Fourteen, biotechnology allows scientists to transfer DNA from one organism to another. When the active salt tolerance gene was put into a tomato plant, the transgenic plant became salt tolerant. What is more, the salt was in the leaves; the tomato fruits were fine.

3. While they are important, tomatoes are not nearly as important as the major grain crops. So Blumwald and others are busily transferring the salt-tolerance gene from *Arabidopsis* to these plants. This may make salty soils usable for farming. Salinity ruined the soils where farming began, in the Fertile Crescent of the Near East. Salt-tolerant transgenic plants may make this desert bloom again.

4. There are several lessons in this story.

a. Biotechnology is a powerful, specific, and rapid way to transfer genes from one organism to another.

b. Biotechnology can lead to a fundamental change in the relationship between people, their crop plants, and the environment. Until now, we have made enormous efforts to adapt the environment to the plant. Now, we may be able to adapt the plant to the environment.

II. The methods of agricultural biotechnology are similar to those for other organisms.

A. Totipotency and recombinant DNA are essential to plant biotechnology.

1. As described in Lecture Twenty-Two, plant cells are totipotent and can be cloned to make new plants.

a. Cloning is valuable when uniformity is necessary. For example, in forestry, trees are cloned and plantlets put in the soil so they grow to the same size for harvesting.

b. Single plant cells can be transformed, selected, and cloned. This allows rapid screening of the phenotype.

2. There are two ways to get plant DNA into cells for transformation.

a. Vector: A bacterial infection causes crown gall tumors in plants. The bacterium has a small chromosome that is injected into the plant cells. This chromosome DNA was isolated and its genes for plant cell alteration (but not infection) removed. It has single restriction enzyme sites for gene splicing. In addition, there are numerous plant promoters that have been described in terms of organ and time (e.g., seed formation).

b. Gene gun: An inert pellet can be coated with DNA and literally shot into the plant cell nucleus. This is needed at times because plant cells are surrounded by a thick cell wall.

B. Plant biotechnology overcomes some of the limitations of traditional plant genetics that were listed in Lecture Twenty-Three.

1. When traditional plant genetics uses many genes in selections and crosses, other, hidden genes that are not desirable may be transferred also. In biotechnology, only single genes are transferred.

2. Many genes in nature cannot be crossed into plants by traditional genetics because they are of different species. In biotechnology, genes from any organism can be transferred.

3. Traditional plant genetics methods are slow. Biotechnology is rapid; results appear in weeks.

4. In traditional genetics, the ecological thrust of agriculture is to use genetics and technology to adapt the environment to the plant. In biotechnology, the plant can be genetically adapted to the environment.

III. Genetically modified plants are in widespread use.

A. Plants can make their own insecticides.

1. Many insecticides are not specific. They target many insects, not just the pest. And some are toxic to the environment in other ways.

2. Insect larvae (the grub or worm immature life stage) eat bacteria. *Bacillus thuringiensis* (Bt) bacteria solve their insect problem by making a protein that binds to the insect larva (grub, worm) intestine to make it lose fluid. The insect dies.

3. The gene coding for this protein has now been introduced to corn, cotton, soybeans, and tomato cells, which were cloned to plants that expressed the toxin in leaves. The larvae eat the leaves and die, and their population goes down. This reduced insecticide use by 90%.

B. Plants can be made resistant to herbicides.

1. Weeds are killed by herbicides, but a problem is that these chemicals also kill beneficial plants, and even the crops themselves. So, great care is needed in their use. Genes have been identified from bacteria and other sources that code for proteins that break down herbicides or make protein targets that are functional but unaffected by the herbicides. Such genes have been inserted into cotton, corn, soybeans, and rice.

2. The transgenic crops are now resistant to the herbicide, and it can be applied without risk of damaging the crop.

C. Golden rice has improved nutritional characteristics.

1. People require in their diet beta-carotene, which gets converted to vitamin A. Rice plants do not have the genes to make beta-carotene. As a result, about 250,000 children go partially blind each year because they are rice eaters and do not get enough beta carotene.

2. Other organisms have genes coding for enzymes that can complete the biochemical pathway in rice for beta-carotene. Ingo Potrykus isolated DNA for one of these enzymes from a bacterium and two other genes from a daffodil plant. One by one, these genes were introduced to rice plants along with a promoter that would stimulate gene expression in the developing rice grain.

3. The resulting rice plant made grains with beta-carotene that were golden and contain adequate amounts of beta-carotene, thanks to the transgenic biochemical pathway.

4. These plants are being crossed with local varieties to make the beta-carotene phenotype part of rice that is used in different regions.

IV. There is public concern about plant biotechnology.

A. At the start in the 1970s, there was concern about biotechnology in general. When it was shown to be safe, these concerns abated. Plant biotechnology and genetically modified foods (they are not necessarily the same—most cotton plants are now genetically engineered) have raised serious concerns, especially in Europe.

B. The objections to biotechnology in agriculture fall into three categories.

1. Genetic manipulation is an unnatural interference with nature.

a. This is what a philosopher calls the "yuck factor." Eating food from a plant that has genes from bacteria is just going too far.

b. There is no real scientific response to this emotional argument.

c. All major crops have been genetically manipulated, but not to the extent of biotechnology.

2. Genetically modified foods are unsafe to eat.

a. Some modifications may create allergic proteins.

b. Most genetically modified plants are not altered phenotypically in the food part of the plant, except for some extra DNA sequences.

3. Genetically modified plants are dangerous to the environment. Although a single gene has been transferred to the crop plant, it may be inadvertently transferred to neighboring plants, such as weeds, in the field.

V. Overview.

A. In this course, we have come from describing heredity to understanding its mechanism to controlling it.

1. In the first lectures, I described what genetics is and how a monk, Gregor Mendel, and others described the rules of inheritance. Then came the identification of what a gene is (DNA) and how it works (expressed as protein).

2. With these basics of descriptive genetics in hand, I described how this knowledge has been applied to human use in the biotechnology industry. From making products to solving crimes and cleaning up the environment, biotechnology is growing in importance.

3. Modern genetics is rewriting the "book of life" as explained by Charles Darwin's theory of evolution by natural selection. We are finding out more about how the amazing array of organisms on Earth are

related, and even how we as humans may have evolved.

4. Genetics and DNA are having growing impact on medicine. With precise descriptions of genetic causes of diseases in hand, we are now able to diagnose and screen for people with the diseases, and we are beginning to design specific treatments. The areas of gene therapy, cloning, and stem cells are at the leading edge of modern medicine and hold great promise for the future.

5. I ended the course with agriculture, that other application of biology to human welfare, not just because it is important (we have to eat!) but also because so many of us in the rich world take it for granted. The applications of modern genetics and DNA to the problem of feeding the world are already profound, and an exciting future is in store.

B. The genetic genie is out of the bottle.

Essential Reading:

Maarten Chrispeels and David Sadava, *Plants, Genes and Crop Biotechnology* (Sudbury, MA: Jones and Bartlett, 2003), chaps. 6, 17, and 18.

Nina Federoff and Nancy Brown, *Mendel in the Kitchen* (Washington, DC: National Academies Press, 2004).

