Basic Methods in Cellular and Molecular Biology

(DNA)

ANALYZING AND MANIPULATING DNA

- Until the early 1970s, DNA was the most difficult biological molecule for the biochemist to analyze.

- Enormously long and chemically monotonous, the string of nucleotides that forms the genetic material of an organism could be examined only indirectly, by protein sequencing or by genetic analysis.

- Today, the situation has changed entirely. From being the most difficult macromolecule of the cell to analyze, DNA has become the easiest.

- It is now possible to determine the entire nucleotide sequence of a bacterial or fungal genome in a matter of hours, and the sequence of an individual human genome in less than a day.

- Once the nucleotide sequence of a genome is known, any individual gene can be easily isolated, and large quantities of the gene product (be it RNA or protein) can be made either by introducing the gene into bacteria or animal cells and coaxing these cells to **overexpress the foreign gene** or by **synthesizing the gene product in vitro**. In this way, proteins and RNA molecules that might be present in only tiny amounts in living cells can be produced in large quantities for biochemical and structural analyses.

- And this approach can also be used to produce large quantities of human proteins (such as insulin, or interferon, or blood-clotting proteins) for use as human pharmaceuticals.

- It is also possible for scientists to **alter** an isolated gene and transfer it back into the germ line of an animal or plant, so as to become a functional and heritable part of the organism's genome.

- In this way, the biological roles of any gene can be assessed by observing—in the whole organism—the results of modifying it.

ANALYZING AND MANIPULATING DNA

The ability to manipulate DNA with precision in a test tube or an organism, known as **recombinant DNA technology** has had a dramatic impact on all aspects of cell and molecular biology, allowing us to routinely study cells and their macromolecules in ways that were unimaginable even twenty years ago.

Central to the technology are the following manipulations:

1. Cleavage of DNA at specific sites by **restriction nucleases**, which greatly facilitates the isolation and manipulation of individual pieces of a genome.

2. **DNA ligation**, which makes it possible to seamlessly join together DNA molecules from widely different sources.

3. **DNA cloning** (through the use of either cloning vectors or the polymerase chain reaction) in which a portion of a genome (often an individual gene) is "purified" away from the remainder of the genome by repeatedly copying it to generate many billions of identical molecules.

4. **Nucleic acid hybridization**, which makes it possible to identify any specific sequence of DNA or RNA with great accuracy and sensitivity based on its ability to selectively bind a complementary nucleic acid sequence.

5. **DNA synthesis**, which makes it possible to **chemically** synthesize DNA molecules with any sequence of nucleotides, whether or not the sequence occurs in nature.

6. Rapid determination of the **sequence of nucleotides** of any DNA or RNA molecule.

Restriction Nucleases Cut Large DNA Molecules into Specific Fragments

- Unlike a protein, a gene does not exist as a discrete entity in cells, but rather as a small region of a much longer DNA molecule.

- Although the DNA molecules in a cell can be **randomly** broken into small pieces by mechanical force, a fragment containing a single gene in a mammalian genome would still be only one among a hundred thousand or more DNA fragments, indistinguishable in their average size.

How could such a gene be separated from all the others?

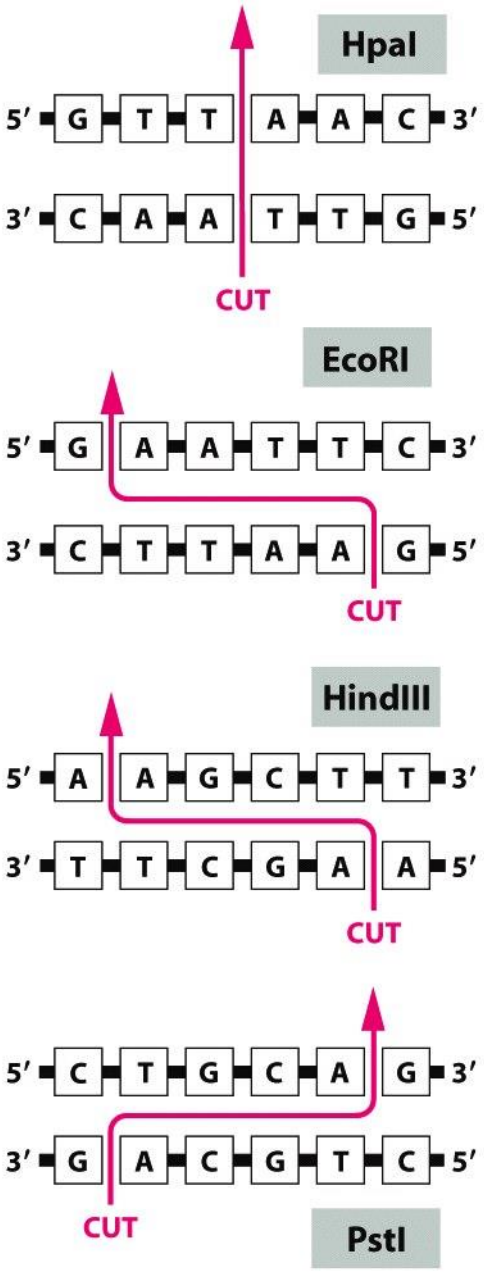
- Because all DNA molecules consist of an **approximately equal mixture of the same four nucleotides**, they cannot be readily separated, as proteins can, on the basis of their different charges and biochemical properties.

- The solution to this problem began to emerge with the discovery of **restriction nucleases**.

- These enzymes, which are purified from bacteria, cut the **DNA double helix at specific sites defined by the local nucleotide sequence**, thereby cleaving a long, double-stranded DNA molecule into fragments of strictly **defined sizes**.

