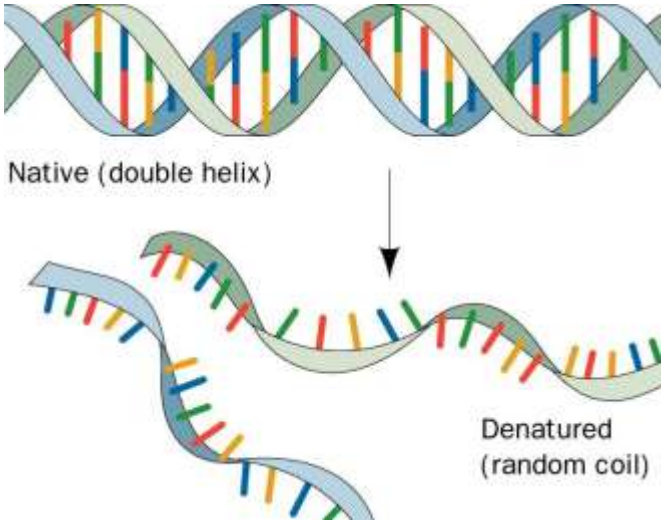
C. Kohn, Waterford, WI

Genomics – Reading What we Can't See

Reading DNA

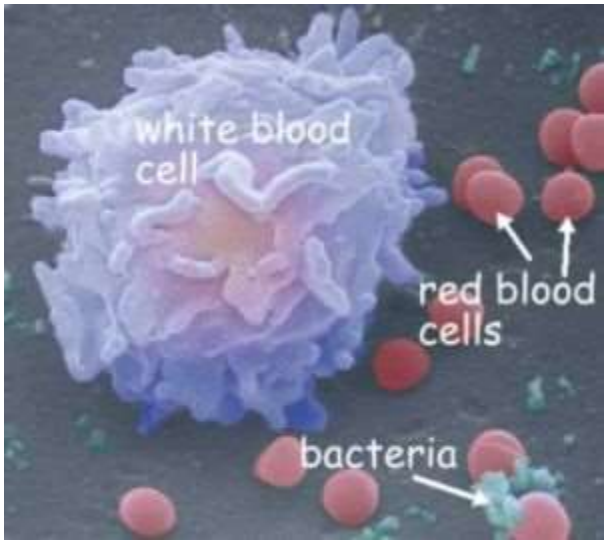
- In today's society, we can take any living organism and read its DNA letter by letter.
  - Whether it is a cow, ear of corn, or a newborn baby, scientists now routinely read DNA to acquire details about species and individuals.
  - This has revolutionized science and agriculture, providing data that until recently was completely inaccessible.



- So how do scientists read something that can't be seen?

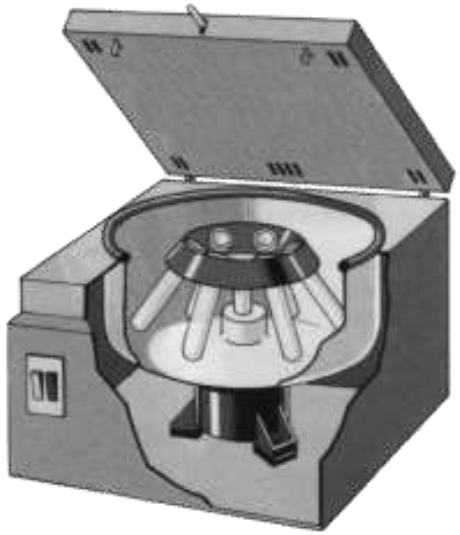
Studying DNA

- To study DNA, scientists must first separate it from everything else that's inside your cell
- For example, if a doctor needed to run a genetic test on you, they might take a blood sample.
- Each white blood cell contains your entire genome – all 3 billion base pairs!
  - The red blood cell do not have a nucleus and therefore do not have DNA



Isolating DNA

- The first step is to centrifuge your blood
  - We'll put a sample of your blood in a machine that spins it rapidly.
- This rapid spinning will cause the denser, heavier portions to move to the bottom of a test tube.
  - The lighter components will move to the top.

