

17 Use of Recombinant DNA Techniques in Medicine

The rapid development of techniques in the field of molecular biology is revolutionizing the practice of medicine. The potential uses of these techniques for the diagnosis and treatment of disease are vast.

Clinical applications. **Polymorphisms**, inherited differences in DNA base sequences, are abundant in the human population, and many alterations in DNA sequences are associated with diseases. Tests for DNA sequence variations are more sensitive than many other techniques (such as enzyme assays) and permit recognition of diseases at earlier and therefore potentially more treatable stages. These tests can also identify carriers of inherited diseases so they can receive appropriate counseling. Because genetic variations are so distinctive, **DNA “fingerprinting”** (analysis of DNA sequence differences) can be used to determine family relationships or to help identify the perpetrators of a crime.

Techniques of molecular biology are used in the **prevention** and **treatment** of disease. For example, recombinant DNA techniques provide human insulin for the treatment of diabetes, Factor VIII for the treatment of hemophilia, and vaccines for the prevention of hepatitis. Although treatment of disease by gene therapy is in the experimental phase of development, the possibilities are limited only by the human imagination and, of course, by ethical considerations.

Techniques. To recognize normal or pathologic genetic variations, DNA must be isolated from the appropriate source, and adequate amounts must be available for study. Techniques for **isolating** and **amplifying** genes and studying and **manipulating** DNA sequences involve the use of **restriction enzymes**, **cloning vectors**, **polymerase chain reaction (PCR)**, **gel electrophoresis**, **blotting onto nitrocellulose paper**, and the preparation of **labeled probes that hybridize** to the appropriate target DNA sequences. **Gene therapy** involves isolating normal genes and inserting them into diseased cells so that the normal genes are expressed, permitting the diseased cells to return to a normal state. Students must have at least a general understanding of recombinant DNA techniques to appreciate their current use and the promise they hold for the future.



THE WAITING ROOM



Erna Nemdy, a third-year medical student, has started working in the hospital blood bank two nights per week (see Chapter 15 for an introduction to Erna Nemdy and her daughter, Beverly). Because she will be handling human blood products, she must have a series of hepatitis B vaccinations. She has reservations about having these vaccinations and inquires about the efficacy and safety of the vaccines currently in use.



Cystic fibrosis is a disease caused by an inherited deficiency in the CFTR (cystic fibrosis transmembrane conductance regulator) protein, which is a chloride channel (see Chapter 10, Fig. 10.11). In the absence of chloride secretion, dried mucus blocks the pancreatic duct, resulting in decreased secretion of digestive enzymes into the intestinal lumen. The resulting malabsorption of fat and other foodstuffs decreases growth and may lead to varying degrees of small bowel obstruction. Liver and gallbladder secretions may be similarly affected. Eventually, atrophy of the secretory organs or ducts may occur. Dried mucus also blocks the airways, markedly diminishing air exchange and predisposing the patient to stasis of secretions, diminished immune defenses, and increased secondary infections. Defects in the CFTR chloride channel also affect sweat composition, increasing the sodium and chloride contents of the sweat, thereby providing a diagnostic tool.



Sissy Fibrosa is a 3-year-old Caucasian girl who is diagnosed with cystic fibrosis. Her growth rate has been in the lower 30th percentile over the last year. Since birth, she has had occasional episodes of spontaneously reversible and minor small bowel obstruction. These episodes are superimposed on gastrointestinal symptoms that suggest a degree of dietary fat malabsorption, such as bulky, glistening, foul-smelling stools two or three times per day. She has experienced recurrent flare-ups of bacterial bronchitis in the last 10 months, each time caused by *Pseudomonas aeruginosa*. A quantitative sweat test was unequivocally positive (Excessive sodium and chloride were found in her sweat on two occasions.). Based on these findings, the pediatrician informed Sissy's parents that Sissy probably has cystic fibrosis (CF). A sample of her blood was sent to a DNA testing laboratory to confirm the diagnosis and to determine specifically which one of the many potential genetic mutations known to cause CF was present in her cells.



Carrie Sichel, Will Sichel's 19-year-old sister, is considering marriage. Her growth and development have been normal, and she is free of symptoms of sickle cell anemia. Because a younger sister, Amanda, was tested and found to have sickle trait, and because of Will's repeated sickle crises, Carrie wants to know whether she also has sickle trait (see Chapters 6 and 7 for Will Sichel's history). A hemoglobin electrophoresis is performed that shows the composition of her hemoglobin to be 58% HbA, 39% HbS, 1% HbF, and 2% HbA₂, a pattern consistent with the presence of sickle cell trait. The hematologist who saw her in the clinic on her first visit is studying the genetic mutations of sickle cell trait and asks Carrie for permission to draw additional blood for more sophisticated analysis of the genetic disturbance that causes her to produce HbS. Carrie informed her fiancé that she has sickle cell trait and that she wants to delay their marriage until he is tested.



Victoria Tim (Vicky Tim) was a 21-year-old woman who was the victim of a rape and murder. She left her home and drove to the local convenience store. When she had not returned home an hour later, her father drove to the store, looking for Vicky. He found her car still parked in front of the store and called the police. They searched the area around the store, and found Vicky Tim's body in a wooded area behind the building. She had been sexually assaulted and strangled. Medical technologists from the police laboratory collected a semen sample from vaginal fluid and took samples of dried blood from under the victim's fingernails. Witnesses identified three men who spoke to Vicky Tim while she was at the convenience store. DNA samples were obtained from these suspects to determine whether any of them was the perpetrator of the crime.



Ivy Sharer's cough is slightly improved on a multidrug regimen for pulmonary tuberculosis, but she continues to have night sweats. She is tolerating her current AIDS therapy well but complains of weakness and fatigue. The man with whom she has shared "dirty" needles to inject drugs accompanies Ivy to the clinic and requests that he be tested for the presence of HIV.

I. RECOMBINANT DNA TECHNIQUES

Techniques for joining DNA sequences into new combinations (recombinant DNA) were originally developed as research tools to explore and manipulate genes but are now also being used to identify defective genes associated with disease and to correct genetic defects. Even a cursory survey of the current literature demonstrates that these techniques will soon replace many of the current clinical testing procedures. At least a basic appreciation of recombinant DNA techniques is required to understand the ways in which genetic variations among individuals are determined

and how these differences can be used to diagnose disease. The first steps in determining individual variations in genes involve isolating the genes (or fragments of DNA) that contain variable sequences and obtaining adequate quantities for study.

A. Strategies for Obtaining Fragments of DNA and Copies of Genes

1. RESTRICTION FRAGMENTS

Enzymes called restriction endonucleases enable molecular biologists to cleave segments of DNA from the genome of various types of cells or to fragment DNA obtained from other sources. A restriction enzyme is an endonuclease that specifically recognizes a short sequence of DNA, usually 4 to 6 base pairs (bp) in length, and cleaves a phosphodiester bond in both DNA strands within this sequence (Fig. 17.1). A key feature of restriction enzymes is their specificity. A restriction enzyme always cleaves at the same DNA sequence and only cleaves at that particular sequence. Most of the DNA sequences recognized by restriction enzymes are palindromes, that is, both strands of DNA have the same base sequence when read in a 5' to 3' direction. The cuts made by these enzymes are usually “sticky” (that is, the products are single-stranded at the ends, with one strand overhanging the other). However, sometimes they are blunt (the products are double-stranded at the ends, with no overhangs). Hundreds of restriction enzymes with different specificities have been isolated (Table 17.1).

Restriction fragments of DNA can be used to identify variations in base sequence in a gene. However, they also can be used to synthesize a recombinant DNA (also called chimeric DNA), which is composed of molecules of DNA from different

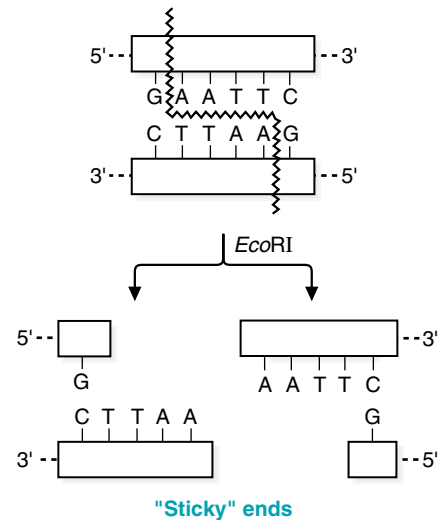


Fig. 17.1. Action of restriction enzymes. Note that the DNA sequence shown is a palindrome; each strand of the DNA, when read in a 5' to 3' direction, has the same sequence. Cleavage of this sequence by *EcoRI* produces single-stranded (or “sticky”) ends or tails. Not shown is an example of an enzyme that generates blunt ends (see Table 17.1).



In sickle cell anemia, the point mutation that converts a glutamate residue to a valine residue (GAG to GTG) occurs in a site that is cleaved by the restriction enzyme *MstII* (recognition sequence CCTNAGG, where N can be any base) within the normal β -globin gene. The sickle cell mutation causes the β -globin gene to lose this *MstII* restriction site. Therefore, because **Will Sichel** is homozygous for the sickle cell gene, neither of the two alleles of his β -globin gene will be cleaved at this site.

Table 17.1. Sequences Cleaved by Selected Restriction Enzymes*

Restriction Enzyme	Source	Cleavage Site
<i>AluI</i>	<i>Arthrobacter luteus</i>	5' - A G C T - 3' 3' - T C G A - 5'
<i>BamHI</i>	<i>Bacillus amyloliquefaciens</i> H	5' - G G A T C C - 3' 3' - C C T A G G - 5'
<i>EcoRI</i>	<i>Escherichia coli</i> RY13	5' - G A A T T C - 3' 3' - C T T A A G - 5'
<i>HaeIII</i>	<i>Haemophilus aegyptius</i>	5' - G G C C - 3' 3' - C C G G - 5'
<i>HindIII</i>	<i>Haemophilus influenzae</i> R _d	5' - A A G C T T - 3' 3' - T T C G A A - 5'
<i>MspI</i>	<i>Moraxella</i> species	5' - C C G G - 3' 3' - G G C C - 5'
<i>MstII</i>	<i>Microcoleus</i>	5' - C C T N A G G - 3' 3' - G G A N T C C - 5'
<i>NotI</i>	<i>Nocardia otitidis</i>	5' - G C G G C C G C - 3' 3' - C G C C G G C G - 5'
<i>PstI</i>	<i>Providencia stuartii</i> 164	5' - C T G C A G - 3' 3' - G A C G T C - 5'
<i>SmaI</i>	<i>Serratia marcescens</i> S _b	5' - C C C G G G - 3' 3' - G G G C C C - 5'

*Restriction enzymes are named for the bacterium from which they were isolated (e.g., *ECO* is from *Escherichia coli*).