Supplemental Reading:

Miguel Altieri, *Genetic Engineering in Agriculture: The Myths, Environmental Risks and Alternatives* (Oakland, CA: Food First, 2005).

Jon Entine, *Let them Eat Precaution: How Politics is Undermining a Genetic Revolution in Agriculture* (Washington, DC: AEI Press, 2006).

Questions to Consider:

1. Besides the applications described in this lecture, can you suggest other genetic characteristics that you would like to see in crop plants?

2. Much of the concern about plant biotechnology boils down to the "precautionary principle." This states that if any action *might* cause harm, don't do it. The onus is on the proponents to show that it *does not* cause harm. Predictions are not enough. Can you think of other technologies where we as a society use the precautionary principle? Do you feel it is justified for plant biotechnology?

Timeline

B.C.

2500.....Oral records of deliberate breeding for desirable characteristics of the date palm and the horse.

350.....Aristotle, the Greek philosopher, proposes that the genetic material is carried in sperm and that the female menstrual fluid "organizes" this to form offspring.

A.D.

200.....Report of the inheritance of hemophilia in the Babylonian Talmud, a biblical commentary.

1721.....Zabdiel Boylston uses inoculation to prevent smallpox in Boston (Edward Jenner provides experimental evidence for the effectiveness of inoculation in 1796).

1766.....The Dutch microscopist, Antonie Van Leeuwenhoek, observes human sperm under the microscope and "sees" tiny humans, as predicted by Aristotle.

1831.....The naval survey ship *Beagle* leaves England for a round-the-world, five-year expedition with Charles Darwin aboard as naturalist. Darwin's careful observations lead him to propose a mechanism for

evolution.

1852–1854.....Gregor Mendel attends the University of Vienna, where he studies mathematics, chemistry, and biology.

1858.....Publication of *On the Origin of Species* by Charles Darwin, explaining how natural selection of variations passed on genetically to offspring leads to the evolution of organisms through time.

1866.....Gregor Mendel reports on his experiments on garden peas, showing the particulate nature of the genetic determinants.

1868.....Friedrich Miescher isolates DNA.

1873.....The term “intelligent design” is used to describe the origin of complex living systems.

1895.....Albrecht Kossel finds that DNA is a long polymer of nucleotides A, T, G, and C.

1900.....Botanists Hugo DeVries, Carol Correns, and Erich von Tschermak independently verify Mendel’s conclusions about genes and alleles—and then discover his paper that had been published 34 years previously.

1903.....Chromosomes are identified in dividing cells as the probable carriers of genes.

1908.....George Shull crosses two pure lines of corn, producing plants whose seeds are very high yielding, demonstrating hybrid vigor (heterosis).

1910.....Peyton Rous discovers that the cause of a type of cancer in chickens is transmitted from one animal to another (he is awarded the Nobel Prize in 1966). Archibald Garrod describes the disease alkaptonuria as an inherited mutation that is expressed as a defective enzyme; the one gene–one enzyme idea is confirmed 40 years later.

1918.....Tsar Nicholas II of Russia and most of his family are killed and buried (DNA forensics is used to identify the remains in 1992).

1928.....Frederick Griffith discovers genetic transformation in bacteria, showing that nonliving extracts of one type can turn another type into the first one; this indicates a chemical nature for genes.

1934.....Asbjorn Folling describes phenylketonuria as a genetic disease.

1935.....Semi-dwarf, high-yielding wheat is developed in Japan by genetic selection and crosses.

1944.....Oswald Avery shows that DNA causes genetic transformation in bacteria, pointing to DNA as the genetic material.

1948.....The protein in hemoglobin is identified as the primary phenotype in sickle cell disease, pointing to protein as the expression of a gene.

1950.....Ernst Wynder finds a linkage between smoking and lung cancer in careful population studies, confirming data that were first reported almost 300 years before.

1952.....Alfred Hershey and Martha Chase show that DNA is the genetic material of a virus.

1953.....James Watson and Francis Crick propose the double-helix structure of DNA.

1954.....Sickle cell disease is described as a balanced polymorphism, existing in high frequency in some populations because it confers resistance to malaria.

1956.....Arthur Kornberg describes DNA polymerase, the enzyme that catalyzes DNA replication.

1957.....Semiconservative replication of DNA is demonstrated in bacteria, and later in eukaryotes.

1958.....A carrot is cloned from a single specialized cell, thus showing that each specialized cell is totipotent.

1959.....The roles of messenger RNA and transfer RNA in gene expression are described.

1960.....The immunoassay is invented to test for tiny amounts of substances; it soon has wide applications, especially to detect hormones. An unusual chromosome called the Philadelphia chromosome is found to be diagnostic in chronic myelogenous leukemia. Forty years later, a drug is developed specifically for the gene product from this chromosome.

1961.....The first codon is identified in the genetic code, relating the nucleotides in messenger RNA to an amino acid in a protein; the other codons are soon identified. High-yielding, semi-dwarf, adaptable wheat is developed by Norman Borlaug in Mexico. It soon results in bumper crops there and in India.

1962.....A frog is cloned by nuclear transplantation from a specialized cell nucleus. This shows totipotency in animal cells. Werner Arber describes bacteriophage restriction and proposes specific restriction endonucleases made by bacteria that cleave incoming phage DNA at specific sequences. Newborn screening for the genetic disease phenylketonuria begins. Its public health success results in other screening programs.

1963.....A high-lysine variety of corn is described from a screen of thousands of genetically different varieties. This leads to quality-protein maize and improvement in human nutrition.

1964.....Robert Holley determines the first sequence of a nucleic acid.

1965.....Nucleic acid hybridization becomes widely used to study relationships between nucleic acids.

1968.....Motoo Kimura proposes evolution by

accumulation of neutral mutations not subject to natural selection. The term "Green Revolution" is used to describe the impact of new wheat and rice varieties on poor regions of the world.

1970.....Norman Borlaug is awarded the Nobel Peace Prize for breeding high-yielding varieties of wheat for use in the poor regions of the world.

1971.....Daniel Nathans cuts and maps a viral genome using a restriction enzyme, showing the potential use of these enzymes for manipulating DNA in the lab. The two-hit model for cancer involving tumor suppressor genes is proposed. This leads to the discovery of these genes and their control of cell division.

1973.....First report of genetically functional recombinant DNA, as genes from different bacteria are spliced together in the lab and then put into a single cell. Soon, human genes are put into bacteria and expressed.

1976.....Recombination of alleles explains the diversity of antibodies.

1977.....DNA sequencing methods are developed and soon automated. This ultimately leads to genome sequencing projects.

1979.....The human insulin gene is expressed in bacteria. This is first drug made by DNA biotechnology, and a new industry is born. Smallpox is eradicated through vaccination.

1982.....A genetically modified bacterium is patented; upheld by the U.S. Supreme Court, this leads to many more patents of organisms and genes.

1983.....The polymerase chain reaction is invented.

1984.....The Human Genome Project is first proposed.

1985.....Instruments are invented to make DNA sequences in the lab; custom DNA is now possible. The dystrophin gene that is defective in Duchenne muscular dystrophy is isolated. DNA identification by repeated sequences is invented by Alec Jeffreys. It soon has many applications in forensics.