Restriction Nucleases Cut Large DNA Molecules into Specific Fragments

- Like many of the tools of recombinant DNA technology, restriction nucleases were discovered by researchers trying to understand an intriguing biological phenomenon.
- It had been observed that certain bacteria always degraded "**foreign**" DNA that was introduced into them experimentally.
- A search for the mechanism responsible revealed a then unanticipated class of **bacterial nucleases** that cleave DNA at specific nucleotide sequences.
- The bacterium's own DNA is protected from cleavage by **methylation** of these same sequences, thereby protecting a bacterium's own genome from being overrun by foreign DNA.
- Because these enzymes restrict the transfer of DNA into bacteria, they were called **restriction nucleases**.
- Different bacterial species produce different restriction nucleases, each cutting at a different, specific nucleotide sequence.



Restriction Nucleases Cut Large DNA Molecules into Specific Fragments

Because these target sequences are short—generally **four to eight nucleotide pairs**—many sites of cleavage will occur, purely by chance, in any long DNA molecule.

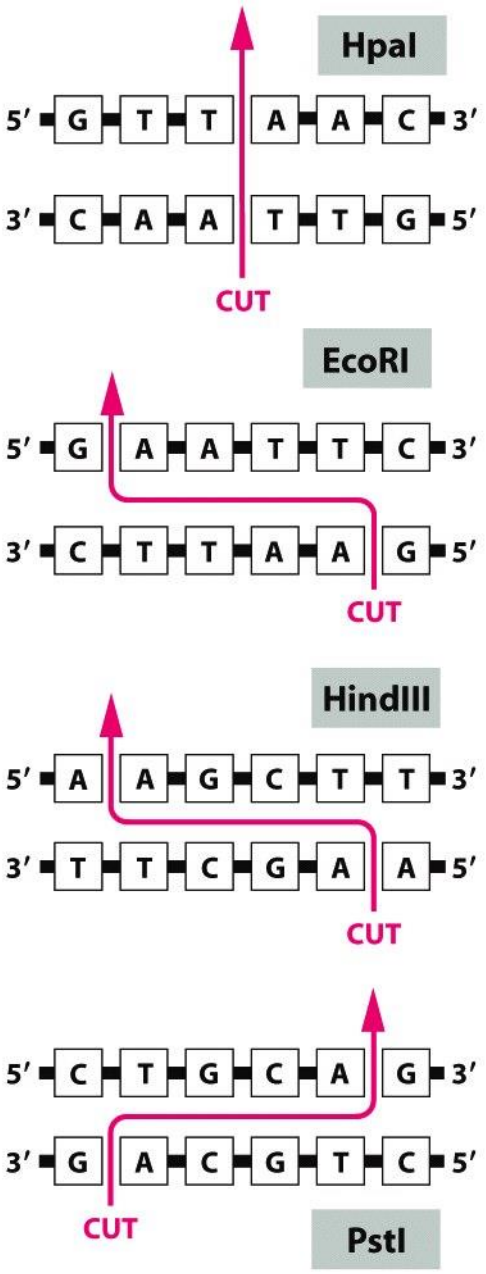
The reason restriction nucleases are so useful in the laboratory is that each enzyme will always cut a particular DNA molecule at the same sites. Thus for a given sample of DNA (which contains many identical molecules), a particular restriction nuclease will reliably generate the **same set of DNA fragments.**

The size of the resulting fragments depends on the length of the target sequences of the restriction nucleases.

The enzyme HaeIII cuts at a sequence of four nucleotide pairs; a sequence this long would be expected to occur purely by chance approximately once every 256 nucleotide pairs (1 in 44).

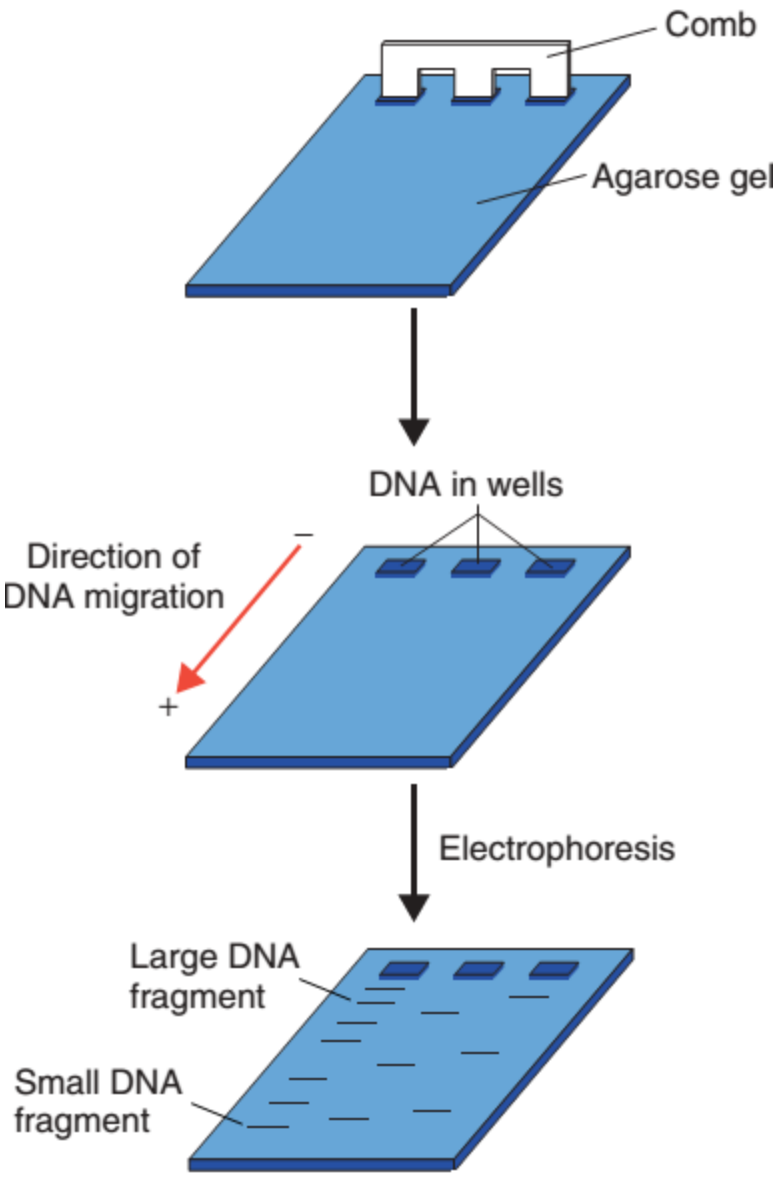
In comparison, a restriction nuclease with a target sequence that is eight nucleotides long would be expected to cleave DNA on average once every 65,536 nucleotide pairs (1 in 48).