Which **ONE** of the following sequences is most likely to be a restriction enzyme recognition sequence?

- (A) (5') G T C C T G (3')
C A G G A C
- (B) (5') T A C G A T (3')
A T G C T A
- (C) (5') C T G A G (3')
G A C T C
- (D) (5') A T C C T A (3')
T A G G A T



Restriction endonucleases were discovered in bacteria in the late 1960s and 1970s. These enzymes were named for the fact that bacteria use them to “restrict” the growth of viruses (bacteriophage) that infect the bacterial cells. They cleave the phage DNA into smaller pieces so the phage cannot reproduce in the bacterial cells. However, they do not cleave the bacterial DNA, because its bases are methylated at the restriction sites by DNA methylases. Restriction enzymes also restrict uptake of DNA from the environment, and they restrict mating with nonhomologous species.



The answer is C. C follows a palindromic sequence of CTNAG, where N can be any base. None of the other sequences is this close to a palindrome. Although most restriction enzymes recognize a “perfect” palindrome, where the sequence of bases in each strand are the same, others can have intervening bases between the regions of identity, as in this question. Note also the specificity of the enzyme *Mst*II in Table 1.

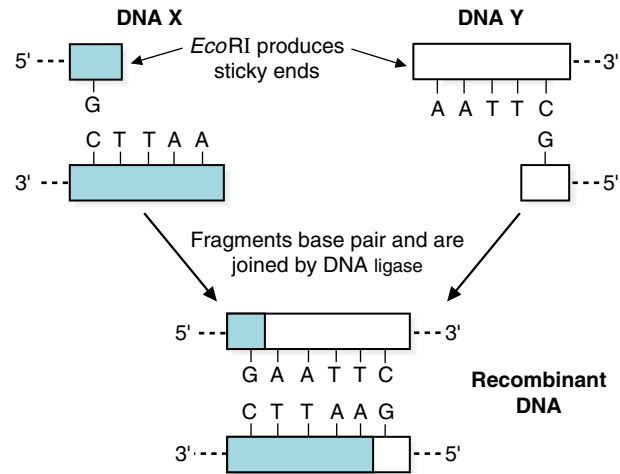


Fig. 17.2. Production of recombinant DNA molecules with restriction enzymes and DNA ligase. The dashes at the 5' and 3'-ends indicate that this sequence is part of a longer DNA molecule.

sources that have been recombined in vitro (outside the organism, e.g., in a test tube). The sticky ends of two unrelated DNA fragments can be joined to each other if they have sticky ends that are complementary. Complementary ends are obtained by cleaving the unrelated DNAs with the same restriction enzyme (Fig. 17.2). After the sticky ends of the fragments base-pair with each other, the fragments can be covalently attached by the action of DNA ligase.

2. DNA PRODUCED BY REVERSE TRANSCRIPTASE

If mRNA transcribed from a gene is isolated, this mRNA can be used as a template by the enzyme reverse transcriptase, which produces a DNA copy (cDNA) of the RNA. In contrast to DNA fragments cleaved from the genome by restriction enzymes, DNA produced by reverse transcriptase does not contain introns because mRNA, which has no introns, is used as a template.

3. CHEMICAL SYNTHESIS OF DNA

Automated machines can synthesize oligonucleotides (short molecules of single-stranded DNA) up to 100 nucleotides in length. These machines can be programmed to produce oligonucleotides with a specified base sequence. Although entire genes cannot yet be synthesized in one piece, oligonucleotides can be prepared that will base-pair with segments of genes. These oligonucleotides can be used in the process of identifying, isolating, and amplifying genes.

B. Techniques for Identifying DNA Sequences

1. PROBES

A probe is a single-stranded polynucleotide of DNA or RNA that is used to identify a complementary sequence on a larger single-stranded DNA or RNA molecule (Fig. 17.3). Formation of base pairs with a complementary strand is called annealing or hybridization. Probes can be composed of cDNA (produced from mRNA by reverse transcriptase), fragments of genomic DNA (cleaved by restriction enzymes from the genome), chemically synthesized oligonucleotides, or, occasionally, RNA.

To identify the target sequence, the probe must carry a label (see Fig. 17.3). If the probe has a radioactive label such as ^{32}P , it can be detected by autoradiography. An autoradiogram is produced by covering the material containing the probe with a sheet of x-ray film. Electrons (β particles) emitted by disintegration of the

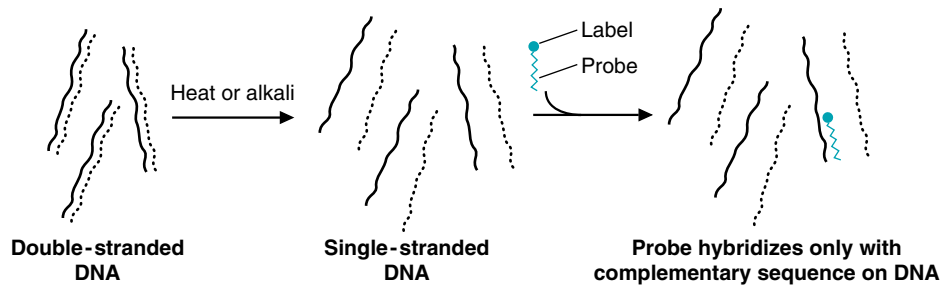


Fig. 17.3. Use of probes to identify DNA sequences. The probe can be either DNA or RNA.

radioactive atoms expose the film in the region directly over the probe. A number of techniques can be used to introduce labels into these probes. Not all probes are radioactive. Some are chemical adducts (compounds that bind covalently to DNA) that can be identified, for example, by fluorescence.

2. GEL ELECTROPHORESIS

Gel electrophoresis is a technique that uses an electrical field to separate molecules on the basis of size. Because DNA contains negatively charged phosphate groups, it will migrate in an electrical field toward the positive electrode (Fig. 17.4). Shorter molecules migrate more rapidly through the pores of a gel than do longer molecules, so separation is based on length. Gels composed of polyacrylamide, which can separate DNA molecules that differ in length by only one nucleotide, are used to determine the base sequence of DNA. Agarose gels are used to separate longer DNA fragments that have larger size differences.

The bands of DNA in the gel can be visualized by various techniques. Staining with dyes such as ethidium bromide allows direct visualization of DNA bands under ultraviolet light. Specific sequences are generally detected by means of a labeled probe.

3. DETECTION OF SPECIFIC DNA SEQUENCES

To detect specific sequences, DNA is usually transferred to a solid support, such as a sheet of nitrocellulose paper. For example, if bacteria are growing on an agar plate, cells from each colony will adhere to a nitrocellulose sheet pressed against the

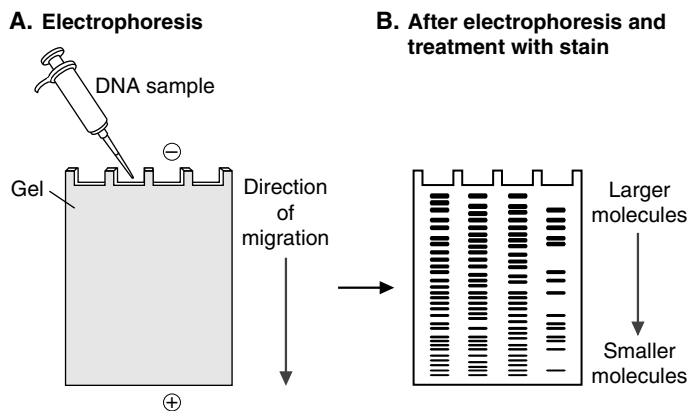


Fig. 17.4. Gel electrophoresis of DNA. **A.** DNA samples are placed into depressions ("wells") at one end of a gel, and an electrical field is applied. The DNA migrates toward the positive electrode at a rate that depends on the size of the DNA molecules. Shorter molecules migrate more rapidly than longer molecules. **B.** The gel is removed from the apparatus. The bands are not visible until techniques are performed to visualize them (see Fig. 17.6).

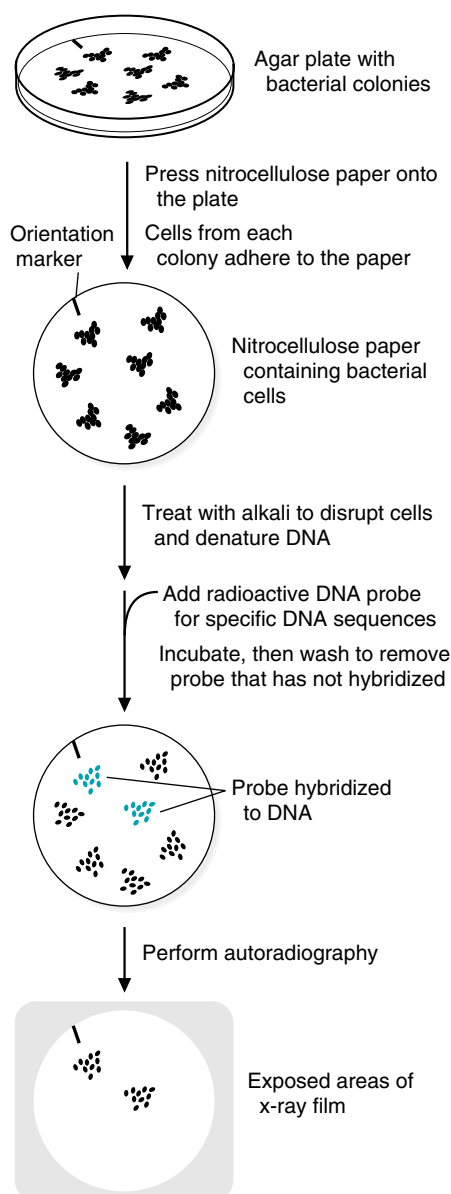



Fig. 17.5. Identification of bacterial colonies containing specific DNA sequences. The autoradiogram can be used to identify bacterial colonies on the original agar plate that contain the desired DNA sequence. Note that an orientation marker is placed on the nitrocellulose and the agar plate so the results of the autoradiogram can be properly aligned with the original plate of bacteria.

 Western blots are one of the tests for the AIDS virus. Viral proteins in the blood are detected by antibodies. Tests performed on **Ivy Sharer's** friend showed that he was HIV positive. Unlike Ivy, however, he has not yet developed the symptoms of AIDS.

agar, and an exact replica of the bacterial colonies can be transferred to the nitrocellulose paper (Fig. 17.5). A similar technique is used to transfer bands of DNA from electrophoretic gels to nitrocellulose sheets. After bacterial colonies or bands of DNA are transferred to nitrocellulose paper, the paper is treated with an alkaline solution. Alkaline solutions denature DNA (that is, separate the two strands of each double helix). The single-stranded DNA is then hybridized with a probe, and the regions on the nitrocellulose blot containing DNA that base-pairs with the probe are identified.