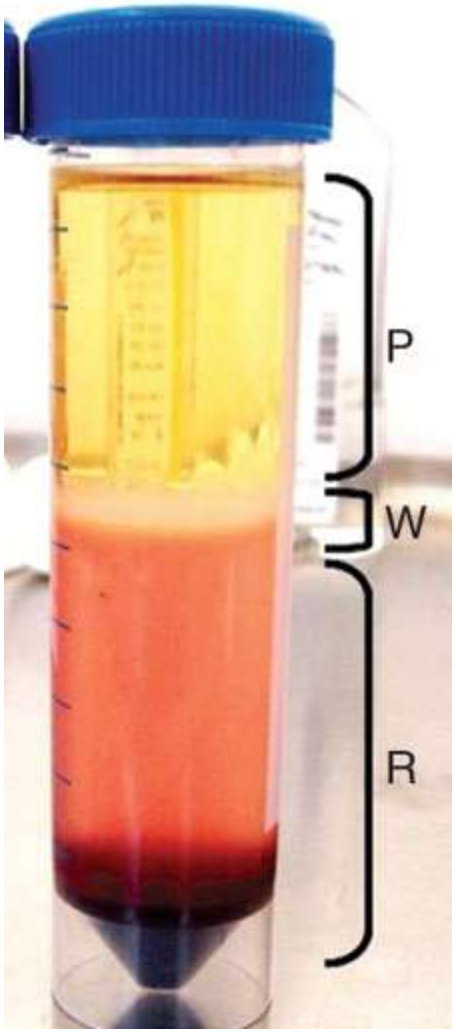


Centrifuging as a Science

- In the late 1800s and early 1900s, Dr. Stephen Babcock used his centrifuges to perform much of his work.
  - His centrifuge is central in his most famous experiment – the butterfat test
- His work led many other researchers to use this as a tool to explore cell biology.
  - We still use this as a primary tool today!

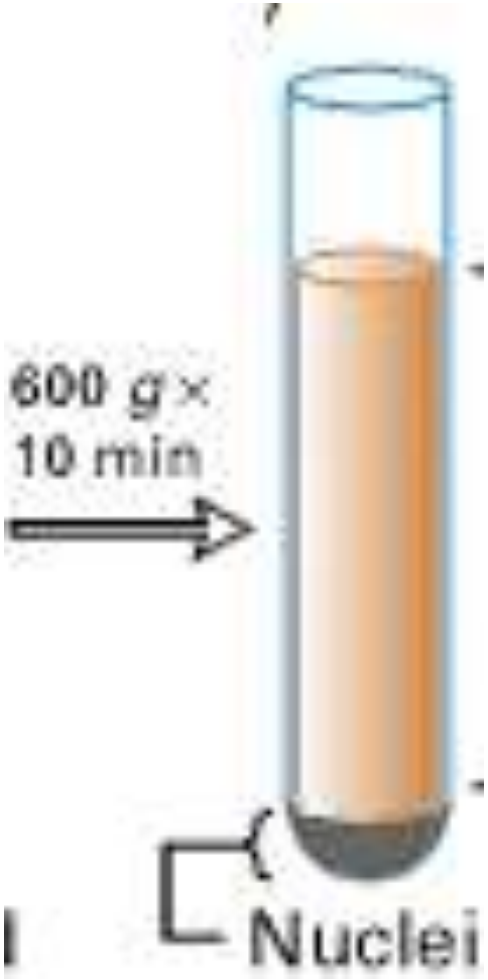
Centrifuging Your Blood

- The fluid (plasma) of the blood will rise to the top after being centrifuged
  - The denser white blood cells will sink to the bottom
  - Below the white blood cells will be the red blood cells



Isolating and Lysing WBCs

- After centrifuging your blood, the white blood cells can be removed.
- The cells are then mixed in a solution that causes them to lyse, or burst open
  - *Lysol* works by bursting open bacterial cells
- The sample is then centrifuged again
  - This separates the nuclei from the rest of the inside of the cell (ribosomes, mitochondria, etc.)



Breaking Open the Nuclei

- Next, we have to break open the nuclei to get to the DNA.
  - The nuclei can be broken open by a mild detergent (like soap)
  - The remaining fats and proteins can be destroyed by special enzymes that kill them but leave your DNA alone.

Final Steps

- After the 'cleaning' process is done, alcohol is added
  - DNA is not soluble in alcohol, and so it clumps and rises to the top
- You now have pure DNA to study with a method so simple, we can do it right here at school.
  - In modern labs, commercial machines do this automatically to speed up the process.