1989.....Bacteria with genes for digesting oil are used in bioremediation of the oil spill from the tanker *Exxon Valdez*.

1990.....The novel *Jurassic Park* brings DNA technology to public attention. Functional human antibodies are made in transgenic plants. Gene therapy is successfully done on a patient with immunodeficiency, leading to much hope and hype. A transgenic cow makes a human protein in its milk; this leads to "pharming." Preimplantation genetic diagnosis is done on an eight-celled embryo to screen for mutant alleles for cystic fibrosis. None are found, and a normal baby is born.

1994.....Quality-protein maize is developed by plant geneticists. It becomes widely used and improves human nutrition as well as crop

production.

1995.....The first genome of an organism is sequenced: a bacterium that causes meningitis. New genes are found and other sequences rapidly follow. First gene expression analysis by DNA microarray is performed, leading to many applications for basic science and medical diagnosis, as well as the development of computing tools to analyze a mass of data.

1996.....Dolly the sheep, the first cloned mammal, is born. Other mammals are soon cloned.

1998.....Human embryonic stem cells are grown in the laboratory, making possible their use in medicine. The RNAi mechanism is explained to shut off specific gene expression. Many applications from basic research to drug development follow. Gleevec, the prototype drug in molecular medicine, enters clinical trials for a type of leukemia and is very successful.

1999.....The first death due to a gene therapy clinical trial temporarily halts all such therapies in the U.S. Golden rice, rich in beta-carotene, is made by genetic modification of rice plants. Debate continues on the possible dangers of genetically modified crops.

2000.....Drafts of the entire human genome sequence are completed (the final sequence is announced in 2003).

2001.....Salt-tolerant tomato plants are made by genetic modification, opening up the possibility of genetically adapting crops to the environment.

2002.....Stem cells are isolated from fat, one of several non-embryonic stem cells that can be used to develop specialized cells for medicine.

2005.....The chimp genome is completed, leading to comparisons with the human genome to find differences.

2006.....Gene therapy augments the immune system rejection of a melanoma tumor in a patient.

Glossary

adeno-associated virus: A small virus that infects human cells, incorporates its DNA into the human genome, but does not cause disease. It is used as a vector for human gene therapy.

adenovirus: A DNA virus that causes the common cold; when disabled from reproducing, it is used as a vector for human gene therapy. It does not incorporate its DNA into the human genome.

allele: One of the different forms of a particular gene. Alleles have different DNA sequences.

amino acid: A chemical building block for proteins. There are 20 different amino acids, and their chemical properties determine the properties of proteins.

anabolism: The chemical conversions in biochemistry in which energy is used to make more complex molecules from simple ones. The assembly of proteins from individual amino acids is an example.

angiogenesis: The formation of new blood vessels, which occurs when organs form in development, and when tumors and metastases form in cancer.

annotation: The assignment of amino acid sequences and functions to proteins from DNA sequence data.

antibody: A protein made by immune system cells that can bind specifically to an antigen, or nonself chemical grouping.

antigen: A chemical grouping whose three-dimensional structure marks it as nonself and so provokes an immune response.

B cell: A type of white blood cell that makes an antibody.

bacteriophage: A virus that infects and reproduces in bacteria.

bacterium: A single-celled prokaryotic organism. This classification excludes a special group of prokaryotes called the Archaea.

balanced polymorphism: A gene with alleles such that an allele is maintained in a population despite its disadvantage for reproduction because it provides an overriding advantage as well. For instance, the sickle cell allele for hemoglobin is disadvantageous because it can result in sickle cell disease but advantageous because it can result in resistance to malaria.

base pairs (nucleotides): Chemical interactions between nucleotides in the same or opposite strands of nucleic acids. A pairs with T or U; G pairs with C.

bioinformatics: The use of computing to analyze information in biological systems, especially extensive data from DNA and protein sequences.

bioremediation: The use of organisms, especially bacteria, to remove pollutants from the environment. For example, bacteria that digest oil are used to clean up accidental oil spills.

biosensor: The use of an organism or cells to provide information about another organism or cell or environmental condition.

biotechnology: The use of organisms to perform functions useful to people.

blending inheritance: The idea that when hereditary determinants from two parents come together after fertilization of the egg by sperm, the individual determinants disappear and do not have a separate existence but are irreversibly blended. Mendel's experiments led to a rejection of this idea.

blood typing: The use of genetically inherited allele coding for molecules on the surface of the red blood cell to identify people. The ABO blood groups are an example.

carrier: In genetics, an organism with a normal phenotype that is heterozygous for a recessive allele that determines an unusual phenotype. For example, the parents of a child born with sickle cell disease are often carriers for the allele that causes this disease.

catabolism: The biochemical pathways that involve breaking down complex substances into simpler ones, releasing stored chemical energy. Digestion of proteins into amino acids is an example.

catalyst: A chemical that speeds up a conversion of other substances but is not changed after the conversion.

cell: The basic unit of biological structure, function, and continuity. It contains the genome, as well as the chemical components for biochemistry.

cell cycle: The sequence of events by which a cell reproduces (divides).

chemotherapy: The use of drugs to treat a disease; used most commonly with cancer and some infectious diseases.

chromosome: A DNA molecule containing all or part of the genome of an organism, which has the ability to replicate.

cloning: In organisms, producing genetically identical copies of a cell or organism; in molecular genetics, isolating a gene and making multiple, identical copies of it by insertion into an organism.

DNA: Deoxyribonucleic acid, a polymer of nucleotide building blocks (A, T, G, and C) that acts as the genetic material in most living things.

DNA microarray: A collection of many gene sequences, usually affixed to a glass slide, that act as probes for hybridization in studies of gene expression.

DNA polymerase: An enzyme that catalyzes the polymerization of DNA from nucleotide monomers.

dominant: An allele whose phenotype is expressed in an organism that is either homozygous or heterozygous for the allele.

double helix: A molecule, usually DNA, that contains two interacting strands that curl into a helical form.

enzyme: A biological catalyst that speeds up a biochemical transformation without emerging changed by the process; most enzymes are proteins, although some are RNAs.

essential amino acids: Building blocks of protein that an organism cannot synthesize and that must be taken in the diet; humans must eat 8 essential amino acids and can make the other 12.

eukaryotic cell: A cell with a nucleus and other cell components that are each enclosed within membranes; these cells make up animals and plants.

expression vector: A DNA carrier for molecular cloning that contains sequences for the transcription and translation of the gene to be cloned.

extremophiles: Organisms that can live in environments that would be highly unsuitable for almost all other forms of life; these environments include very high temperature and high salt concentration.

fermentation: The breakdown of carbohydrates by organisms in the absence of oxygen gas, usually to alcohol or lactic acid. More generally, in biotechnology, fermentation is the use of microorganisms or their enzymes to form products from the breakdown of carbohydrates.

gene: The unit of heredity; a sequence of nucleotides on a chromosome that is expressed as a product that is part of the phenotype.

gene library: A collection of the DNA fragments from an organism’s genome carried on vectors; these vectors are usually used to clone the library fragments in bacteria.

gene therapy: In medicine, the introduction of new genes into cells for improvement of the symptoms of a disease.

genealogy: The study of family pedigrees for the purpose of tracing ancestors.