E. M. Southern developed the technique, which bears his name, for identifying DNA sequences on gels. Southern blots are produced when DNA on a nitrocellulose blot of an electrophoretic gel is hybridized with a DNA probe. Molecular biologists decided to continue with this geographic theme as they named two additional techniques. Northern blots are produced when RNA on a nitrocellulose blot is hybridized with a DNA probe. A slightly different but related technique, known as a Western blot, involves separating proteins by gel electrophoresis and probing with labeled antibodies for specific proteins (Fig. 17.6).

4. DNA SEQUENCING

The most common procedure for determining the sequence of nucleotides in a DNA strand was developed by Frederick Sanger and involves the use of dideoxynucleotides. Dideoxynucleotides lack a 3'-hydroxyl group (in addition to lacking the 2' hydroxyl group normally absent from DNA deoxynucleotides). Thus, once they are incorporated into the growing chain, the next nucleotide cannot add, and polymerization is terminated. In this procedure, only one of the four dideoxynucleotides (ddATP, ddTTP, ddGTP, or ddCTP) is added to a tube containing all four normal deoxynucleotides, DNA polymerase, a primer, and the template strand for the DNA that is being sequenced (Fig. 17.7). As DNA polymerase catalyzes the sequential addition of complementary bases to the 3' end, the dideoxynucleotide competes with its corresponding normal nucleotide for insertion. Whenever the dideoxynucleotide is incorporated, further polymerization of the strand cannot occur, and synthesis is terminated. Some of the chains will terminate at each of the locations in the template strand that is complementary to the dideoxynucleotide. Consider, for example, a growing polynucleotide strand in which adenine (A) should add at positions 10, 15, and 17. Competition between ddATP and dATP for each position results in some chains terminating at position 10, some at 15, and some at 17. Thus, DNA strands of varying lengths are produced from a template. The shortest strands are closest to the 5'-end of the growing DNA strand because the strand grows in a 5' to 3' direction.

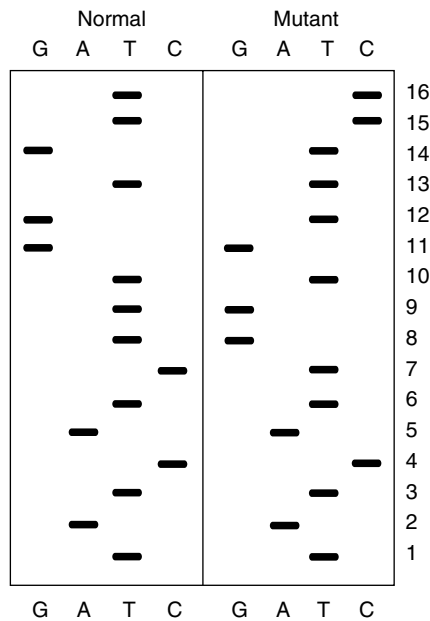
Four separate reactions are performed, each with only one of the dideoxynucleotides present (ddATP, ddTTP, ddGTP, ddCTP) plus a complete mixture of normal nucleotides (see Fig. 17.7B.). In each tube, some strands are terminated whenever the complementary base for that dideoxynucleotide is encountered. If these strands are subjected to gel electrophoresis, the sequence 5'→3' of the DNA strand complementary to the template can be determined by “reading” from the bottom to the top of the gel, that is, by noting the lanes (A, G, C, or T) in which bands appear, starting at the bottom of the gel and moving sequentially toward the top.



Ivy Sharer is being treated with didanosine. This drug is a purine nucleoside composed of the base hypoxanthine linked to dideoxyribose. In cells, didanosine is phosphorylated to form a nucleotide that adds to growing DNA strands. Because dideoxynucleotides lack both 2'- and 3'-hydroxyl groups, DNA synthesis is terminated. Reverse transcriptase has a higher affinity for the dideoxynucleotides than does the cellular DNA polymerase, so the use of this drug will affect reverse transcriptase to a greater extent than the cellular enzyme.



In the early studies on cystic fibrosis, DNA sequencing was used to determine the type of defect in patients. Buccal cells were obtained from washes of the mucous membranes of the mouth, DNA isolated from these cells was amplified by PCR, and DNA sequencing of the CF gene was performed. A sequencing gel for the region in which the normal gene differs from the mutant gene is shown below.



What is the difference between the normal and the mutant CF gene sequence shown on the gel, and what effect would this difference have on the protein produced from this gene?

C. Techniques for Amplifying DNA Sequences

To study genes or other DNA sequences, adequate quantities of material must be obtained. It is often difficult to isolate significant quantities of DNA from the original source. For example, an individual cannot usually afford to part with enough tissue to provide the amount of DNA required for clinical testing. Therefore, the available quantity of DNA has to be amplified.

1. CLONING OF DNA

The first technique developed for amplifying the quantity of DNA is known as cloning (Fig. 17.8). The DNA that you want amplified (the “foreign” DNA) is attached to a vector (a carrier DNA), which is introduced into a host cell that makes multiple copies of the DNA. The foreign DNA and the vector DNA are usually cleaved with the same restriction enzyme, which produces complementary sticky ends in both DNAs. The foreign DNA is then added to the vector. Base pairs form between the complementary single-stranded regions, and DNA ligase joins the molecules to produce a chimera, or recombinant DNA. As the host cells divide, they replicate their own DNA, and they also replicate the DNA of the vector, which includes the foreign DNA.

If the host cells are bacteria, commonly used vectors are bacteriophage (viruses that infect bacteria), plasmids (extrachromosomal pieces of circular DNA that are taken up by bacteria), or cosmids (plasmids that contain DNA sequences from the lambda phage). When eukaryotic cells are used as the host, the vectors are often retroviruses, adenoviruses, free DNA, or DNA coated with a lipid layer (liposomes). The foreign DNA sometimes integrates into the host cell genome or it exists as eiposomes (extrachromosomal fragments of DNA) (See section III.D. of this chapter.)

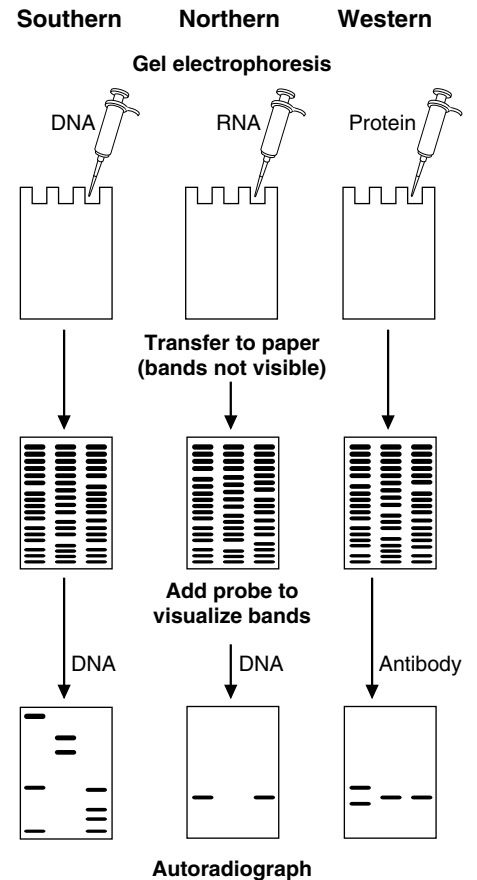


Fig. 17.6. Southern, Northern, and Western blots. For Southern blots, DNA molecules are separated by electrophoresis, denatured, transferred to nitrocellulose paper (by “blotting”), and hybridized with a DNA probe. For Northern blots, RNA is electrophoresed and treated similarly except that alkali is not used. (First, alkali hydrolyzes RNA, and second, RNA is already single stranded.) For Western blots, proteins are electrophoresed, transferred to nitrocellulose, and probed with a specific antibody.



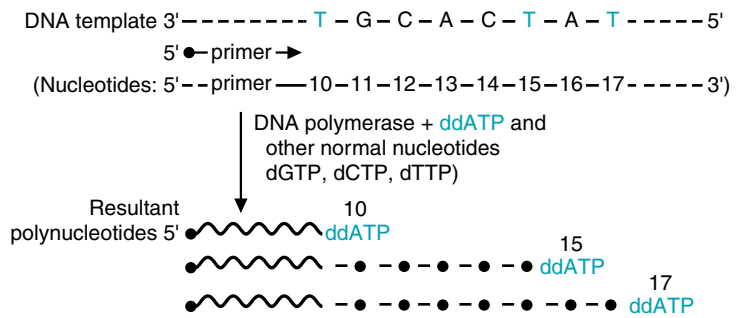
A genomic “library” in molecular biologists’ terms is a set of host cells that collectively contain all of the DNA sequences from the genome of another organism. A cDNA library is a set of host cells that collectively contain all the DNA sequences produced by reverse transcriptase from the mRNA obtained from cells of a particular type. Thus, a cDNA library contains all the genes expressed in that cell type, at the stage of differentiation when the mRNA was isolated.

A: In individuals of northern European descent, 70% of the cases of cystic fibrosis (CF) are caused by a deletion of three bases in the CF gene. In the region of the gene shown on the gels, the base sequence (read from the bottom to the top of the gel) is the same for the normal and mutant gene for the first 6 positions, and the bases in positions 10 through 16 of the normal gene are the same as the bases in positions 7 through 13 of the mutant gene. Therefore, a 3-base deletion in the mutant gene corresponds to bases 7 through 9 of the normal gene.

	Ile	Ile	Phe	Gly
Normal sequence:	T A T C	A T C	T T T	G G T
CF sequence:	T A T C	A T	--	T G G T
	Ile	Ile		Gly

Loss of 3 bp (indicated by the dashes) maintains the reading frame, so only the single amino acid phenylalanine (F) is lost. Phenylalanine would normally appear as residue 508 in the protein. Therefore, the deletion is referred to as ΔF_{508} . The rest of the amino acid sequence of the normal and the mutant proteins is identical.