Review of Steps

- 1. Get a blood sample.
- 2. Centrifuge the blood to separate white blood cells from red blood cells and plasma.
- 3. Break open the white blood cells with a detergent.
- 4. Centrifuge the white blood cell contents to separate the nuclei from the other cell contents
- 5. Break open the nuclei with a detergent
- 6. Add alcohol – DNA is not soluble in alcohol and will float to the top.

Reading DNA

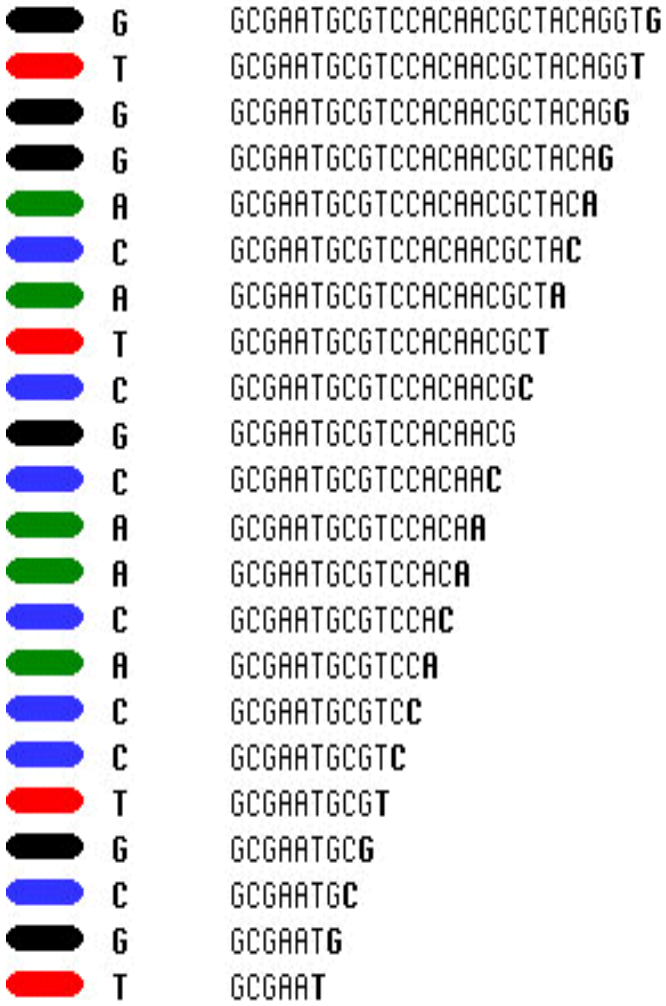
Now that we've isolated it, how do we read it?

Sanger Method

- The Sanger Method is what we use if we want to read a gene or even an entire genome step by step, letter by letter, base by base.
- The Sanger Method works by creating strands of different lengths of the same sequence of DNA.
  - Each copied strand of DNA is color coded.
  - This color is determined by the last base added.
  - By lining up each chunk of the strand from shortest to longest, we can read each strand by its color to determine the next base in a sequence *(next slide).*

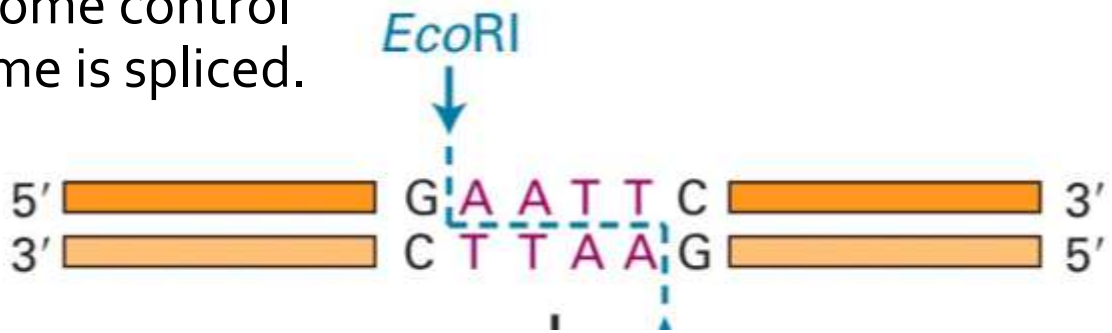
Sanger Method

- As you can see, this is simply the same gene copied over and over.
- Each time, the copying was stopped at a random point.
- The snippets of DNA were then lined up shortest to longest, and the color reflects the last base added (A, T, G, or C)