genetic code: The sequence of nucleotides along mRNA that is used to translate the genome into amino acids in protein. The code is virtually the same in all organisms.

genetics: The science of heredity.

genome: The complete nucleic acid (usually DNA) sequence of an organism.

golden rice: Rice that has been genetically engineered to make beta-carotene, a precursor to vitamin A.

green fluorescent protein: A protein from certain jellyfish that gives a bright green, visible glow under ultraviolet light; the gene for this protein has been used as a marker in gene cloning.

growth factor: A protein made in mammals by one tissue that stimulates cell division in a target tissue.

herbicide tolerance: The ability of a plant to be resistant to the effects of an herbicide, a property sometimes conferred

by genetic engineering.

heterozygous: An organism with two different alleles for a particular gene.

homozygous: An organism with two identical alleles for a particular gene.

human leukocyte antigen (HLA): A set of human genes coding for cell surface proteins that are involved in cellular immunity; HLA proteins mark an individual as unique.

hybridization: In genetics, the mating of two individuals homozygous for different alleles to produce heterozygous offspring. In molecular biology, the binding by complementary base pairing (A with T/U and G with C) of nucleic acids from different sources.

hypothyroidism: A disease caused by inadequate production of thyroid hormone. There are both genetic and nongenetic causes; genetic hypothyroidism is screened for in newborns.

independent assortment: The independent segregation of genes on different chromosomes during the formation of gametes in sexual reproduction.

locus: The location of a gene on a chromosome.

metabolism: The sum total of all of the chemical transformations in an organism.

metastasis: The ability of a tumor to break off cells that travel in the blood or lymphatic system to a new location in the body and grow to a satellite tumor.

minimal genome: The genes absolutely necessary for life as deduced by serial inactivation of all genes and testing of an organism for survival.

mitochondrion: The membrane-enclosed "powerhouse of the cell" that releases chemical energy in usable form and contains a small DNA chromosome coding for some of its proteins.

molecular clock: If a DNA sequence mutates at a constant rate over time, and the mutations are not selected by natural selection, comparing the difference of that DNA sequence in two organisms can give an estimate of the time they last had a common ancestor.

mRNA: Messenger RNA is transcribed from one of the two strands of DNA in a gene and carries coding information to the ribosome, where its information is translated to amino acids in protein synthesis.

mutation: A change in the genetic material that is passed on to both daughter cells after cell division. If the cell is a germ line cell, the change can be passed on to offspring and is inherited. If the change is in a somatic cell, it is passed on only to the cells deriving from the original changed cell.

natural selection: The process by which the changing environment causes some individuals with more favorable genes and alleles to have more offspring and pass those alleles to the next generation. This leads to the organism changing through time, or evolving.

neutral mutation: A genetic change that is not subject to natural selection but is nevertheless passed on to the next generation.

nucleic acid: A large molecule or DNA or RNA made up of nucleotide building blocks.

nucleotide: The building block of a nucleic acid. Each nucleotide has an identical sugar and phosphate group and one of five different bases: A, G, C, T, and U.

oncogene: A cellular gene that stimulates cell division and tumor formation when activated. Active oncogenes can be brought into cells by certain tumor viruses.

one gene–one protein: The hypothesis that each gene is expressed by a unique protein that is responsible for its

phenotype at the biochemical level.

open reading frame: A sequence of DNA in a chromosome that codes for a protein.

pesticide: A substance that is used by people to kill unwanted organisms (pests). For example, insecticides are used to kill insect pests.

phagocyte: A white blood cell that surrounds, ingests, and digests foreign substances.

pharming: The use of animals to express and produce human proteins for use as drugs in medicine. Expression is usually in the animals' milk.

phenotype: The outward appearance of genes. It can be influenced by the environment.

phenylketonuria (PKU): A genetic disease in which affected individuals lack an active enzyme called phenylalanine hydroxylase. This causes mental retardation if untreated. Newborns are screened for PKU and if it is found are put on a special diet that results in normal development.

pluripotent: The ability of a stem cell and its offspring to specialize into a few different cell types. For example, some pluripotent stem cells in bone marrow can specialize into red blood cells and several types of white blood cells.

polymerase chain reaction (PCR): A method of amplifying a DNA sequence in the test tube by adding DNA polymerase and other necessary components for replication. A sequence can be amplified a millionfold in a few hours.

polymorphism: **A** difference in a particular DNA sequence between individuals.

population bottleneck: A severe reduction in a population, followed by an expansion by the few remaining individuals. The new population's frequencies of alleles will be a reflection of these few individuals, and this is probably different from the larger, original population.

prokaryotic cell: A cell that lacks a nucleus or other membrane-enclosed components.

promoter: A DNA sequence adjacent to the coding region of a gene, to which RNA polymerase binds to initiate gene expression. The events at the promoter are highly regulated in location and time.

protein: A large molecule composed of amino acid building blocks linked together.

pure-line selection: In crop plant genetics, the repeated selection of a genetically diverse plant population for a specific characteristic. Pure lines are generally homozygous.

quality-protein maize: A variety of corn (maize) that was developed in Mexico during the 1990s that has increased crop yield and more and better protein for human nutrition.

recessive: An allele that is expressed only when homozygous and not expressed when heterozygous (in which case the dominant allele is expressed).

recombinant DNA: DNA molecule made from DNA of two different organisms spliced together in the laboratory.

restriction enzyme: An enzyme made by a microorganism that catalyzes the cleavage of DNA at a specific nucleotide sequence.

reverse genetics: The process by which a gene coding for a phenotype is isolated first, and then the protein involved in that phenotype is isolated.

ribosome: A particle in the cell that acts as the "workbench" and catalyst for protein synthesis.

RNA: Ribonucleic acid, a polymer of the nucleotides A, G, C, and U. There are several types of RNA in the cell, such as transfer RNA and messenger RNA.

RNA polymerase: An enzyme that catalyzes the formation of RNA from nucleotides, using base pairing to DNA as a

template.

RNAi: The use of double-stranded RNA to inhibit the translation to protein of a specific mRNA. A cellular mechanism cuts a larger RNAi into small, single-stranded fragments that actually do the inhibition.

screening: In medicine, the presumptive identification of an individual with a disease by the use of a rapid test that can be applied to large numbers of people.

segregation: In genetics, the separation of the two alleles for a gene into different cells during gamete formation.

semiconservative replication: The mechanism of duplication of DNA whereby each of the two strands in the parental DNA acts as a template for a new strand by complementary base pairing so that each of the two DNA molecules produced has one parental and one new strand.

sex determination: Primary sex determination is the genetically determined formation of either male or female gametes. Secondary sex determination is the outward appearance of the individual with regard to male and female characteristics.

short tandem repeat: A DNA polymorphism in which a sequence of 2–10 base pairs is repeated a number of times, in an inherited manner.

shotgun sequencing: DNA sequencing in which a large DNA is fragmented into many smaller pieces, each piece sequenced, and the sequences aligned by computer analysis.

single nucleotide polymorphism: A DNA sequence that varies between individuals of species by a single base pair.

stem cell: A cell in the body that replicates continuously and can form certain specialized cells.

substrate: In biochemistry, a molecule that is acted upon by an enzyme to make a product.

synthetic biology: The attempt to use knowledge of the minimal genome to synthesize a DNA genome in the laboratory and put it inside an enclosed space to custom-make a cell.

synthetic DNA: DNA artificially made in the laboratory by nonbiological methods.