A. Terminates with ddATP



B. If synthesis is terminated with:

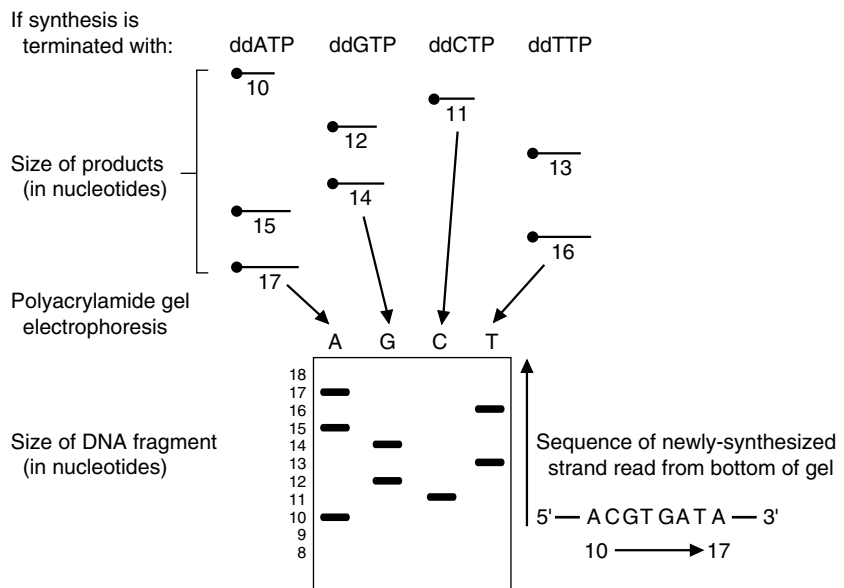


Fig. 17.7. The Sanger method. (A) A reaction mixtures contain one of the dideoxynucleotides, such as ddATP, and some of the normal nucleotide, dATP, which compete for incorporation into the growing polypeptide chain. When a T is encountered on the template strand (position 10), some of the molecules will incorporate a ddATP, and the chain will be terminated. Those that incorporate a normal dATP will continue growing until position 15 is reached, where they will incorporate either a ddATP or the normal dATP. Only those that incorporate a dATP will continue growing to position 17. Thus, strands of different length from the 5' end are produced, corresponding to the position of a T in the template strand. (B) DNA sequencing by the dideoxynucleotide method. Four tubes are used. Each one contains DNA polymerase, a DNA template hybridized to a primer, plus dATP, dGTP, dCTP, and dTTP. Either the primer or the nucleotides must have a radioactive label, so bands can be visualized on the gel by autoradiography. Only one of the four dideoxynucleotides (ddNTPs) is added to each tube. Termination of synthesis occurs where the ddNTP is incorporated into the growing chain. The template is complementary to the sequence of the newly synthesized strand.

Host cells that contain recombinant DNA are called transformed cells if they are bacteria, or transfected cells if they are eukaryotes. Markers in the vector DNA are used to identify cells that have been transformed, and probes for the foreign DNA can be used to determine that the host cells actually contain the foreign DNA. If the host cells containing the foreign DNA are incubated under conditions in which they replicate rapidly, large quantities of the foreign DNA can be isolated from the cells. With the appropriate vector and growth conditions that

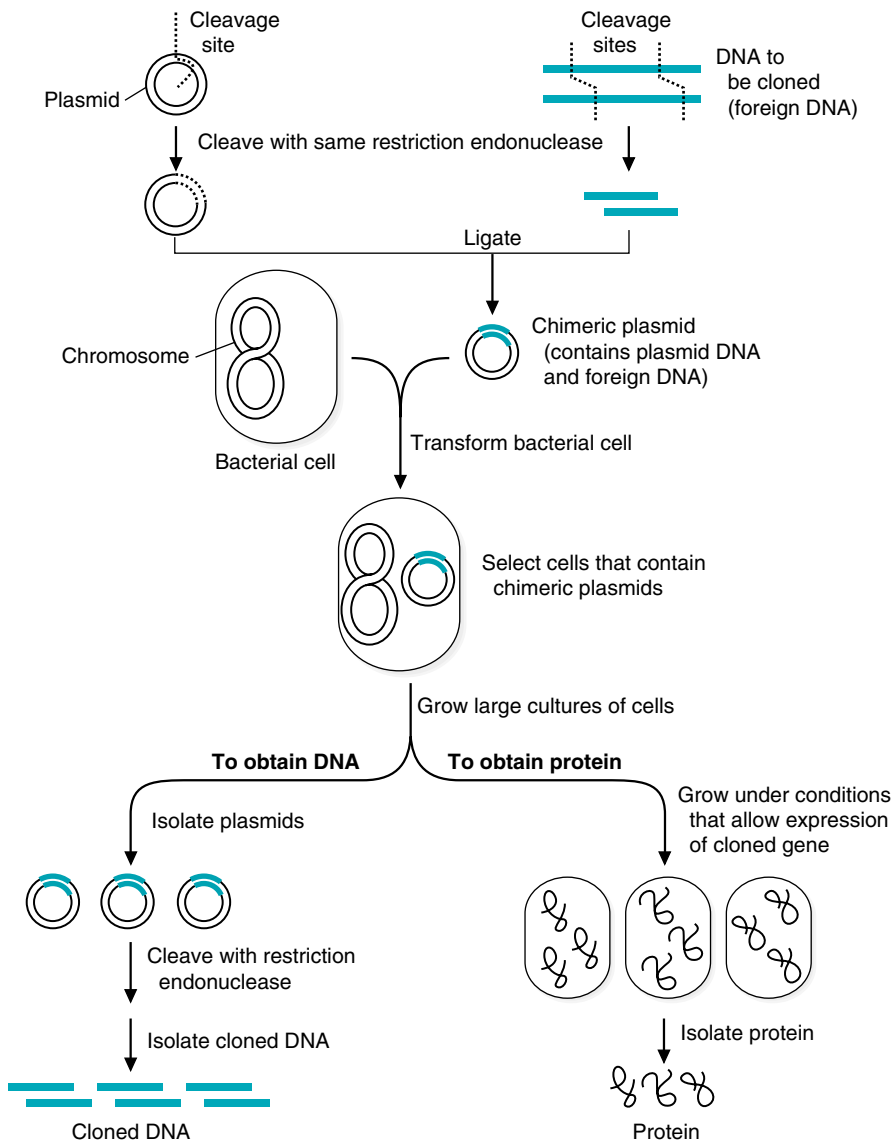


Fig. 17.8. Simplified scheme for cloning of DNA in bacteria. A plasmid is a specific type of vector, or carrier, which can contain inserts of foreign DNA of up to 2.0 kb in size. For clarity, the sizes of the pieces of DNA are not drawn to scale (for example, the bacterial chromosomal DNA should be much larger than the plasmid DNA).

permit expression of the foreign DNA, large quantities of the protein produced from this DNA can be isolated.

2. POLYMERASE CHAIN REACTION (PCR)

PCR is an *in vitro* method that can be used for rapid production of very large amounts of specific segments of DNA. It is particularly suited for amplifying regions of DNA for clinical or forensic testing procedures because only a very small sample of DNA is required as the starting material. Regions of DNA can be amplified by PCR from a single strand of hair or a single drop of blood or semen.

First, a sample of DNA containing the segment to be amplified must be isolated. Large quantities of primers, the four deoxyribonucleoside triphosphates, and a heat-stable DNA polymerase are added to a solution in which the DNA is heated to separate the strands (Fig. 17.9). The primers are two synthetic oligonu-



Although only small amounts of semen were obtained from **Vicky Tim's** body, the quantity of DNA in these specimens could be amplified by PCR. This technique provided sufficient amounts of DNA for comparison with DNA samples from the three suspects.

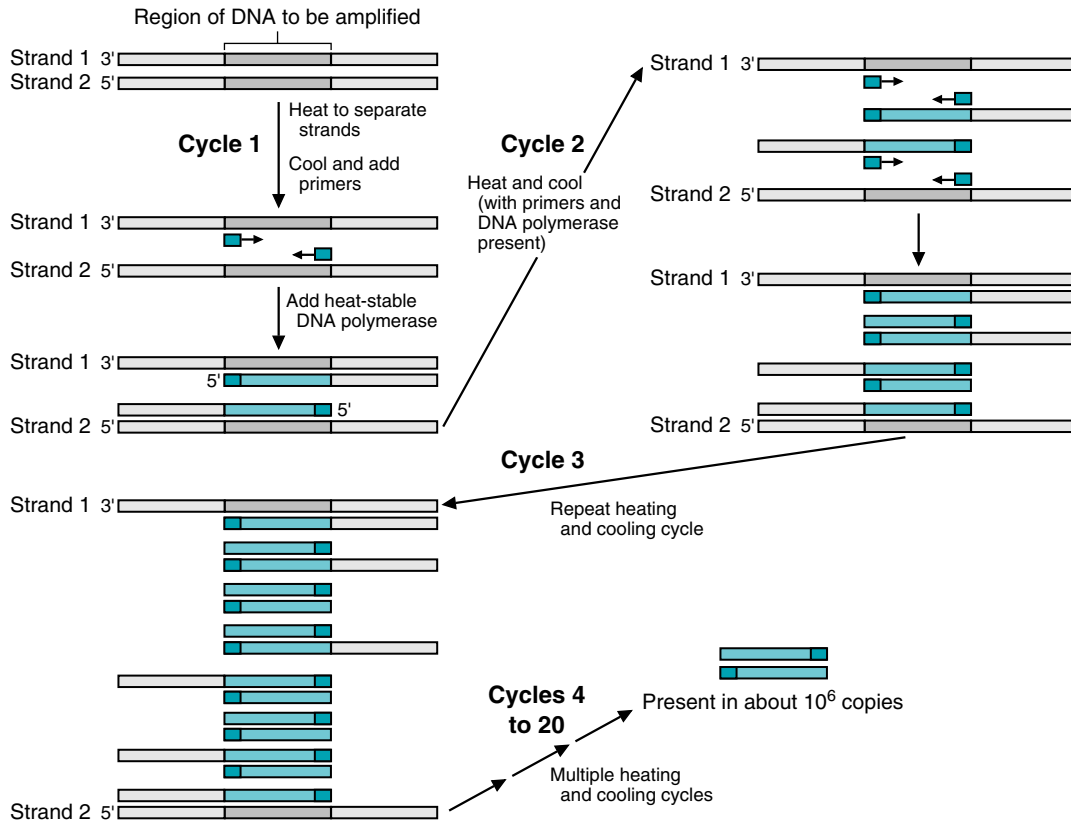



Fig. 17.9. Polymerase chain reaction (PCR). Strand 1 and strand 2 are the original DNA strands. The short dark blue fragments are the primers. After multiple heating and cooling cycles, the original strands remain, but most of the DNA consists of amplified copies of the segment (shown in lighter blue) synthesized by the heat-stable DNA polymerase.