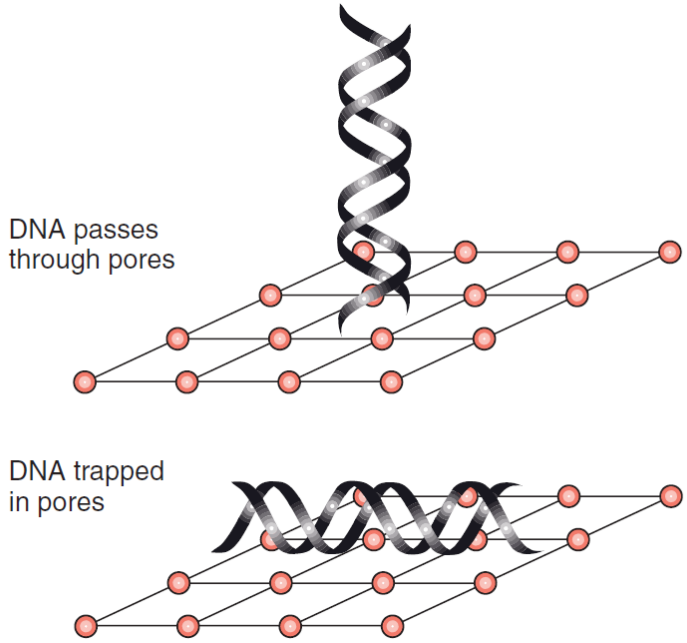
This difference in **sequence selectivity** makes it possible to cleave a long DNA molecule into the fragment sizes that are most suitable for a given application.



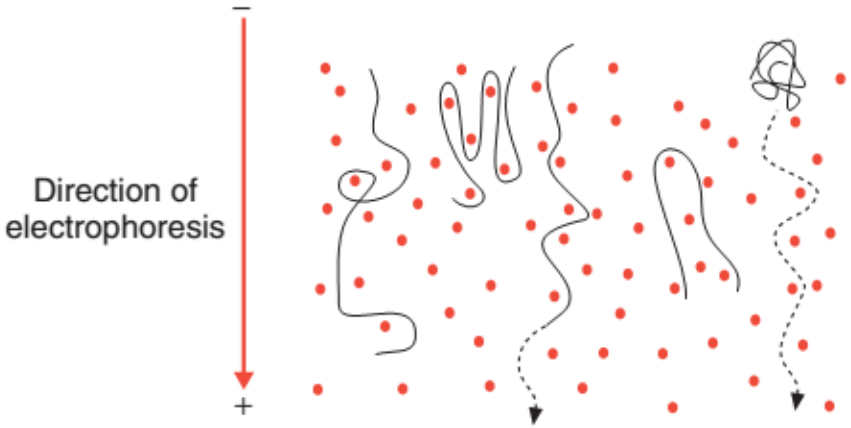
Gel Electrophoresis Separates DNA Molecules of Different Sizes

- The same types of gel-electrophoresis methods that have proved so useful in the analysis of proteins can be applied to DNA molecules.
- The procedure is actually simpler than for proteins: because each nucleotide in a nucleic acid molecule already carries **a single negative charge** (on the phosphate group), there is no need to add the negatively charged detergent SDS that is required to make protein molecules move uniformly toward the positive electrode.
- **Larger DNA fragments will migrate more slowly because their progress is impeded to a greater extent by the gel matrix.**
- Over several hours, the DNA fragments become spread out across the gel **according to size**, forming a ladder of discrete bands, each composed of a collection of DNA molecules of identical length.
- To separate DNA molecules longer than 500 nucleotide pairs, the gel is made of a diluted solution of **agarose** (a polysaccharide isolated from seaweed).





