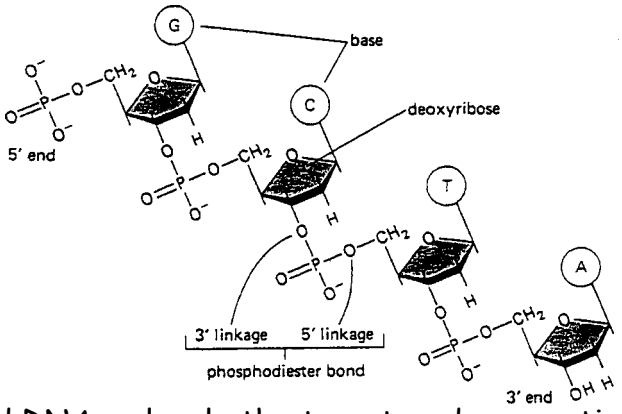
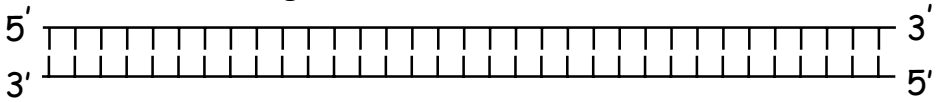
Lecture 13

To see how gene sequences are actually obtained, we will first need to consider some fundamentals of the chemical structure of DNA. Each strand of DNA is directional. The different ends are usually called the 5′ and 3′ ends; referring to different positions on the ribose sugar ring where the linking phosphate residues attach.



In a double stranded DNA molecule the two strands run anti-parallel to one another and the general structure can be diagramed like this:



• Note about representation of DNA sequences.

1) Single strands are always represented in direction of synthesis – 5′ to 3′

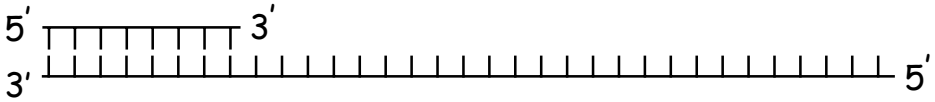
2) For double stranded DNA, usually one strand is represented in the 5′ to 3′ direction. For a gene, the strand represented would correspond to the sequence of the mRNA.

DNA polymersaes are the key players in the methods that we will be considering. The general reaction carried out by DNA polymerase is to synthesize a copy of a DNA template starting with the chemical precursors (nucleotides) dATP, dGTP, dCTP, and dTTP (dNTPs). All DNA polymerases have two fundamental properties in common.

(1) New DNA is synthesized only by elongation of an existing strand at its 3′ end.

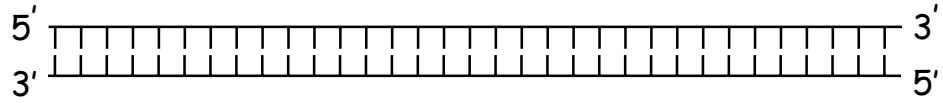
(2) Synthesis requires nucleotide precursors, a free 3′ OH end, and a template strand.

A general substrate for DNA polymerase looks like this:



Note that the template strand can be as short as 1 base or as long as several thousand bases.

After addition of DNA polymerase and nucleotide precursors this product will be readily synthesized:



Polymerase Chain Reaction

Now let's consider how to obtain DNA segments that are suitable for sequencing. At first, DNA sequences were obtained from cloned DNA segments (we will discuss some methods to clone new genes in a subsequent lecture). Presently the entire DNA sequence for *E. coli*, as well as a variety of other bacterial species, has been determined. If we want to find the sequence of a new mutant allele of a known gene we need an easy way to obtain a quantity of this DNA from a culture of bacterial cells. The best way to do this is to use a method known as PCR or polymerase chain reaction that was developed by Kary Mullis in the mid-1980's. The steps in a PCR reaction are as follows.

(1) A crude preparation of chromosomal DNA is extracted from the bacterial strain of interest.

(2) Two short oligo nucleotide primers (each about 18 bases long) are added to the DNA. The primers are designed from the known genomic sequence to be complimentary to opposite strands of DNA and to flank the chromosomal segment of interest.

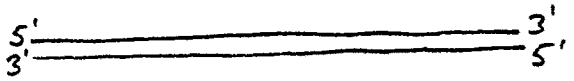
(3) The double stranded DNA is melted by heating to 100°C and then the mixture is cooled to allow the primers to anneal to the template DNA.

(4) DNA polymerase and the four nucleotide precursors are added and the reaction is incubated at 37°C for a period of time to allow a copy of the segment to be synthesized.

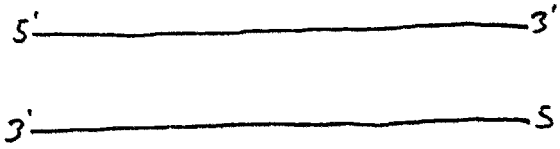
(5) Steps 3 and 4 are repeated multiple times. To avoid the inconvenience of having to add new DNA polymerase in each cycle a special DNA polymerase that can withstand heating to 100°C is used.