 The DNA polymerase used for PCR is isolated from *Thermus aquaticus*, a bacterium that grows in hot springs. This polymerase can withstand the heat required for separation of DNA strands.

cleotides; one oligonucleotide is complementary to a short sequence in one strand of the DNA to be amplified, and the other is complementary to a sequence in the other DNA strand. As the solution is cooled, the oligonucleotides form base pairs with the DNA and serve as primers for the synthesis of DNA strands by the heat-stable DNA polymerase. The process of heating, cooling, and new DNA synthesis is repeated many times until a large number of copies of the DNA are obtained. The process is automated, so that each round of replication takes only a few minutes and in 20 heating and cooling cycles, the DNA is amplified over a million-fold.

II. USE OF RECOMBINANT DNA TECHNIQUES FOR DIAGNOSIS OF DISEASE

A. DNA Polymorphisms

Polymorphisms are variations among individuals of a species in DNA sequences of the genome. They serve as the basis for using recombinant DNA techniques in the diagnosis of disease. The human genome probably contains millions of different polymorphisms. Some polymorphisms involve point mutations, the substitution of one base for another. Deletions and insertions are also responsible for variations in DNA sequences. Some polymorphisms occur within the coding region of genes. Others are found in noncoding regions closely linked to genes involved in the cause of inherited disease, in which case they can be used as a marker for the disease.

B. Detection of Polymorphisms

1. RESTRICTION FRAGMENT LENGTH POLYMORPHISMS

Occasionally, a point mutation occurs in a recognition site for one of the restriction enzymes. The restriction enzyme therefore can cut at this restriction site in DNA from most individuals, but not in DNA from individuals with this mutation. Consequently, the restriction fragment that binds a probe for this region of the genome will be larger for a person with the mutation than for most members of the population. Mutations also can create restriction sites that are not commonly present. In this case, the restriction fragment from this region of the genome will be smaller for a person with the mutation than for most individuals. These variations in the length of restriction fragments are known as restriction fragment length polymorphisms (RFLPs).

In some cases, the mutation that causes a disease affects a restriction site within the coding region of a gene. However, in many cases, the mutation affects a restriction site that is outside the coding region but tightly linked (i.e., physically close on the DNA molecule) to the abnormal gene that causes the disease. This RFLP could still serve as a biologic marker for the disease. Both types of RFLPs can be used for genetic testing to determine whether an individual has the disease.

2. DETECTION OF MUTATIONS BY ALLELE-SPECIFIC OLIGONUCLEOTIDE PROBES

Other techniques have been developed to detect mutations, because many mutations associated with genetic diseases do not occur within restriction enzyme recognition sites or cause detectable restriction fragment length differences when digested with restriction enzymes. For example, oligonucleotide probes (containing 15–20 nucleotides) can be synthesized that are complementary to a DNA sequence that includes a mutation. Different probes are produced for alleles that contain mutations and for those that have a normal DNA sequence. The region of the genome that contains the abnormal gene is amplified by PCR, and the samples of DNA are placed in narrow bands on nitrocellulose paper (“slot blotting”). The paper is then treated with the radioactive probe for either the normal or the mutant sequence. Autoradiograms indicate whether the normal or mutant probe has preferentially base-paired (hybridized) with the DNA, that is, whether the alleles are normal or mutated. Carriers, of course, have two different alleles, one that binds to the normal probe and one that binds to the mutant probe.

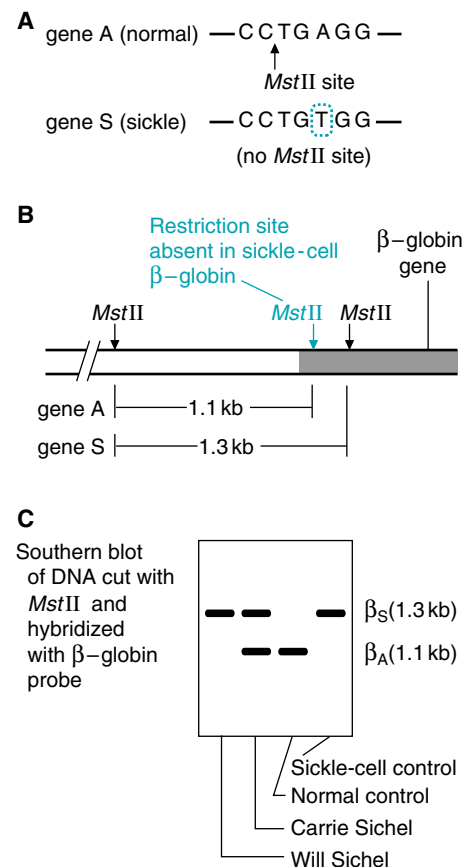


So how does one determine the DNA sequence of a gene that contains a mutation to develop specific probes to that mutation? Initially the gene causing the disease must be identified. This is done by a process known as positional cloning, which involves linking polymorphic markers to the disease. Individuals who express the disease should contain a specific variant of these polymorphic markers, whereas individuals who do not express the disease would not contain these markers. Once such polymorphic markers are identified, they are used as probes to screen a human genomic library. This will identify pieces of human DNA containing the polymorphic marker. These pieces of DNA are then used as probes to expand the region of the genome surrounding this marker. Potential genes within this region are identified (using data available from the sequencing of the human genome), and the sequence of bases within each gene compared with the sequence of bases in the genes of individuals with the disease. The one gene that shows an altered sequence in disease-carrying individuals as compared with normal individuals is the tentative disease gene. Through the sequencing of genes from many people afflicted with the disease, the types of mutations that lead to this disease can be characterized and specific tests developed to look for these specific mutations.



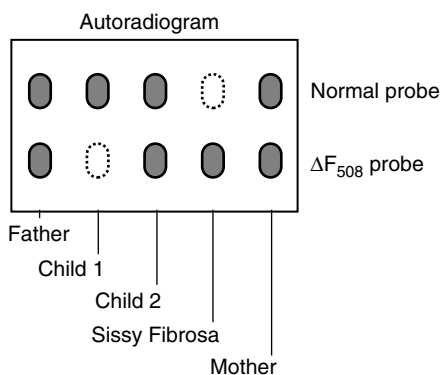
The mutation that causes sickle cell anemia abolishes a restriction site for the enzyme *MstII* in the β -globin gene. The consequence of this mutation is that the restriction fragment produced by *MstII* that includes the 5'-end of the β -globin gene is larger (1.3 kilobases [kb]) for individuals with sickle cell anemia than for normal individuals (1.1 kb). Analysis of restriction fragments provides a direct test for the mutation. In **Will Sichel's** case, both alleles for β -globin lack the *MstII* site and produce 1.3-kb restriction fragments; thus, only one band is seen in a Southern blot.

Carriers have both a normal and a mutant allele. Therefore, their DNA will produce both the larger and the smaller *MstII* restriction fragments. When **Will Sichel's** sister **Carrie Sichel** was tested, she was found to have both the small and the large restriction fragments, and her status as a carrier of sickle cell anemia, initially made on the basis of protein electrophoresis, was confirmed.



Q: Testing for cystic fibrosis by DNA sequencing is time-consuming and expensive. Therefore, another technique that uses allele-specific oligonucleotide probes has been developed. **Sissy Fibrosa** and her family were tested by this method. Oligonucleotide probes, complementary to the region where the 3-base deletion is located, have been synthesized. One probe binds to the mutant (ΔF_{508}) gene, and the other to the normal gene.

DNA was isolated from Sissy, her parents, and two siblings and amplified by PCR. Samples of the DNA were spotted on nitrocellulose paper, treated with the oligonucleotide probes, and the following results were obtained. (Dark spots indicate binding of the probe.)



Which members of Sissy's family have CF, which are normal, and which are carriers?

3. TESTING FOR MUTATIONS BY PCR

If an oligonucleotide complementary to a DNA sequence containing a mutation is used as a primer for PCR, a DNA sample used as the template will be amplified only if it contains the mutation. If the DNA is normal, the primer will not hybridize with it, and the DNA will not be amplified. This concept is extremely useful for clinical testing. In fact, a number of oligonucleotides, each specific for a different mutation and each containing a different label, can be used as primers in a single PCR reaction. This procedure results in rapid and relatively inexpensive testing for multiple mutations.

4. DETECTION OF POLYMORPHISMS CAUSED BY REPETITIVE DNA

Human DNA contains many sequences that are repeated in tandem a variable number of times at certain loci in the genome. These regions are called highly variable regions because they contain a variable number of tandem repeats (VNTR). Digestion with restriction enzymes that recognize sites that flank the VNTR region produces fragments containing these loci, which differ in size from one individual to another, depending on the number of repeats that are present. Probes used to identify these restriction fragments bind to or near the sequence that is repeated (Fig. 17.10).

The restriction fragment patterns produced from these loci can be used to identify individuals as accurately as the traditional fingerprint. In fact, this restriction fragment technique has been called "DNA fingerprinting" and is gaining widespread use in forensic analysis. Family relationships can be determined by this method, and it can be used to help acquit or convict suspects in criminal cases.