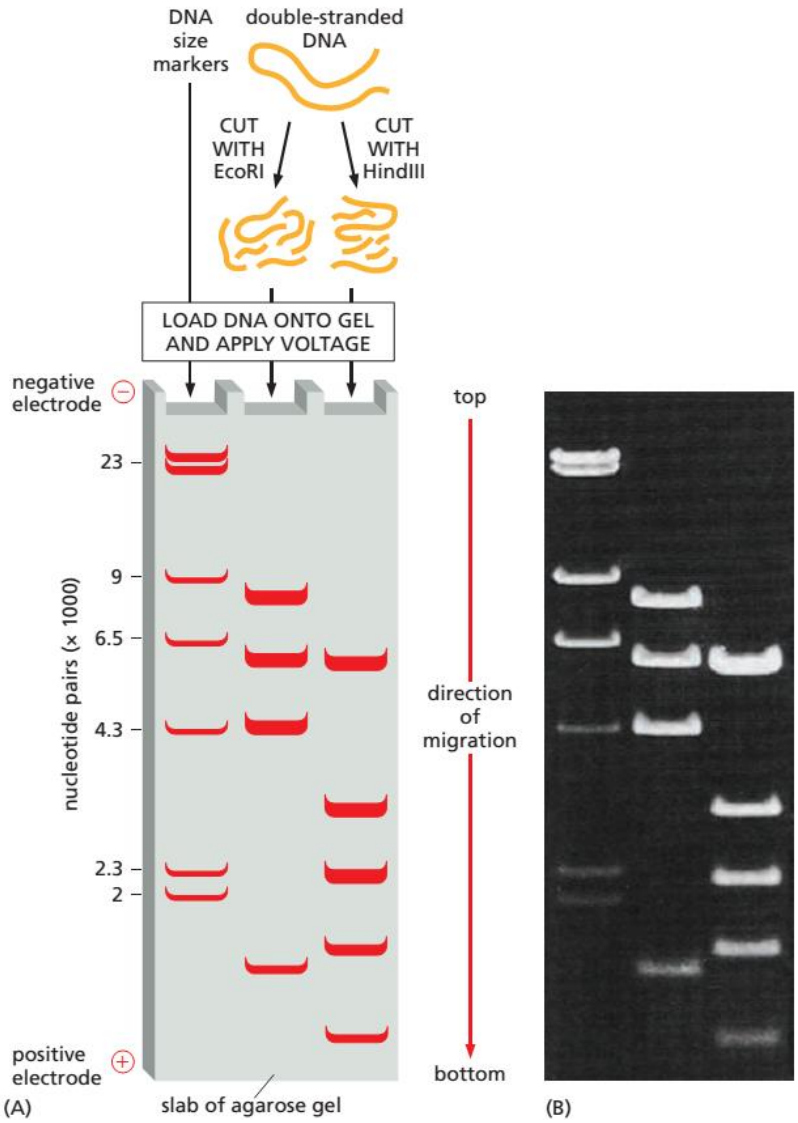
DNA is thought to travel through the pores of a gel in an end-on fashion. If we think of an agarose gel as a meshed network of pores, then we can imagine DNA can more readily pass through the pores if it travels end-on rather than side-on. This end-on movement is sometimes referred to as snaking or reputation.

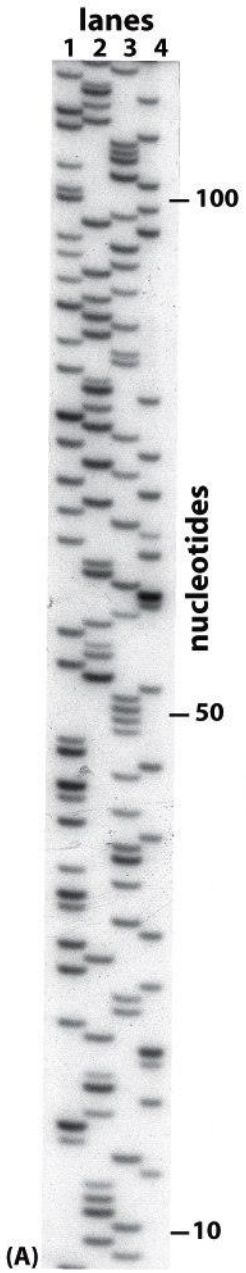


DNA snaking through the pores of a gel. DNA molecules moving through the pores of a gel may become trapped in a variety of ways. Larger DNA molecules are more likely to be trapped due to their length than smaller ones. This may be the reason that DNA molecules larger than about 30 kbp all run in about the same place in a conventional agarose gel

Gel Electrophoresis Separates DNA Molecules of Different Sizes

For DNA fragments less than 500 nucleotides long, specially designed **polyacrylamide gels** allow the separation of molecules that differ in length by as little as a single nucleotide.





Gel Electrophoresis Separates DNA Molecules of Different Sizes

- A variation of agarose-gel electrophoresis, called **pulsed-field gel electrophoresis,** makes it possible to separate extremely long DNA molecules, even those found in whole chromosomes.

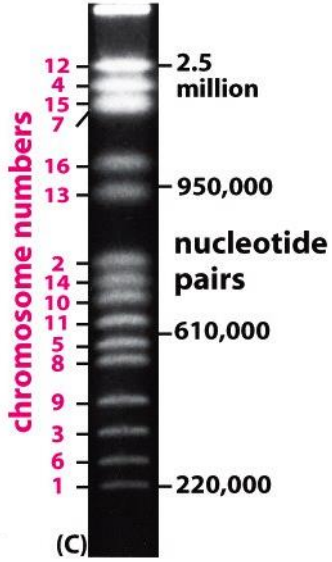
- Ordinary gel electrophoresis fails to separate very large DNA molecules because the steady electric field stretches them out so that they travel **end-first** through the gel in snakelike configurations at a rate that is independent of their length.

- In pulsed-field gel electrophoresis, by contrast, the direction of the electric field changes periodically, which forces the molecules to reorient before continuing to move snakelike through the gel.

- **This re-orientation takes much more time for larger molecules, so that longer molecules move more slowly than shorter ones.**

- As a consequence, entire bacterial or yeast chromosomes separate into discrete bands in pulsed-field gels and so can be sorted and identified on the basis of their size.

- Although a typical mammalian chromosome of 10^8 nucleotide pairs is still too long to be sorted even in this way, large segments of these chromosomes are readily separated and identified if the chromosomal DNA is first cut with a restriction nuclease selected to recognize sequences that occur only rarely.