The idea is that each cycle – of melting, annealing and DNA synthesis – the amount of DNA in the segment between the primers is doubled. This gives an exponential increase in the amount of the specific DNA segment as the cycles proceed. After 10 cycles the DNA amplified 10^3 fold and after 20 cycles the DNA will be amplified 10^6 fold. Usually amplification is continued until all ofthe nucleotide precursors are incorporated into synthesized DNA.

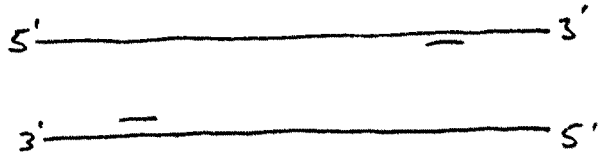
A PCR Reaction



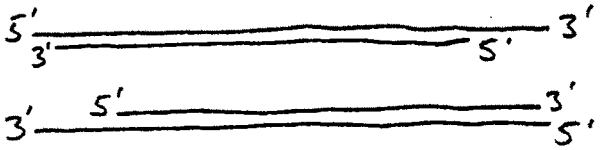
Separate DNA strands



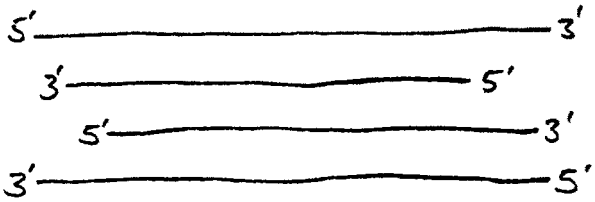
Anneal primers



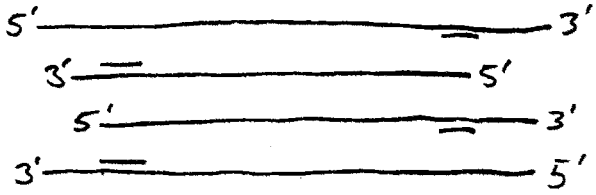
DNA synthesis



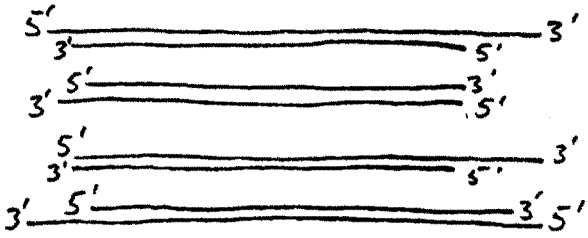
Separate DNA strands



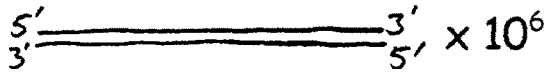
Anneal primers



DNA synthesis



Repeat cycle 20 times



Gene Mutations

Let's say that we are investigating the LacZ gene, which encodes the lactose hydrolyzing enzyme ß-galactosidase. There is a useful compound known as X-gal that can be hydrolyzed by ß-galactosidase to release a dark blue pigment. When X-gal is added to the growth medium in petri plates, Lac^+ E. coli colonies turn blue whereas Lac^- colonies with mutations in the LacZ gene are white. By screening many colonies on such plates it is possible to isolate a collection of E. coli mutants with alterations in the LacZ gene. PCR amplification of the LacZ gene from each mutant followed by DNA sequencing allows the base changes that cause the $LacZ^-$ phenotype to be determined. A very large number of different LacZ mutations can be found but they can be categorized into three general types.

Mutation Type	**Description**
Missense	A base change that converts one codon into another. Many missense mutations are silent because the encoded amino acid remains the same or the amino acid substitution is sufficiently subtle so as not to compromise activity of the enzyme. Missense mutations that have a marked effect often lie in the active site or grossly disrupt protein folding.
Nonsense	A base change that converts a codon within the coding sequence into a stop codon. Note that there is only a limited set of sense codons that can be converted to a stop codon by a single base change. Nonsense mutations lead to a truncated protein product. Nonsense mutations that lie early in the gene sequence will completely inactivate the gene. Sometimes nonsense mutations that lie late in the gene sequence will not disrupt gene function.
Frameshift	The addition or deletion of a base or bases such that the coding sequence is shifted out of register. Note that addition or deletion of a multiple of three bases does not cause a frameshift. After the frameshift mutation is encountered, missense codons will be read up to the first stop codon. Like nonsense mutations, frameshift mutations usually lead to complete inactivation of the gene.

Although many different kinds of mutations occur spontaneously, the frequency with which mutations occur can be increased as much as 10^3 fold by treatment of cells with a mutagen. Here are some general categories of mutagens

Type of Mutagen	Mechanism	Examples	Type of Mutations
Base Analog	Analog is incorporated into DNA and can pair with more than one base	5-bromouracil 2-aminopurine	A•T → G•C, G•C → A•T A•T → G•C
Base modifying agent	Chemical or photo damage to DNA can be repaired, but repair itself is error prone	Hydroxylamine EMS UV	G•C → A•T G•C → A•T, C•G, or T•A All changes
Intercalating agent	Polycyclic compounds can fit between bases and cause mis-copying by polymerase to add or delete bases	Acridine Proflavine ICR-191	Frameshifts (+ or –) " "