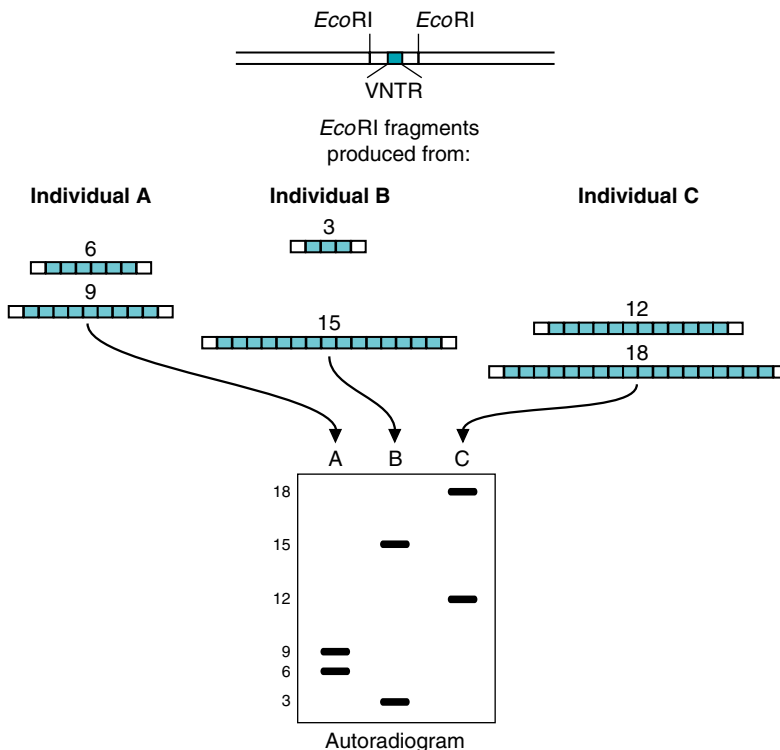
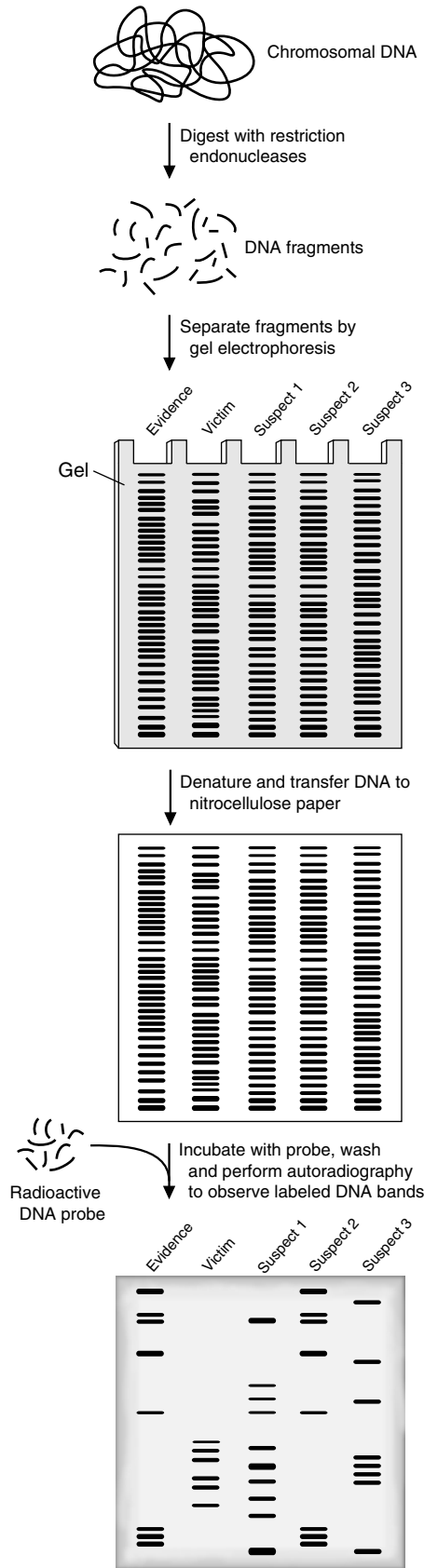


Fig. 17.10. Restriction fragments produced from a gene with a variable number of tandem repeats (VNTR). Each individual has two homologues of every somatic chromosome and thus two genes each containing this region with a VNTR. Cleavage of each individual's genomic DNA with a restriction enzyme produces two fragments containing this region. The length of the fragments depends on the number of repeats they contain. Electrophoresis separates the fragments, and a labeled probe that binds to the fragments allows them to be visualized. Each short blue block represents one repeat.



Obviously, the father and mother are both carriers of the defective allele, as is one of the two siblings (Child 2). Sissy has the disease, and the other sibling (Child 1) is normal.



DNA samples were obtained from each of the three suspects in **Vicky Tim's** rape and murder case, and these samples were compared with the victim's DNA by using DNA fingerprinting. Because **Vicky Tim's** sample size was small, PCR was used to amplify the regions containing the VNTRs. The results, using a probe for one of the repeated sequences in human DNA, are shown below to illustrate the process. For more positive identification, a number of different restriction enzymes and probes were used. The DNA from suspect 2 produced the same restriction pattern as the DNA from the semen obtained from the victim. If the other restriction enzymes and probes corroborate this finding, suspect 2 can be identified by DNA fingerprinting as the rapist/murderer.



The huge amount of information now available from the sequencing of the human genome, and the results available from gene chip experiments, has greatly expanded the field of Bioinformatics. Bioinformatics can be defined as the gathering, processing, data storage, data analysis, information extraction, and visualization of biologic data. Bioinformatics also provides the scientist with the capability of organizing vast amounts of data in a manageable form that allows easy access and retrieval of data. Powerful computers are required to perform these analyses. As an example of an experiment requiring these tools, suppose you want to compare the effects of two different immunosuppressant drugs on gene expression in lymphocytes. Lymphocytes would be treated with either nothing (the control) or with the drugs (experimental samples). RNA would be isolated from the cells during drug treatment, and the RNA converted to fluorescent cDNA using the enzyme reverse transcriptase and a fluorescent nucleotide analog. The cDNA produced from your three samples would be used as probes for a gene chip containing DNA fragments from more than 5,000 human genes. The samples would be allowed to hybridize to the chips, and you would then have 15,000 results to interpret (the extent of hybridization of each cDNA sample with each of the 5,000 genes on the chip). Computers are used to analyze the fluorescent spots on the chips and to compare the levels of fluorescent intensity from one chip to another. In this way, you can group genes showing similar levels of stimulation or inhibition in the presence of the drugs and compare the two drugs with respect to which genes have activated or inhibited expression.



When **Erna Nemdy** began working with patients, she received the hepatitis B vaccine. The hepatitis B virus (HBV) infects the liver, causing severe damage. The virus contains a surface antigen (HBsAg) or coat protein for which the gene has been isolated. However, because the protein is glycosylated, it could not be produced in *E. coli*. (Bacteria, because they lack subcellular organelles, cannot produce glycosylated proteins.) Therefore, a yeast (eukaryotic) expression system was used that produced a glycosylated form of the protein. The viral protein, separated from the small amount of contaminating yeast protein, is used as a vaccine for immunization against HBV infection.

Individuals who are closely related genetically will have restriction fragment patterns (DNA fingerprints) that are more similar than those who are more distantly related. Only monozygotic twins will have identical patterns.

5. DNA CHIPS

A recently introduced technique will permit the screening of many genes to determine which alleles of these genes are present in samples obtained from patients. The surface of a small chip is dotted with thousands of pieces of single-stranded DNA, each representing a different gene or segment of a gene. The chip is then incubated with a sample of a patient's DNA, and the pattern of hybridization is determined by computer analysis. The results of the hybridization analysis could be used, for example, to determine which one of the many known mutations for a particular genetic disease is the specific defect underlying a patient's problem. An individual's gene chip also may be used to determine which alleles of drug-metabolizing enzymes are present and, therefore, the likelihood of that individual having an adverse reaction to a particular drug.

Another use for a DNA chip is to determine which genes are being expressed. If the mRNA from a tissue specimen is used to produce a cDNA by reverse transcriptase, the cDNA will hybridize with only those genes being expressed in that tissue. In the case of a cancer patient, this technique could be used to determine the classification of the cancer much more rapidly and more accurately than the methods traditionally used by pathologists. The treatment then could be more specifically tailored to the individual patient. This technique also can be used to identify the genes required for tissue specificity (e.g., the difference between a muscle cell and a liver cell) and differentiation (the conversion of precursor cells into the different cell types). Experiments using gene chips are helping us to understand differentiation and may open the opportunity to artificially induce differentiation and tissue regeneration in the treatment of disease.

Although development of this DNA or biochip technology is in its infancy, the technique has astonishing potential for the future diagnosis and treatment of disease.

III. USE OF RECOMBINANT DNA TECHNIQUES FOR THE PREVENTION AND TREATMENT OF DISEASE

A. Vaccines

Before the advent of recombinant DNA technology, vaccines were made exclusively from infectious agents that had been either killed or attenuated (altered so that they can no longer multiply in an inoculated individual). Both types of vaccines were potentially dangerous because they could be contaminated with the live, infectious agent. In fact, in a small number of instances, disease has actually been caused by vaccination. The human immune system responds to antigenic proteins on the surface of an infectious agent. By recombinant DNA techniques, these antigenic proteins can be produced, completely free of the infectious agent, and used in a vaccine. Thus, any risk of infection is eliminated. The first successful recombinant DNA vaccine to be produced was for the hepatitis B virus.

B. Production of Therapeutic Proteins

1. INSULIN AND GROWTH HORMONE

Recombinant DNA techniques are used to produce proteins that have therapeutic properties. One of the first such proteins to be produced was human insulin. Recombinant DNA corresponding to the A chain of human insulin was prepared and inserted into plasmids that were used to transform *E. coli* cells. The bacteria then synthesized the insulin chain, which was purified. A similar process was used to obtain B chains. The A and B chains were then mixed and allowed to fold and form disulfide bonds, producing active insulin molecules (Fig. 17.11). Insulin is not glycosylated, so there was no problem with differences in glycosyltransferase activity between *E. coli* and human cell types.

Human growth hormone has also been produced in *E. coli* and is used to treat children with growth hormone deficiencies. Before production of recombinant growth hormone, growth hormone isolated from cadaver pituitary tissue was used, which was in short supply.

2. COMPLEX HUMAN PROTEINS

More complex proteins have been produced in mammalian cell culture using recombinant DNA techniques. The gene for Factor VIII, a protein involved in blood clotting, is defective in individuals with hemophilia. Before genetically engineered Factor VIII became available, a number of hemophiliac patients died of AIDS or hepatitis that they contracted from transfusions of contaminated blood or from Factor VIII isolated from contaminated blood.

Tissue plasminogen activator (TPA) is a protease in blood that converts plasminogen to plasmin. Plasmin is a protease that cleaves fibrin (a major component of blood clots), and, thus, TPA administration dissolves blood clots. Recombinant TPA, produced in mammalian cell cultures, is frequently administered during or immediately after a heart attack to dissolve the thrombi that occlude coronary arteries and prevent oxygen from reaching the heart muscle.

Hematopoietic growth factors also have been produced in mammalian cell cultures by recombinant DNA techniques. Erythropoietin can be used in certain types of anemias to stimulate the production of red blood cells. Colony-stimulating factors (CSFs) and interleukins (ILs) can be used after bone marrow transplants and after chemotherapy to stimulate white blood cell production and decrease the risk of infection.

Recombinant β -interferon is the first drug known to decrease the frequency and severity of episodes resulting from the effects of demyelination in patients with multiple sclerosis.

A method for producing human proteins that is being tested involves transgenic animals. These animals (usually goat or sheep) have been genetically engineered to produce human proteins in the mammary gland and secrete them into milk. The gene of interest is engineered to contain a promoter that is only active in the mammary glands under lactating conditions. The vector containing the gene and promoter is inserted into the nucleus of a freshly fertilized egg, which is then implanted into a foster mother. The female animal progeny are tested for the presence of this transgene, and milk from the positive animals is collected. Large quantities of the protein of interest can then be isolated from the relatively small number of proteins present in milk.

C. Genetic Counseling

One means of preventing disease is to avoid passing defective genes to offspring. If individuals are tested for genetic diseases, particularly in families known to carry a



Di Abietes is using a recombinant human insulin called lispro (Humalog) (see Chapter 6, Fig. 6.13).