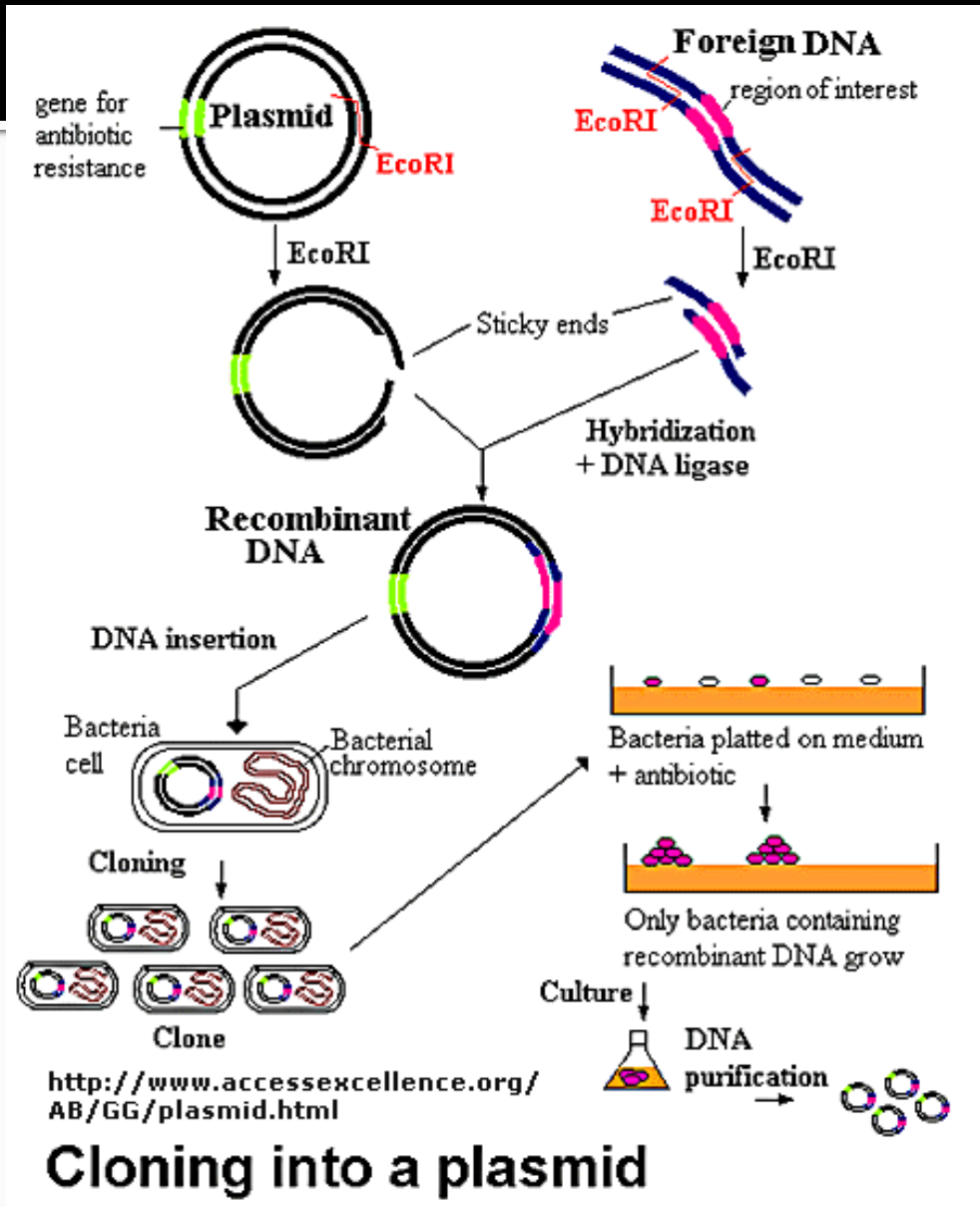
Step 1: Cutting

- 1. The first step in the Sanger Method is to break up the DNA into more manageable chunks
  - 6 billion bases are far too many to read all at once!
  - These chunks are created using a Restriction Enzyme.
  - The restriction enzyme is sort of like a chemical scissors.
    - However, it only cuts DNA when it sees a specific DNA sequence (e.g. GAATTC)
    - This gives scientists some control over where the genome is spliced.