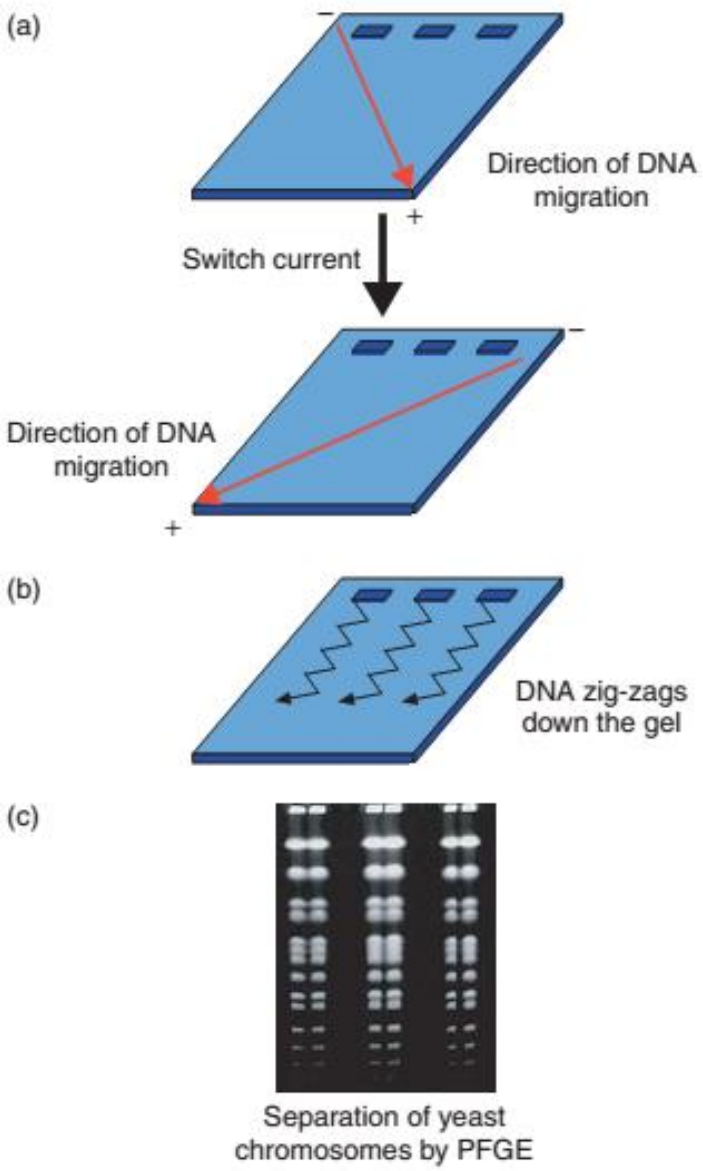


Pulsed-field gel electrophoresis (PFGE)

(a) The switching of the electric current during PFGE. Current is applied across the gel for a defined period – the pulse time – which is often in the range of 0.5–2 min. After this time the direction of the current is switched.

(b) The repetitive switching of the current means that the DNA will **zig-zag** down the gel. The original PFGE technique used two non-homogenous electric fields to change the direction of DNA migration during electrophoresis (Schwartz and Cantor, 1984).
The zig-zagging motion allowed the separation of large DNA molecules, but the non-homogenous electric field resulted in **bowed DNA banding patterns**.

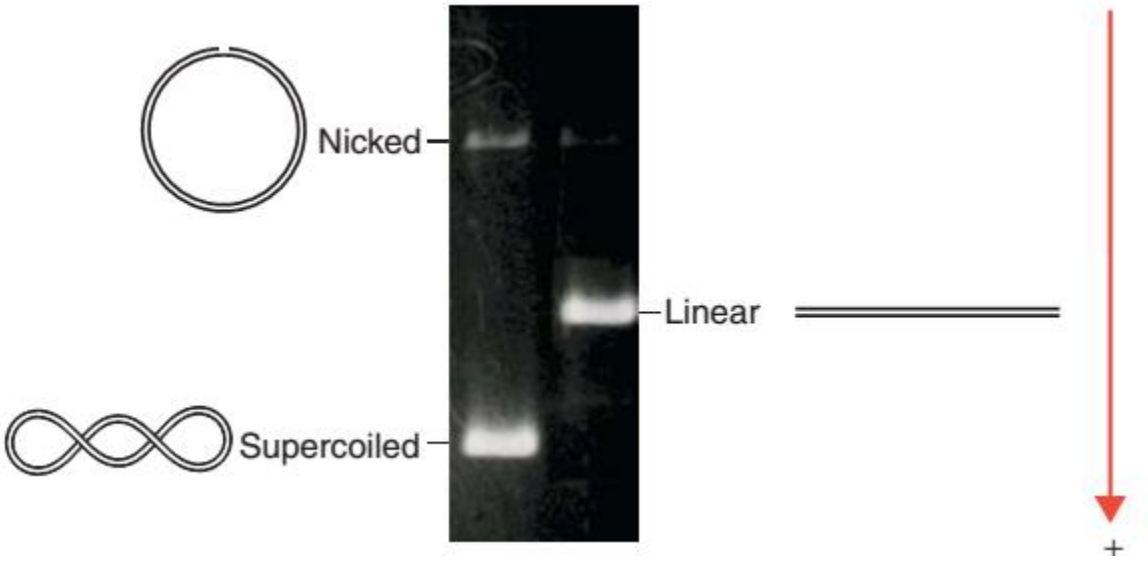
(c) Using homogenous electric fields, straight-line separation patterns are obtained, like the separation of whole yeast chromosomes shown here.