Lispro was genetically engineered so that lysine is at position 28 and proline is at position 29 of the B chain (the reverse of their positions in normal human insulin). Di injects a Humalog mixture that contains 75% lispro protamine suspension (intermediate-acting) and 25% lispro solution (rapid-acting). The switch of position of the two amino acids leads to a faster-acting insulin homolog. The lispro is absorbed from the site of injection much more quickly than other forms of insulin, and it acts to lower blood glucose levels much more rapidly than the other insulin forms.

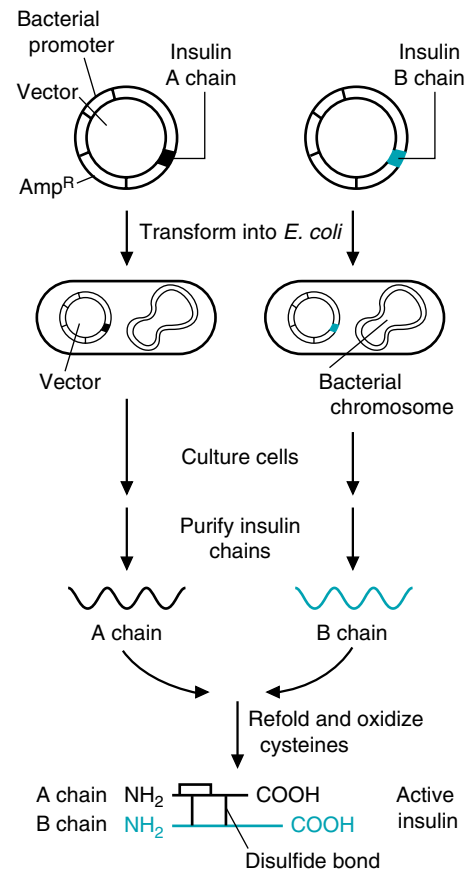


Fig. 17.11. Production of human insulin in *E. coli*. Amp^R = the gene for ampicillin resistance. The presence of Amp^R allows bacterial cells that contain the vector to grow in the presence of ampicillin. Cells lacking the Amp^R gene die in the presence of ampicillin. Because *E. coli* cannot process preproinsulin, a synthetic scheme was developed whereby each individual chain of insulin was expressed, produced, and purified, and then the two chains were linked together in a test tube.

Carrie Sichel's fiancé decided to be tested for the sickle cell gene. He was found to have both the 1.3-kb and the 1.1-kb *Mst*II restriction fragments that include a portion of the β -globin gene. Therefore, like Carrie, he also is a carrier for the sickle cell gene.

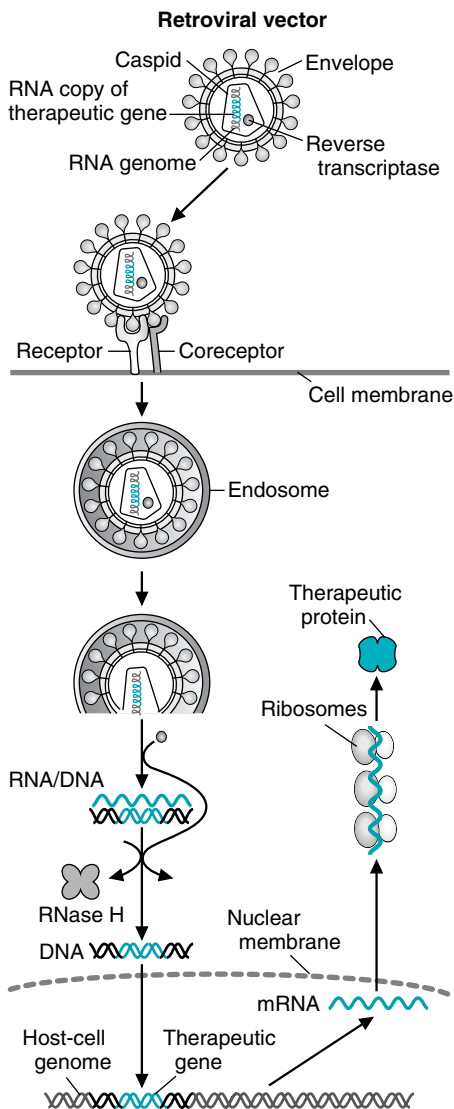


Fig. 17.12. Use of retroviruses for gene therapy. The retrovirus carries an RNA copy of the therapeutic gene into the cell. The endosome that contains the virus dissolves, and the RNA and viral reverse transcriptase are released. This enzyme copies the RNA, making a double-stranded DNA that integrates into the host cell genome. Transcription and translation of this DNA (the therapeutic gene) produces the therapeutic protein. (The virus does not multiply because its genes were removed and replaced by the RNA copy of the therapeutic gene.)

defective gene, genetic counselors can inform them of their risks and options. With this information, individuals can decide in advance whether to have children.

Screening tests, based on the recombinant DNA techniques outlined in this chapter, have been developed for many inherited diseases. Although these tests are currently rather expensive, particularly if entire families have to be screened, the cost may be trivial compared with the burden of raising children with severe disabilities. Obviously, ethical considerations must be taken into account, but recombinant DNA technology has provided individuals with the opportunity to make choices.

Screening can be performed on the prospective parents before conception. If they decide to conceive, the fetus can be tested for the genetic defect. In some cases, if the fetus has the defect, treatment can be instituted at an early stage, even in utero. For certain diseases, early therapy leads to a more positive outcome.

D. Gene Therapy

The ultimate cure for genetic diseases is to introduce normal genes into individuals who have defective genes. Currently, gene therapy is being attempted in animals, cell cultures, and human subjects. It is not possible at present to replace a defective gene with a normal gene at its usual location in the genome of the appropriate cells. However, as long as the gene is expressed at the appropriate time and produces adequate amounts of the protein to return the person to a normal state, the gene does not have to integrate into the precise place in the genome. Sometimes the gene does not even have to be in the cells that normally contain it.

Retroviruses were the first vectors used to introduce genes into human cells. Normally, retroviruses enter target cells, their RNA genome is copied by reverse transcriptase, and the double-stranded DNA copy is integrated into the host cell genome (see Fig. 14.22). If the retroviral genes (e.g., *gag*, *pol*, and *env*) are first removed and replaced with the therapeutic gene, the retroviral genes integrated into the host cell genome will produce the therapeutic protein rather than the viral proteins (Fig. 17.12). This process works only when the human host cells are undergoing division, so it has limited applicability. Other problems with this technique are that it can only be used with small genes (=8 kb), and it may disrupt other genes because the insertion point is random, thereby possibly resulting in cancer.



A defect in the adenosine deaminase (ADA) gene causes severe combined immunodeficiency syndrome (SCID). When ADA is defective, deoxyadenosine and dATP accumulate in rapidly dividing cells, such as lymphocytes, and prove toxic to these cells. Cells of the immune system cannot proliferate at a normal rate, and children with SCID usually die at an early age because they cannot combat infections. To survive, they must be confined to a sterile, environmental "bubble." When an appropriate donor is available, bone marrow transplantation can be performed with a reasonable degree of success.

In 1990, a 4-year-old girl, for whom no donor was available, was treated with infusions of her own lymphocytes that had been treated with a retrovirus containing a normal ADA gene. Although she had not responded to previous therapy, she improved significantly after this attempt at gene therapy. This disease is still being treated with gene therapy, in combination with replacement enzyme infusion.



In familial hypercholesterolemia, a condition associated with a high incidence of heart attacks, the low-density lipoprotein (LDL) receptor is deficient. Attempts to correct this defect with gene therapy involved removal of a segment of liver and preparation of hepatocytes that had been grown in tissue culture. After these dividing cells were infected with a retrovirus containing the gene for the LDL receptor, they were reinfused into the hepatic portal vein of the patient. The early efforts using this approach met with only limited success.

Adenoviruses, which are natural human pathogens, can also be used as vectors. As in retroviral gene therapy, the normal viral genes required for synthesis of viral particles are replaced with the therapeutic genes. The advantages to using an adenovirus are that the introduced gene can be quite large (~36 kb), and infection does not require division of host cells. The disadvantage is that genes carried by the adenovirus do not stably integrate into the host genome, resulting in only transient expression of the therapeutic proteins (but preventing disruption of host genes). Thus, the treatment must be repeated periodically. Another problem with adenoviral gene therapy is that the host can mount an immune response to the pathogenic adenovirus, causing complications.

To avoid the problems associated with viral vectors, researchers are employing treatment with DNA alone or with DNA coated with a layer of lipid (i.e., in liposomes). Adding a ligand for a receptor located on the target cells could aid delivery of the liposomes to the appropriate host cells. Many problems still plague the field of gene therapy. In many instances, the therapeutic genes must be targeted to the cells where they normally function—a difficult task at present. Deficiencies in dominant genes are more difficult to treat than those in recessive genes, and the expression of the therapeutic genes often needs to be carefully regulated. Although the field is moving forward, progress is slow.

Another approach to gene therapy involves the use of antisense oligonucleotides rather than vectors. These oligonucleotides are designed to hybridize either with the target gene to prevent transcription or with mRNA to prevent translation. Again technical problems have plagued the development of therapy based on this theoretically promising idea.

E. Transgenic Animals

The introduction of normal genes into somatic cells with defective genes corrects the defect only in the treated individuals, not in their offspring. To eliminate the defect for future generations, the normal genes must be introduced into the germ cell line (the cells that produce sperm in males or eggs in females). Experiments with animals indicate that gene therapy in germ cells is feasible. Genes can be introduced into fertilized eggs from which transgenic animals develop, and these transgenic animals can produce apparently normal offspring.

In fact, if the nucleus isolated from the cell of one animal is injected into the enucleated egg from another animal of the same species and the egg is implanted in a foster mother, the resulting offspring is a “clone” of the animal from which the nucleus was derived (Fig. 17.13). Clones of sheep and pigs have been produced, and similar techniques could be used to clone humans. Obviously, these experiments raise many ethical questions that will be difficult to answer.



Adenoviral vectors have been used in a aerosol spray to deliver normal copies of the CFTR (cystic fibrosis transmembrane conductance regulator) gene to cells of the lung. Some cells took up this gene, and the patients experienced moderate improvement. However, stable integration of the gene into the genome did not occur, and cells affected by the disease other than those in the lung (e.g., pancreatic cells) did not benefit. Nevertheless, this approach marked a significant forward step in the development of gene therapy.



In an attempt to treat ornithine transcarbamoylase deficiency (a disorder of nitrogen metabolism) using adenoviral vectors, a volunteer died of a severe immune response to the vector. This unfortunate result has led to a reevaluation of the safety of viral vectors for gene therapy.