T cell: A white blood cell involved in cellular immunity.

therapeutic cloning: The use of cloning by donor nuclear transplantation to produce an embryo and then embryonic stem cells that can provide specialized cells for therapy on the nuclear donor.

totipotent: A cell able to produce all other cells of an organism. For example, a fertilized egg is totipotent.

transcription: The expression of a gene by the production of RNA from a DNA template, catalyzed by RNA polymerase.

transformation: The introduction of DNA from an outside source to a cell, causing it to become genetically different.

transgenic: A eukaryotic organism, usually an animal, that has received and integrated DNA from a different organism.

translation: The synthesis of a chain of amino acids as a protein in response to the information of nucleotide sequence in a gene as appearing in mRNA.

tumor suppressor gene: A gene in mammals that inhibits cell division and therefore cancer formation.

vector: A DNA molecule that is used as a carrier to bring a foreign gene into a recipient cell.

virus: An infectious particle usually composed of DNA and protein that requires a host cell to replicate.

X-ray diffraction: A physical method that involves measuring the changes in orientation of X rays as they pass through a crystal. It gives information on the three-dimensional arrangement of atoms in the crystal.

Biographical Notes

Allison, Anthony: A British geneticist who found in the late 1940s that the frequency of the allele for sickle cell hemoglobin was highest in Africa in regions where the incidence of malaria was highest. He proposed that the sickle allele was an example of a balanced polymorphism, where the selective advantages of having the allele outweighed the disadvantages.

Arber, Werner: A Swiss biologist whose studies on how bacteria defend themselves against viral infection led to the discovery in 1962 of restriction endonucleases. These enzymes are tools for cutting DNA in the laboratory to prepare recombinant DNA. He was awarded the Nobel Prize in 1978.

Aristotle: A supreme philosopher in ancient Greece, he made careful observations of the natural world, including biology. His ideas on genetics, promulgated around 340 B.C., included one that an individual's inheritance was primarily from the father. He believed that sperm contained tiny humans that were then organized by the menstrual fluid during sexual intercourse.

Avery, Oswald: An American medical researcher, born in Canada, who found in 1944 that the active cellular ingredient that caused genetic transformation in bacteria was DNA. This was a key line of evidence for DNA as the genetic material.

Blumwald, Eduardo: An American plant biologist who studied the mechanisms of tolerance to salty soils. In the late 1990s, he found a gene in the model plant *Arabidopsis* that when active conferred salt tolerance, and when this gene was transferred to tomato plants, they too were salt tolerant. This gene is under active investigation for transfer into other crops that could then grow on salty soils.

Borlaug, Norman: Plant geneticist who led the team in Mexico that bred high-yielding strains of wheat that were adaptable for cultivation in poor regions. The adoption of these strains led to spectacular gains in food production in many areas of the world. In 1970, Borlaug was awarded the Nobel Peace Prize for his discovery, which staved off famines.

Boyer, Herbert: An American biochemist who, along with Stanley Cohen, made the first functional recombinant DNA in a test tube in 1973 by joining genes from two different bacteria and inserting them into a third strain. He then went on to be a founder of the biotechnology company Genentech.

Boyleston, Zabdiel: An American physician who followed the advice of minister Cotton Mather to inoculate people in Boston during a smallpox epidemic in 1721. Almost all of these people survived.

Brock, Thomas: An American microbiologist and author of a leading textbook on the subject, he studied the biochemistry of bacteria that tolerate the extreme heat of hot springs at Yellowstone National Park. In 1970, he discovered that they had a gene coding for a heat-tolerant DNA polymerase. This was later used for the polymerase chain reaction.

Cano, Raul: A Cuban American microbiologist, he and his students isolated and amplified in the late 1980s DNA from insects preserved in amber millions of years ago. This became the scientific basis for a fictional story of cloning dinosaurs in the novel and film *Jurassic Park.*

Chakrabarty, Ananda: A biologist working for General Electric who isolated a genetically modified bacterium capable of breaking down oil. He applied for a patent, and the case went to the U.S. Supreme Court, which in 1980 ruled that his patent was valid. This led to many more patents of organisms and DNA.

Chargaff, Erwin: An American biochemist who studied the chemical composition of DNA and found in 1951 that the ratios of the bases A:T and G:C were always about one and that every species had its own unique composition of the bases. "Chargaff's rules" were important evidence used by Watson and Crick to decipher the double-helical structure of DNA.

Chase, Martha and Hershey, Alfred: In 1952, these American geneticists performed a key experiment that showed that DNA is the genetic material of viruses. Hershey was awarded the Nobel Prize in 1969.

Cohen, Stanley: An American geneticist who worked with Herbert Boyer to make the first functional recombinant DNA in 1973, ushering in a new era of biotechnology.

Collins, Francis: An American physician-geneticist, he led the publicly funded effort to sequence the entire human genome. A draft sequence was announced in 2000 and a final sequence in 2003.

Crichton, Michael: An American novelist trained as a physician, he has written several books with themes related to genetics, notably *Jurassic Park* (1990), in which dinosaurs are cloned using DNA extracted from fossil insects. He is also one of the creators and a producer of the television series *ER*, which often has themes related to genetics.

Crick, Francis: An English physicist and biologist, he was the codiscoverer of the structure of DNA in 1953. In 1959, he proposed a role for RNA in gene expression and predicted the triplet genetic code. He was awarded the Nobel Prize in 1962.

Darwin, Charles: The English naturalist whose theory of evolution by natural selection, proposed in 1858, unified biology and still does. Darwin's idea was that there is a lot of genetic variation among individuals of a species, and those variants best adapted to the environment for reproduction are passed on to the next generation. Thus, species evolve.

DeVries, Hugo: Dutch botanist who, with Carl Correns and Erich von Tschermak, independently verified Gregor Mendel's conclusions about genes and alleles and then discovered Mendel's paper that had been published 34 years previously.

Doll, Richard: English epidemiologist who showed in 1950 that lung cancer is linked to cigarette smoking. His study led to further investigations linking smoking to heart disease.

Druker, Brian: American physician who led the team that developed the drug Gleevec to treat chronic myelogenous leukemia. His 2001 study showing the effectiveness of this targeted drug is a landmark in molecular medicine.

Dulbecco, Renato: Italian-born American virologist whose studies in the late 1950s and 1960s of the molecular genetics of tumor viruses laid the foundation for work on other viruses, including HIV. He was an early proponent of the Human Genome Project. He was awarded the Nobel Prize in 1975.

Fire, Andrew: American geneticist who, along with Craig Mello, discovered the mechanism of RNA interference (RNAi) in 1998. He was awarded the Nobel Prize in 2006.

Folling, Asbjorn: Norwegian physician and chemist who in 1934 discovered the genetic disease phenylketonuria in two children with mental retardation. Newborn screening for phenylketonuria and dietary control since 1963 have led to a significant reduction in these symptoms in affected individuals.