- The second major factor influencing **migration** through a gel is the **topology or structure of a particular DNA fragment**:

✓ For instance, plasmid DNA isolated from E. coli cells is invariably negatively **supercoiled** closed-circular molecules. These are relatively compact structures that run quickly through agarose gels.
✓ If one strand of the plasmid double helix becomes broken (**nicked**) then the supercoiling within the plasmid will be lost, and the more open structure of the relaxed plasmid will migrate more slowly through an agarose gel.
✓ If the same plasmid is treated with a restriction enzyme that cleaves it once, then this **linearized DNA** will run with a mobility intermediate between those of the supercoiled and the nicked molecules.

Therefore, DNA molecules that all contain precisely the same number of base pairs can run in several different locations on an agarose gel depending upon the topology of the DNA.

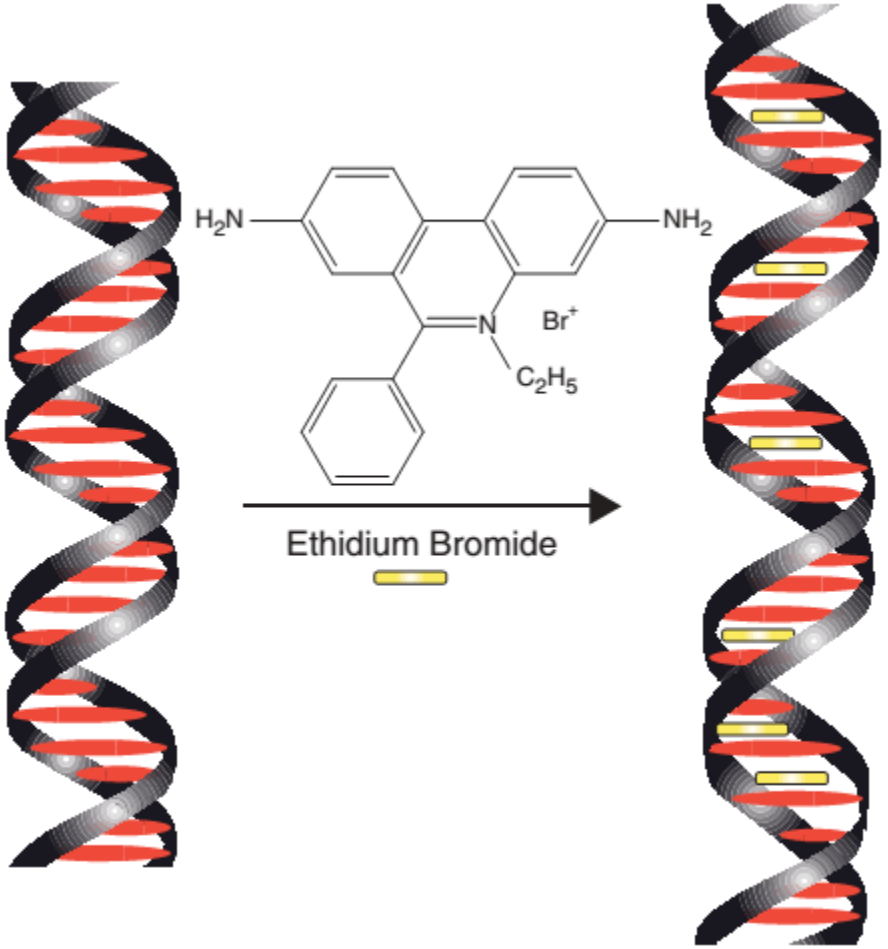


Gel Electrophoresis Separates DNA Molecules of Different Sizes

The DNA bands on agarose or polyacrylamide gels are invisible unless the DNA is labeled or stained in some way.

A particularly sensitive method of staining DNA is to soak the gel in the dye **ethidium bromide**, which fluoresces under **ultraviolet light** when it is bound to DNA.

Even more sensitive detection methods incorporate a **radioisotope** or **chemical marker** into the DNA molecules before electrophoresis.



The binding of ethidium bromide to DNA.

Ethidium bromide is a flat, planar molecule that is able to intercalate in between the stacked bases of double stranded DNA.

The binding of ethidium bromide distorts the double helix and increases its overall length.

DNA to which ethidium bromide is bound fluoresces when viewed under ultraviolet light.

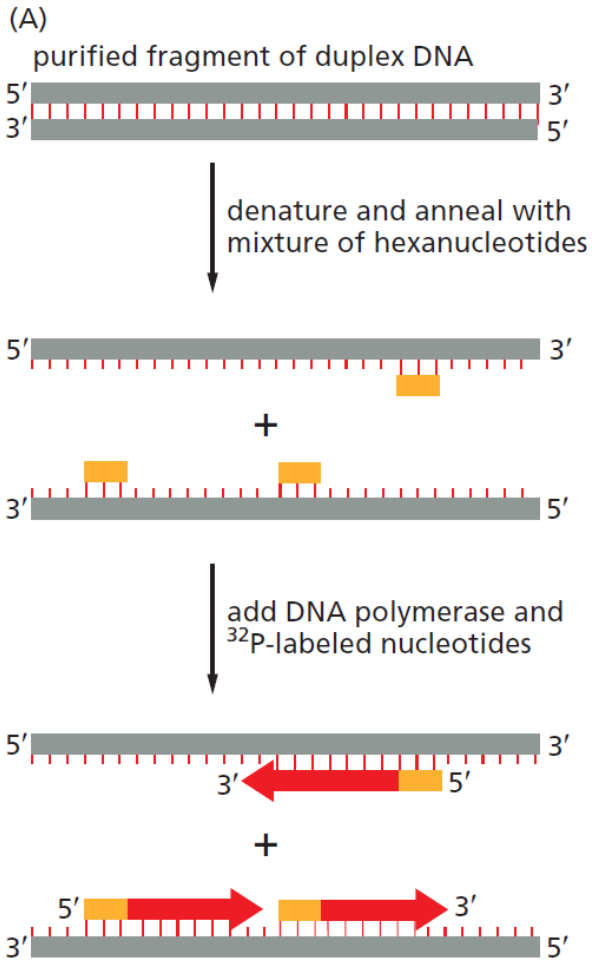
Purified DNA Molecules Can Be Specifically Labeled with Radioisotopes or Chemical Markers in vitro