One approach to in vivo gene therapy involves the direct injection of DNA for certain HLA antigens into malignant melanomas (skin cancers). The HLA gene chosen for therapy should not be the one expressed by the patient. Thus, if the gene is incorporated into the cells and expressed, the body should recognize the tumor cells as foreign tissue and mount an immune attack. Preliminary results using this strategy were encouraging.

CLINICAL COMMENTS



Erna Nemdy. In reading about development of the hepatitis B vaccine, Erna Nemdy learned that the first vaccine available for HBV, marketed in 1982, was a purified and “inactivated” vaccine containing HBV virus that had been chemically killed. The virus was derived from the blood of known HBV carriers. Later, “attenuated” vaccines were used in which the virus remained live but was altered so that it no longer multiplied in the inoculated host. Both the inactivated and the attenuated vaccines are potentially dangerous because they can be contaminated with live infectious HBV.

The modern “subunit” vaccines, first marketed in 1987, were made by recombinant DNA techniques described earlier in this chapter. Because this vaccine consists solely of the viral surface protein or antigen to which the immune system responds, there is no risk for infection with HBV.

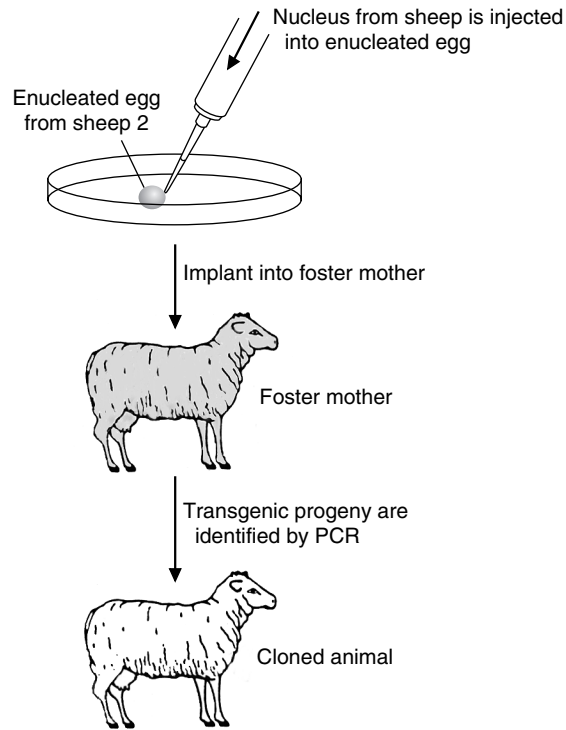


Fig. 17.13. Cloning of a mammalian organism.



The most common CF mutation is a 3-bp deletion that causes the loss of phenylalanine at position 508 (delta 508). This mutation is present in more than 70% of CFTR patients. The defective protein is synthesized in the endoplasmic reticulum, but is misfolded. It is therefore not transported to the Golgi, but is degraded by a proteolytic enzyme complex called the proteasome. Other mutations responsible for CF generate an incomplete mRNA because of premature stop signals, frame shifts, or abnormal splice sites or create a CFTR channel in the membrane that does not function properly.



Sissy Fibrosa. Cystic fibrosis (CF) is a genetically determined autosomal recessive disease that can be caused by a variety of mutations within the CF gene located on chromosome 7. **Sissy Fibrosa** was found to have a 3-bp deletion at residue 508 of the CF gene (the mutation present in approximately 70% of Caucasian patients with CF in the United States). This mutation is generally associated with a more severe clinical course than many other mutations causing the disease. However, other genes and environmental factors may modify the clinical course of the disease, so it is not currently possible to counsel patients accurately about prognosis based on their genotype.

CF is a relatively common genetic disorder in the United States, with a carrier rate of approximately 5% in Caucasians. The disease occurs in 1 per 1,600 to 2,000 Caucasian births in the country (1 per 17,000 in African Americans and 1 per 100,000 in Asians).



Carrie Sichel. After learning the results of their tests for the sickle cell gene, **Carrie Sichel** and her fiancé consulted a genetic counselor. The counselor informed them that, because they were both carriers of the sickle cell gene, their chance of having a child with sickle cell anemia was fairly high (approximately 1 in 4). She told them that prenatal testing was available with fetal DNA obtained from cells by amniocentesis or chorionic villus sampling. If these tests indicated that the fetus had sickle cell disease, abortion was a possibility. Carrie, because of her religious background, was not sure that abortion was an option for her. But having witnessed her brother's sickle cell crises for many years, she also was not sure that she wanted to risk having a child with the disease. Her fiancé also felt that, at 25 years of age, he was not ready to deal with such difficult problems. They mutually agreed to cancel their marriage plans.



Vicky Tim. DNA fingerprinting represents an important advance in forensic medicine. Before development of this technique, identification of criminals was far less scientific. The suspect in the rape and murder of **Vicky Tim** was arrested and convicted mainly on the basis of the results of DNA fingerprint analysis.

This technique has been challenged in some courts on the basis of technical problems in statistical interpretation of the data and sample collection. It is absolutely necessary for all of the appropriate controls to be run, including samples from the victim's DNA as well as the suspect's DNA. Another challenge to the fingerprinting procedure has been raised because PCR is such a powerful technique that it can amplify minute amounts of contaminating DNA from a source unrelated to the case.

BIOCHEMICAL COMMENTS



Mapping of the Human Genome. The Human Genome Project began in 1990, and by the summer of 2000, the entire human genome had been mapped. This feat was accomplished in far less than the expected time as a result of both cooperative and competitive interactions of laboratories in the private as well as the public sector.

The human genome contains over 3×10^9 (3 billion) bp. A large percentage of this genome (<95%) does not code for the amino acid sequences of proteins or for functional RNA (such as rRNA or tRNA) but is composed of repetitive sequences, introns, and other noncoding elements of unknown function. The human genome is estimated to contain only about 30,000 to 50,000 genes.

As the announcement of the identification of a wayward gene appears in the morning newspaper, the average citizen expects the cure for the genetic disease to be described in the evening edition. Although knowledge of the chromosomal location and the sequence of genes will result in the rapid development of tests to determine whether an individual carries a defective gene, the development of a treatment for the genetic disease caused by the defective gene will not be that easy or that rapid. As outlined in the section on gene therapy above, many technical problems need to be solved before gene therapy becomes commonplace. In addition to solving the molecular puzzles involved in gene therapy, we also will have to deal with many difficult questions.

Is it appropriate to replace defective genes in somatic cells to relieve human suffering? Many people may agree with this goal. But there is a related question: is it appropriate to replace defective genes **in the germ cell line** to relieve human suffering? Fewer people may agree with this goal. Genetic manipulation of somatic cells affects only one generation; these cells die with the individual. Germ cells, however, live on, producing each successive generation.

The techniques developed to explore the human genome could be used for many purposes. What are the limits for the application of the knowledge gained by advances in molecular biology? Who should decide what the limits are, and who should serve as the genetic police? If we permit experiments that involve genetic manipulation of the human germ cell line, however nobly conceived, could we, in our efforts to "improve" ourselves, genetically engineer the human race into extinction?



What are the statistical issues relating to DNA fingerprinting? Through the analysis of 1,000 individuals of different ethnicities, one can determine the frequency of a particular DNA polymorphism within that distinct population. By matching four or five polymorphisms from DNA at the crime scene with DNA from a suspect, one can determine the odds of that match happening by chance. For example, let us assume that a suspect's DNA was compared with DNA found at the crime scene for four unique polymorphisms within the suspect's ethnic group. The frequency of polymorphism A in that population is 1 in 20; of polymorphism B, 1 in 30; of polymorphism C, 1 in 50; and of polymorphism D, 1 in 100. The odds of the suspect DNA matching the DNA found at the crime scene for all four polymorphisms would be the product of each individual probability, or $(1/20) \times (1/30) \times (1/50) \times (1/100)$. This comes out to a 1 in 3 million chance that an individual would have the same polymorphisms in their DNA as that found at the crime scene. The question left to the courts is whether the 1 in 3 million match is sufficient to convict the suspect of the crime. Given that there may be 30 million individuals in the United States within the same ethnic group as the suspect, there would then be 30 people within the country who would match the DNA polymorphisms found at the scene of the crime. Can the court be sure that the suspect is the correct individual? Clearly, the use of DNA fingerprinting is much clearer when a match is not made, for that immediately indicates that the suspect was not at the scene of the crime.

Suggested References

- Housman D. DNA on trial—the molecular basis of DNA fingerprinting. *N Engl J Med* 1995;332:534–535.
- Trent RJ. *Molecular Medicine*. New York: Churchill Livingstone, 1993.

- Watson JD, Gilman M, Witkowski J, Zoller M. Recombinant DNA. 2nd Ed. New York: Scientific American Books, WH Freeman, 1992.
- Gelehrter TD, Collins FS, Ginsburg D. Principles of Medical Genetics. 2nd Ed. Baltimore: Williams & Wilkins, 1998.
- Welsh MJ. Cystic Fibrosis. Chapter 76. In: Goldman L, Bennett JC, eds. Cecil Textbook of Medicine. 21st Ed, Volume 1, pp 401–405. Philadelphia: WB Saunders, 2000.



REVIEW QUESTIONS—CHAPTER 17

- Electrophoresis resolves double-stranded DNA fragments based on which of the following?
 - Sequence
 - Molecular weight
 - Isoelectric point
 - Frequency of CTG repeats
 - Secondary structure
- If a restriction enzyme recognizes a six-base sequence, how frequently, on average, will this enzyme cut a large piece of DNA?
 - Once every 16 bases
 - Once every 64 bases
 - Once every 256 bases
 - Once every 1,024 bases
 - Once every 4,096 bases
- Which of the following sets of reagents are required for dideoxy chain DNA synthesis in the Sanger technique for DNA sequencing?
 - Deoxyribonucleotides, Taq polymerase, DNA primer
 - Dideoxyribonucleotides, deoxyribonucleotides, template DNA
 - Dideoxyribonucleotides, DNA primer, reverse transcriptase
 - Two DNA primers, template DNA, Taq polymerase
 - mRNA, dideoxynucleotides, reverse transcriptase
- Which of the following statements correctly describe a feature of DNA electrophoresis?
 - Larger DNA fragments migrate farther in the gel.
 - DNA fragments migrate toward the negative charge (anode).
 - DNA can be visualized using UV light and the dye ethidium bromide.
 - Total human genomic DNA cut by a specific restriction endonuclease will generate three distinctly separable bands.
 - DNA must be denatured before it can be run in the gel.
- The best method to determine whether albumin is transcribed in the liver of a mouse model of hepatocarcinoma is which of the following?
 - Genomic library screening
 - Genomic Southern blot
 - Tissue Northern blot
 - Tissue Western blot
 - VNTR analysis