Franklin, Rosalind: British physical chemist whose studies in 1952 of DNA using X-ray diffraction provided clear evidence for the helical nature of the molecule and were a key line of evidence used by Watson and Crick in developing the double-helical structure for DNA.

Garrod, Archibald: British physician who in 1910 proposed that the disease alkaptonuria is due to a genetically controlled absence of an active enzyme. The enzyme was specifically identified years later, and the genetic mutation in DNA in the 1990s.

Gehring, Walter: Swiss geneticist whose pioneering work on the molecular genetics of development led to his codiscovery in 1984 of a DNA sequence called the homeobox that is part of genes that control positional information in development. These genes were initially shown to occur in fruit flies, but similar genes occur throughout the animal world.

Griffith, Frederick: A British public health scientist, he accidentally discovered genetic transformation in bacteria in 1928. His studies laid the groundwork for the proof that DNA is the genetic material responsible for transformation.

Guthrie, Robert: An American pediatrician and microbiologist who developed a simple screening test for phenylketonuria in 1963 and then led a major study that proved its usefulness in early detection of the disease in

newborns.

Hedrick, Marc: American surgeon who isolated stem cells from fat in 2001 and showed that they are pluripotent.

Henslow, John Stevens: An English geologist and botanist, he was the professor at Cambridge who most influenced Charles Darwin when he was a student there. In 1831, he helped get Darwin the position as naturalist on the survey ship *Beagle*.

Hershey, Alfred and Chase, Martha: In 1952, these American geneticists performed a key experiment that showed that DNA is the genetic material of viruses. Hershey was awarded the Nobel Prize in 1969.

Holley, Robert: An American chemist, he led a team that was the first to determine the sequence of a nucleic acid, an 80-base RNA, in 1964 after five years of work. It takes a machine minutes to do this today. He was awarded the Nobel Prize in 1968.

Itakura, Keiichi: A Japanese-born American chemist, he developed methods to make DNA molecules in the chemistry laboratory. In 1978, he synthesized the gene for human insulin, which was then put into bacteria in the first widespread use of recombinant DNA to produce a drug.

Itano, Harvey: An American pathologist who in 1948 discovered that abnormal hemoglobin is the molecular basis of the phenotype in sickle cell disease.

Jeffreys, Alec: A British geneticist, he developed the use of DNA markers to identify organisms in 1985. DNA analysis quickly became widespread in forensics.

Kimura, Motoo: A Japanese evolutionary geneticist, in 1968 he developed the mathematical and theoretic bases for the theory that evolution can occur through the accumulation of neutral mutations that are not subject to natural selection.

Knudson, Arthur: An American physician, his careful examination of a hereditary cancer led in 1971 to the "two-hit" hypothesis, in which he proposed that cancer was partially due to two mutations in alleles of a tumor suppressor gene. Later, these genes were shown to indeed exist, and his hypothesis has been borne out.

Kornberg, Arthur: An American biochemist, in 1956 he was the first to describe DNA polymerase, the enzyme that catalyzed DNA replication. This was a key line of evidence in support of the double-helix model for DNA proposed a few years earlier. He was awarded the Nobel Prize in 1959.

Kunkel, Louis: An American medical geneticist, he showed in 1985 that the phenotype in Duchenne muscular dystrophy, the most common form of this disease, is due to a minor muscle protein called dystrophin. He isolated this protein by reverse genetics, first describing its gene.

Mello, Craig: An American geneticist who, along with Andrew Fire, discovered the mechanism of RNA interference (RNAi) in 1998. He was awarded the Nobel Prize in 2006.

Mendel, Gregor: An Austrian monk whose careful studies of crosses he made on pea plants laid the foundations for the modern science of genetics. His work, published in 1866, was ignored for 35 years but is now revered.

Miescher, Freidrich: A Swiss biologist who was the first person to isolate DNA, which he called "nuclein" when he extracted it from white blood cells in 1868.

Mullis, Kary: An American biochemist, he invented the polymerase chain reaction in 1983. It was an instant hit and is widely used to amplify DNA. He was awarded the Nobel Prize in 1993.

Nathans, Daniel: In 1971, he used a restriction enzyme to cut and map a viral genome, showing the potential use of these enzymes for manipulating DNA in the lab. He was awarded the Nobel Prize in 1978.

Nirenberg, Marshall: An American biochemist, his 1961 experiment identifying the first genetic codeword (a specific triplet of nucleotides that gets translated to a specific amino acid) led to the complete description of the near-universal genetic code. He was awarded the Nobel Prize in 1968.

Pauling, Linus: One of the greatest chemists of the 20th century, in the 1940s and 1950s he described the structure of proteins and was instrumental in identifying the alterations in hemoglobin in sickle cell disease. He was awarded two Nobel Prizes: for chemistry in 1954 and for peace (for leading the opposition to atmospheric testing of nuclear bombs) in 1962.

Potrykus, Ingo: Swiss plant biologist who spent his career studying crop plants of importance to the poor regions of the world. Using DNA technologies, he developed golden rice in 1999 to help improve the diets of millions.

Riggs, Arthur: An American biologist who did important early work on chromosome replication in eukaryotes. In 1978, he inserted the gene for human insulin into bacteria in the first widespread use of recombinant DNA to produce a drug.

Rous, Peyton: An American physician who in 1910 discovered that the cause of a type of cancer in chickens is transmitted from one animal to another; it was the first demonstration of a tumor virus. He was awarded the Nobel Prize in 1966.

Sanger, Frederick: An English biochemist who developed the methods to determine the sequence of the amino acids in proteins, for which he was awarded the Nobel Prize in 1958. Then he developed a method to determine the sequence of nucleotides in DNA, for which he was awarded a second Nobel Prize in 1980. He has been a leader in the genome sequencing effort.

Shull, George: An American geneticist who in 1908 crossed two pure lines of corn and first described the hybrid vigor of heterosis. Hybrid corn varieties soon became the major ones planted.

Skorecki, Karl: An Israeli physician who in 1997 used DNA polymorphisms to show that Jewish priests in the Cohanim tradition were all descended in a male line from an ancestor of several thousand years ago.

Smith, Hamilton: An American physician and microbiologist who was one of the first scientists to characterize the chemistry of restriction enzymes and in 1978 was awarded the Nobel Prize. He has also been a leader in DNA sequencing efforts and in 1975 was in the team that sequenced the first bacterial genome.

Spiegelman, Sol: An American microbiologist and biochemist who used nucleic acid hybridization, a technique he pioneered in 1961, to show that only one of the two strands of DNA in a given region is expressed and transcribed into RNA.

Steward, Frederick: An American plant physiologist who cloned entire carrot plants from specialized cells, thus demonstrating totipotency in 1958.

Thomson, James: An American cell biologist who isolated human embryonic stem cells in 1998 and grew them in the laboratory, a major advance in potentially using them for therapy.

van Leeuwenhoek, Antonie: The Dutch founder of cell biology was the first person to report the existence of living cells that he observed under the recently-invented microscope.

van 't Veer, Laura: A Dutch physician and cancer researcher who used DNA microarrays to study gene expression in different breast cancers and in 2002 developed a gene expression signature to differentiate tumors with good from poor prognosis, a major event in molecular medicine.