- The **DNA polymerases** that synthesize and repair DNA have become important tools in experimentally manipulating DNA.
- Because they synthesize sequences complementary to an existing DNA molecule, they are often used in the test tube to create exact copies of existing DNA molecules.
- **The copies can include specially modified nucleotides**.

- To synthesize DNA in this way, the DNA polymerase is presented with a **template** and a **pool** of **nucleotide precursors** that contain the modification.
- As long as the polymerase can use these precursors, it automatically makes new, modified molecules that match the sequence of the template.

- Modified DNA molecules have many uses.

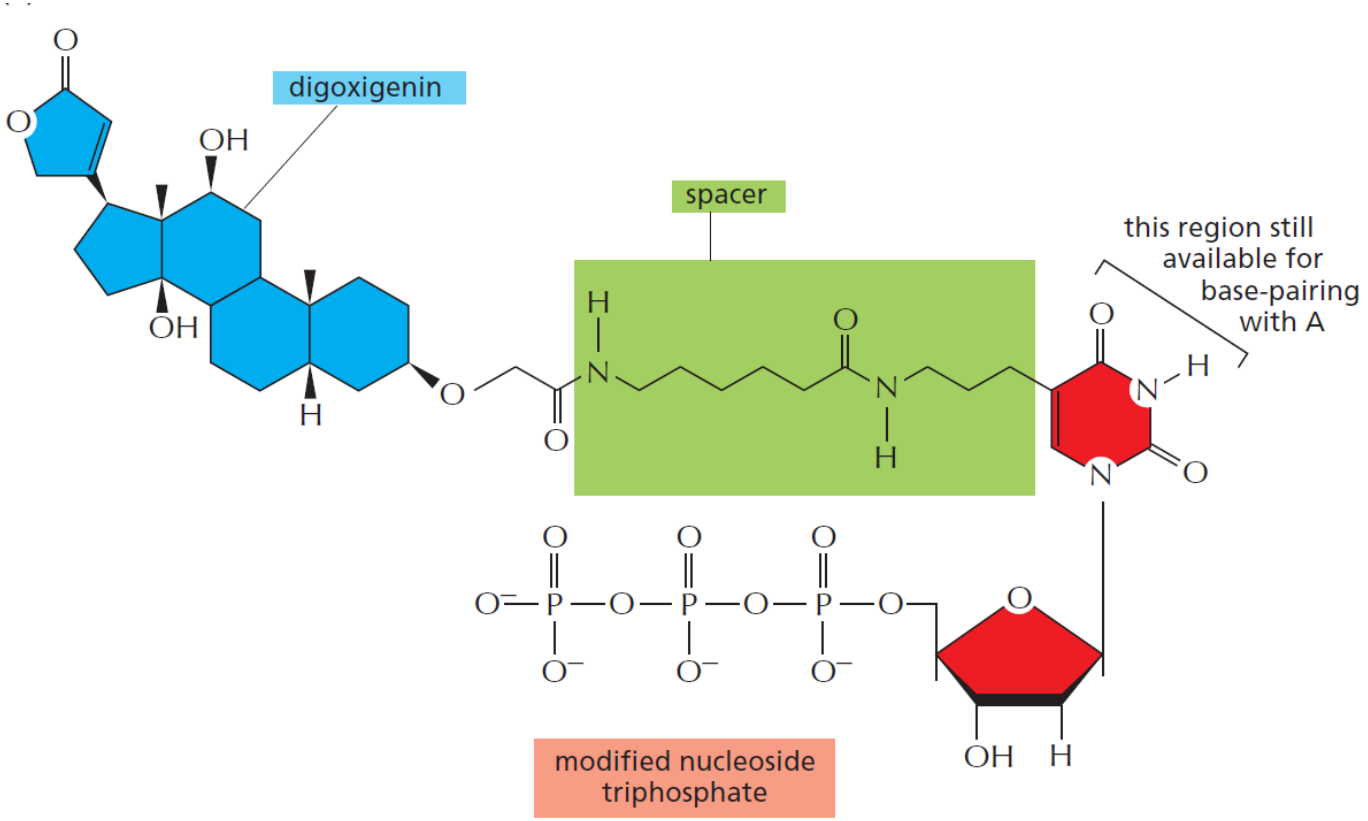
- DNA labeled with the **radioisotope ^{32}P** can be detected following gel electrophoresis by placing the gel next to a piece of photographic film.
- The ^{32}P atoms emit β particles which expose the film, producing a visible record of every band on the gel. Alternatively, the gel can be scanned by a detector that measures the **β emissions** directly.



DNA polymerase incorporates ^{32}P nucleotides, resulting in a population of radiolabeled DNA molecules that contain sequences from both strands

Purified DNA Molecules Can Be Specifically Labeled with Radioisotopes or Chemical Markers in vitro

Other types of modified DNA, such as that labeled by **digoxigenin**, are useful for visualizing DNA molecules in whole cells.