Vasal, Surinder: An Indian geneticist and plant breeder who works in Mexico, he developed quality-protein maize (corn) in 1990; this variety has improved the diets of millions of people while yielding high amounts of grain.

Venter, Craig: An American biologist whose private industry team invented shotgun sequencing, a way to rapidly get the sequence of a large genome. This led to the sequences of bacteria, the fruit fly, and the human genomes. He has also been active in finding new organisms through genomics, and in synthetic biology.

Watson, James: An American biologist who was the codiscoverer of the double-helix model for DNA in 1953. He was

awarded the Nobel Prize in 1962.

Wilkins, Maurice: An English physical chemist who did studies on DNA by X-ray diffraction in early 1953 that led to the double-helix model. He was awarded the Nobel Prize in 1962.

Wilmut, Ian: An English biologist who with Keith Campbell led the team that in 1996 cloned Dolly the sheep by nuclear transplantation into an egg. Dolly was the first mammal to be cloned.

Yalow, Rosalyn: An American medical physicist who in 1960, with colleague Sidney Berson, invented immunoassay—a way to determine tiny quantities of substances. The ability to measure hormones revolutionized medicine. She was awarded the Nobel Prize in 1977.

Bibliography

Essential Reading:

Acquah, George. *Understanding Biotechnology.* Upper Saddle River, NJ: Pearson Prentice-Hall, 2004. A self-testing, internet-based outline.

Altieri, Miguel. *Genetic Engineering in Agriculture: The Myths, Environmental Risks and Alternatives*. Oakland, CA: Food First, 2005. Makes the case that genetically modified crops might not be a good thing.

Barnum, Susan. *Biotechnology: An Introduction.* Belmont, CA: Thomson Brooks-Cole, 2005. A somewhat advanced, but accessible, textbook for college-level courses.

Bellomo, Michael. *The Stem Cell Divide.* New York: AOMCOM Press, 2005. A balanced view of the science and the controversy.

Berg, Jeremy M., John Tymoczko, and Lubert Stryer. *Biochemistry*. 6th ed. New York: W. H. Freeman, 2006. Elegantly written and illustrated, this is an authoritative and comprehensive textbook.

Brookes, Martin. *Fly: The Unsung Hero of 20th-Century Science.* New York: HarperCollins, 2001. The history of genetics through its most popular organism of study.

Brooks, Gavin. *Gene Therapy: Using DNA as a Drug.* London: Pharmaceutical Press, 2003. Comprehensive introduction to the strategies and pitfalls of human gene therapy.

Cain, Michael, Hans Damman, Robert Lue, and Carol Yoon. *Discover Biology*. 3rd ed. New York: W. W. Norton, 2007. An excellent and interesting approach to general biology for the nonspecialist.

Carlson, Elof Axel. *The Unfit: A History of a Bad Idea.* Woodbury, NY: Cold Spring Harbor Laboratory Press, 2001. A very readable and fascinating account of the history of eugenics.

Chrispeels, Maarten and David Sadava. *Plants, Genes and Crop Biotechnology.* Sudbury, MA: Jones and Bartlett, 2003. A comprehensive introduction to plant biology in an agricultural context, with extensive coverage of applied genetics.

Crick, Francis. *What Mad Pursuit: A Personal View of Scientific Discovery.* New York: Basic Books, 1988. The discovery of the structure of DNA through the eyes of one of the scientists involved. See also Watson, James D.

Darwin, Charles. *The Origin of Species*. New York: Random House, 1979. Originally published in 1859. If there is one biology book that everyone should read, this is it.

Davies, Kevin. *Cracking the Genome: Inside the Race to Unlock Human DNA.* New York: The Free Press, 2001. The story of how publicly- and privately-financed groups sequenced the human genome, finishing at the same time.

Dawkins, Richard. *The Selfish Gene: 30th Anniversary Edition.* Oxford: Oxford University Press, 2006. What if genes use organisms to propagate themselves, instead of the other way around? This book and its author revolutionized the way some biologists think about evolution and genetics.

Denniston, Katherine and Joseph Topping. *Introduction to General, Organic and Biochemistry.* 4th ed. New York: McGraw-Hill, 2003. In one compact and well-illustrated volume, this text delivers what the title says: three college courses.

DeSalle, Rob and Michael Yudell. *Welcome to the Genome: A User's Guide to the Genetic Past, Present and Future.*

Hoboken, NJ: John Wiley and Sons, 2005. An easy-to-read, fascinating summary of what genome studies mean.

DeVita Jr., Vincent, Samuel Hellman, and Steven Rosenberg. *Cancer: Principles and Practice of Oncology.* Philadelphia: Lippincott Williams and Wilkins, 2005. The authoritative book on cancer in all of its aspects.

Devlin, Thomas. *Textbook of Biochemistry with Clinical Correlations.* 6th ed. Hoboken, NJ: Wiley-Liss, 2006. Excellent, in-depth treatment of medical biochemistry, with many examples of genetic diseases.

Entine, Jon. *Let Them Eat Precaution: How Politics is Undermining a Genetic Revolution in Agriculture.* Washington, DC: AEI Press, 2006. Makes the case that genetically modified crops are a good thing.

Evans, Gareth and Judith Furlong. *Environmental Biotechnology: Theory and Application.* West Sussex, England: John Wiley and Sons, 2004. Although written for specialists, this collection of articles is surprisingly accessible and comprehensive.

Federoff, Nina and Nancy Brown. *Mendel in the Kitchen.* Washington, DC: National Academies Press, 2004. Excellent overview of genetics and plant breeding.

Fisher, Barry, David Fisher, and Jason Kowolski. *Forensics Demystified.* New York: McGraw-Hill, 2007. An excellent, readable, and well-illustrated introduction to the use of genetics and DNA technology in forensics.

Futuyma, Douglas. *Evolution.* Sunderland, MA: Sinauer Associates, 2005. A comprehensive textbook on the subject, with excellent chapters on DNA studies.

Goldsby, Richard, Thomas Kindt, and Barbara Osborne. *Kuby Immunology.* 6th ed. New York: W. H. Freeman, 2006. An excellent, well-illustrated, and comprehensive text, with medical applications throughout.

Gralla, Jay and Preston Gralla. *The Complete Idiot's Guide to Understanding Cloning.* New York: Penguin, 2004. Don't let the title repel you: This is a solid book written by a professor-scientist and journalist.

Griffiths, Anthony, Susan Wessler, Richard Lewontin, William Gelbart, David Suzuki, and Jeffrey Miller. *An Introduction to Genetic Analysis*. 8th ed. New York: W. H. Freeman, 2007. An excellent, comprehensive text written for undergraduates.

Henig, Robin. *The Monk in the Garden: The Lost and Found Genius of Gregor Mendel, the Father of Genetics.* Boston: Houghton-Mifflin, 2001. A vivid portrayal of Mendel and his times.

Judson, Horace Freeland. *The Eighth Day of Creation: Makers of the Revolution in Biology.* Woodbury, NY: Cold Spring Harbor Laboratory Press, 1996. If you ever wanted to know how science is really practiced and how real people laid the foundations of the revolution in molecular biology, read this book.

Lewin, Benjamin. *Genes VIII.* Upper Saddle River, NJ: Pearson Prentice-Hall, 2005. A well-written, authoritative textbook of DNA and molecular biology. Assumes knowledge of university-level general biology.

Lewis, Ricki. *Human Genetics*. 7th ed. New York: McGraw-Hill, 2006. A superb, readable textbook aimed at nonscientists and written by a talented science writer–biologist.

Lindee, Susan. *Moments of Truth in Genetic Medicine.* Baltimore: Johns Hopkins, 2005. A historian uses several well-known cases to show how far we have come as a society in dealing with human genetic diseases.

Litwack, Gerald. *Human Biochemistry and Disease*. New York: Academic Press, 2007. Using human diseases as a starting point, this book has an innovative and interesting approach to a complex subject.

Lodish, Harvey, Arnold Berk, Paul Matsudaira, Chris Kaiser, Monty Krieger, Matthew P. Scott, Lawrence Zipursky, and James Darnell. *Molecular Cell Biology.* 5th ed. New York: W. H. Freeman, 2005. The authoritative text of cell biology and biochemistry.

Maddox, Brenda. *Rosalind Franklin: The Dark Lady of DNA.* New York: HarperCollins, 2002. The riveting story of the scientist whose role in the discovery of the structure of DNA was long underappreciated.

Massie, Robert. *The Romanovs: The Final Chapter*. New York: Random House, 1995. The fascinating story of the remains of the Russian royal family, and how DNA technology was used to identify them.

McCabe, Linda L. and Edward E. R. B. McCabe. *DNA: Promise and Peril*. Berkeley: University of California Press,

2008. Two leading human geneticists look at how the exploding knowledge of the human genome impacts medical practice and how we look at ourselves.

McConkey, Edwin. *How the Human Genome Works.* Sudbury, MA: Jones and Bartlett, 2004. Excellent, concise coverage of the genome, written for the nonscientist.

Meyer, Anna. *The DNA Detectives: How the Double Helix is Solving Puzzles of the Past.* New York: Avalon Publishing, 2005. A readable account of the use of DNA forensics in history and human evolution.

Mullis, Kerry. *Dancing Naked in the Mind Field.* New York: Vantage Books, 2000. How an iconoclastic scientist discovered the polymerase chain reaction (PCR), won the Nobel Prize, and revolutionized DNA studies.

Panno, Joseph. *Gene Therapy: Treating Disease by Repairing Genes.* New York: Facts on File, 2005. A readable account of how gene therapy can be done and the challenges it faces in medicine.

Pasternak, Jack. *An Introduction to Human Molecular Genetics*. 2nd ed. Bethesda, MD: Fitzgerald Science Press, 2005. An authoritative text that is well organized and is a useful reference.

Pierce, Benjamin. *Genetics: A Conceptual Approach.* New York: W. H. Freeman, 2005. An excellent textbook for undergraduates with a modest background in biology.

Reilly, Philip. *Abraham Lincoln's DNA and Other Adventures in Genetics*. Woodbury, NY: Cold Spring Harbor Laboratory Press, 2000.

———. *Is It in Your Genes? How Genes Influence Common Disorders and Diseases That Affect You and Your Family*. Woodbury, NY: Cold Spring Harbor Laboratory Press, 2004. One of the best and liveliest books written for the general public on human genetics.

———. *The Strongest Boy in the World and Other Adventures in Genetics.* Woodbury, NY: Cold Spring Harbor Laboratory Press, 2006. Fascinating stories on how the human genome is changing medicine, written by a physician-lawyer.

Ridley, Mark. *Evolution (Oxford Readers).* Oxford: Oxford University Press, 2005. A compact yet in-depth discussion of what the great thinkers on the science of evolution thought.

Ridley, Matt. *Genome: The Autobiography of a Species in 23 Chapters (P.S.)* New York: HarperCollins, 2006. A science writer delves into the human genome in a lively and interesting way.

Robbins-Roth, Cynthia. *From Alchemy to IPO: The Business of Biotechnology.* Cambridge, MA: Perseus Group, 2001. How discoveries of DNA manipulation led companies listed on the stock exchange.

Sadava, David, Craig Heller, Gordon Orians, William Purves, and David Hillis. *Life: The Science of Biology*. 8th ed. Sunderland, MA: Sinauer Associates; New York: W. H. Freeman, 2008. An introductory biology text at the college level, with many illustrations of classic experiments and extensive coverage of genetics and DNA.

Scott, Christopher Thomas. *Stem Cell Now: From the Experiment That Shook the World to the New Politics of Life.* New York: Penguin, 2005. Written by a molecular biologist–journalist, this balanced treatment of a controversial subject is scientifically accurate and presents the ethical issues dispassionately.

Smith, John. *Biotechnology.* 4th ed. Cambridge: Cambridge University Press, 2004. An excellent textbook on genetic manipulation and its social consequences.

Sompayrac, Lauren. *How Cancer Works.* Sudbury, MA: Jones and Bartlett, 2004. Excellent, concise coverage of cancer, written for the nonscientist.

———. *How the Immune System Works.* Malden, MA: Blackwell Sciences, 2002. Readable, well-illustrated, and compact, this is a remarkable book for the nonscientist on a complicated subject.

Stone, Linda, Paul Lurquin, and Luca Cavalli-Sforza. *Genes, Culture and Human Evolution: A Synthesis.* Malden, MA: Blackwell, 2007. Lucidly written by leaders in the field, the book examines how DNA studies of human groups relate to anthropology.

Tannock, Ian, Richard Hill, Robert Bristow, and Lea Harrington. *The Basic Science of Oncology.* 4th ed. New York: McGraw-Hill, 2005. Although written by authorities for advanced readers, this text is highly accessible and provides

clear explanations of many phenomena in cancer.

Thieman, William and Michael Palladino. *Introduction to Biotechnology.* San Francisco: Benjamin-Cummings, 2004. A readable text for nonscientists, including issues for discussion. Venter, J. Craig. *My Life Decoded: My Genome*. New York: Viking Press, 2007. A lively view of the present and future of genome sequencing by a leader in the field.

Walker, Sharon. *Biotechnology Demystified*. New York: McGraw-Hill, 2007. A short, self-paced textbook with learning guides that skims the subject well. Watson, James D. *Avoid Boring People: Lessons from a Life in Science*. New York: Oxford, 2007. An inside view of how science really works, by one of the codiscoverers of the model of DNA.

———. *The Double Helix.* New York: Athenaeum, 1968. The discovery of the structure of DNA through the eyes of one of the scientists involved. See also Crick, Francis. ———, Jan Witkowski, Richard Myers, and Amy Caudy. *Recombinant DNA: Genes and Genomics.* 3rd ed. New York:

W. H. Freeman, 2007. A detailed guide to the major discoveries of molecular biology. Weinberg, Robert. *The Biology of Cancer.* New York: Garland Science, 2007. Written by a leading cancer researcher, an up-to-date, well-illustrated, and beautifully written textbook of the science behind cancer. For readers with a desire for advanced knowledge. Zaikov, George. *Biotechnology and Environment*. Haupaugge, NY: Nova Science, 2005. A short, accessible volume from experts on using microbes in environmental cleanup.