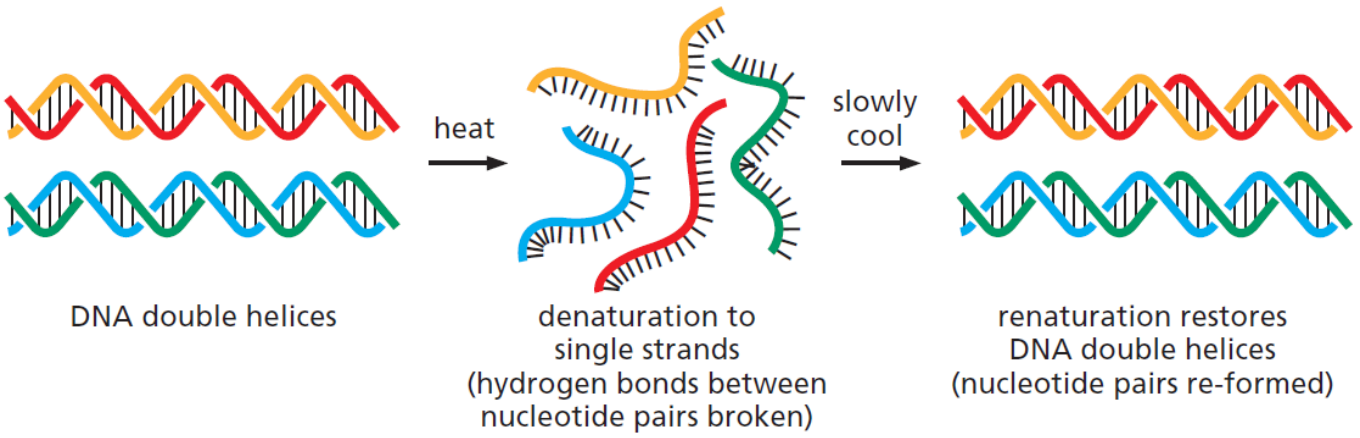


- Modified nucleotide with a specific chemical marker is used to produce nonradioactive DNA molecules that can be detected with an **appropriate antibody**.
- The base on the nucleoside triphosphate shown is an analog of thymine, in which the methyl group on T has been replaced by a **spacer arm linked to the plant steroid digoxigenin**.
- An anti-digoxigenin antibody coupled to a visible marker such as a **fluorescent dye** is then used to visualize the DNA.
- Other chemical labels, such as **biotin**, can be attached to nucleotides and used in the same way.

Hybridization Provides a Powerful, But Simple Way to Detect Specific Nucleotide Sequences

- Under normal conditions, the two strands of a DNA double helix are held together by **hydrogen bonds** between the complementary base pairs.
- But these relatively weak, noncovalent bonds can be fairly easily broken.
- Such **DNA denaturation** will release the two strands from each other, but does not break the covalent bonds that link together the nucleotides within each strand.
- Perhaps the simplest way to achieve this separation involves **heating the DNA to around 90°C**.
- When the conditions are reversed—by slowly lowering the temperature— the complementary strands will readily come back together to re-form a double helix.
- This **hybridization**, or **DNA renaturation**, is driven by the re-formation of the hydrogen bonds between complementary base pairs.

- **This fundamental capacity of a single-stranded nucleic acid molecule, either DNA or RNA, to form a double helix with a single-stranded molecule of a complementary sequence provides a powerful and sensitive technique for detecting specific nucleotide sequences.**



Hybridization Provides a Powerful, But Simple Way to Detect Specific Nucleotide Sequences

- Today, one simply designs a **short, single-stranded DNA molecule** (often called a **DNA probe**) that is complementary to the nucleotide sequence of interest.

- Because the nucleotide sequences of so many genomes are known—and are stored in publicly accessible databases—designing a probe to hybridize anywhere in a genome is straightforward.
-
- Probes are single-stranded, typically **30** nucleotides in length, and are usually synthesized chemically by a commercial service for pennies per nucleotide.

- A DNA sequence of 30 nucleotides will occur by chance only once every 1×10^{18} nucleotides (4^{30}); so, even in the human genome of 3×10^{9} nucleotide pairs, a DNA probe designed to match a particular 30-nucleotide sequence will be highly unlikely to hybridize—by chance— anywhere else on the genome.

- This, of course, presumes that the sequence complementary to the probe does not occur multiple times in the genome, a condition that can be checked beforehand by scanning the genomic sequence in silico (using a computer) and designing probes that match only one spot.

Hybridization Provides a Powerful, But Simple Way to Detect Specific Nucleotide Sequences

The hybridization conditions can be set so that even **a single mismatch** will prevent hybridization to "near-miss" sequences.

The exquisite specificity of nucleic acid hybridization can be easily appreciated by the in situ (Latin for "in place") hybridization experiment.
Nucleic acid hybridization has many uses in modern cell and molecular biology; one of the most powerful is in the cloning of DNA by the **polymerase chain reaction**.

In situ hybridization can be used to locate genes on isolated chromosomes. Here, six different DNA probes have been used to mark the locations of their complementary nucleotide sequences on human Chromosome 5, isolated from a mitotic cell in metaphase. The DNA probes have been labeled with different chemical groups and are detected using fluorescent antibodies specific for those groups.
The chromosomal DNA has been partially denatured to allow the probes to base-pair with their complementary sequences.
Both the maternal and paternal copies of Chromosome 5 are shown, aligned side by side.
Each probe produces two dots on each chromosome because chromosomes undergoing mitosis have already replicated their DNA; therefore, each chromosome contains two identical DNA helices.
The technique employed here is nicknamed **FISH**, for **fluorescence in situ hybridization.**