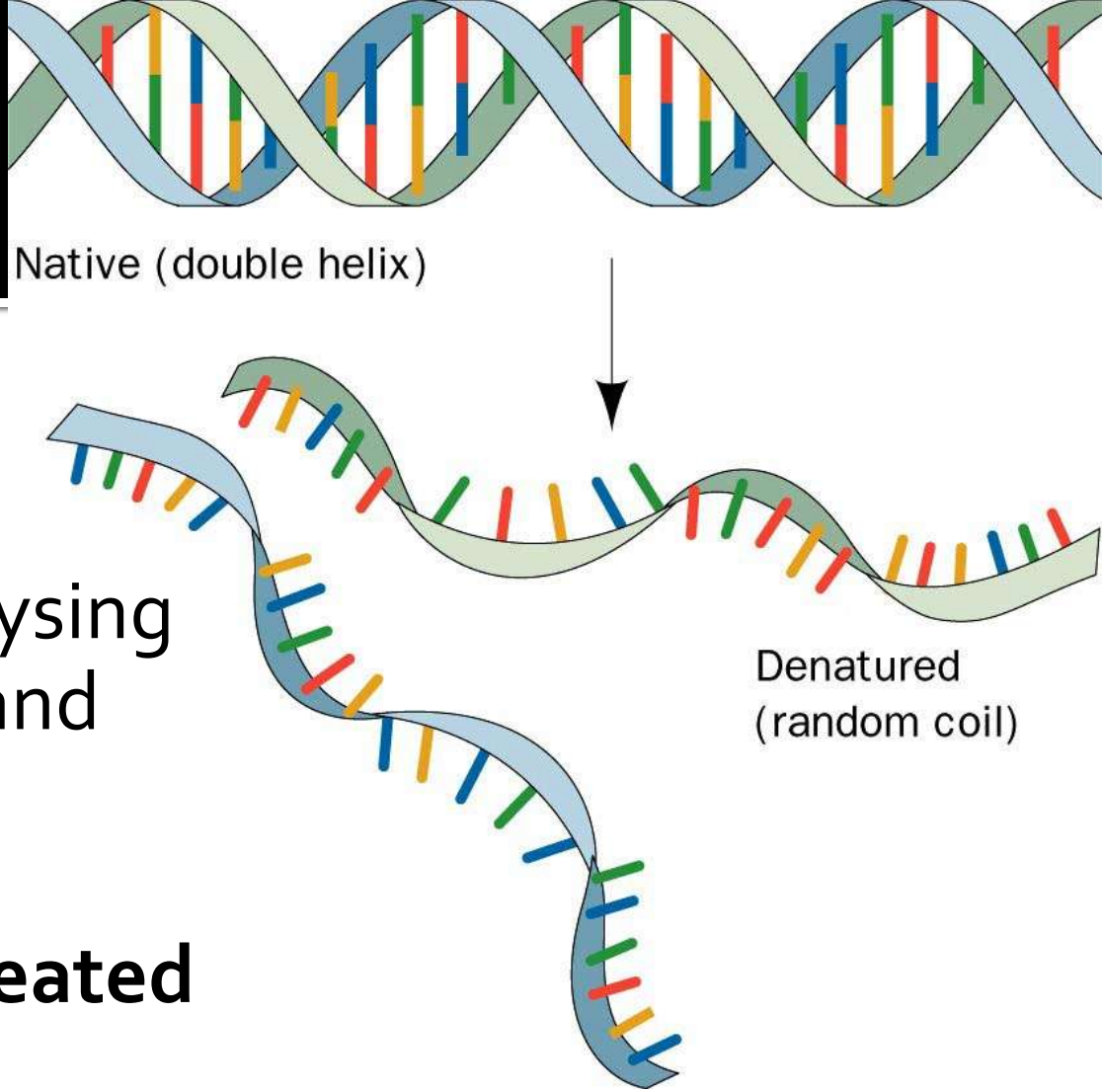


2. Copying

- Step 2: once we've cut out the gene we're interested in, we can put it into a bacterial genome.
  - The bacterial cell will divide, copying its DNA and the inserted gene over and over again.
  - Usually, scientists use a harmless version of *E. coli*, the same kind found in our stomachs



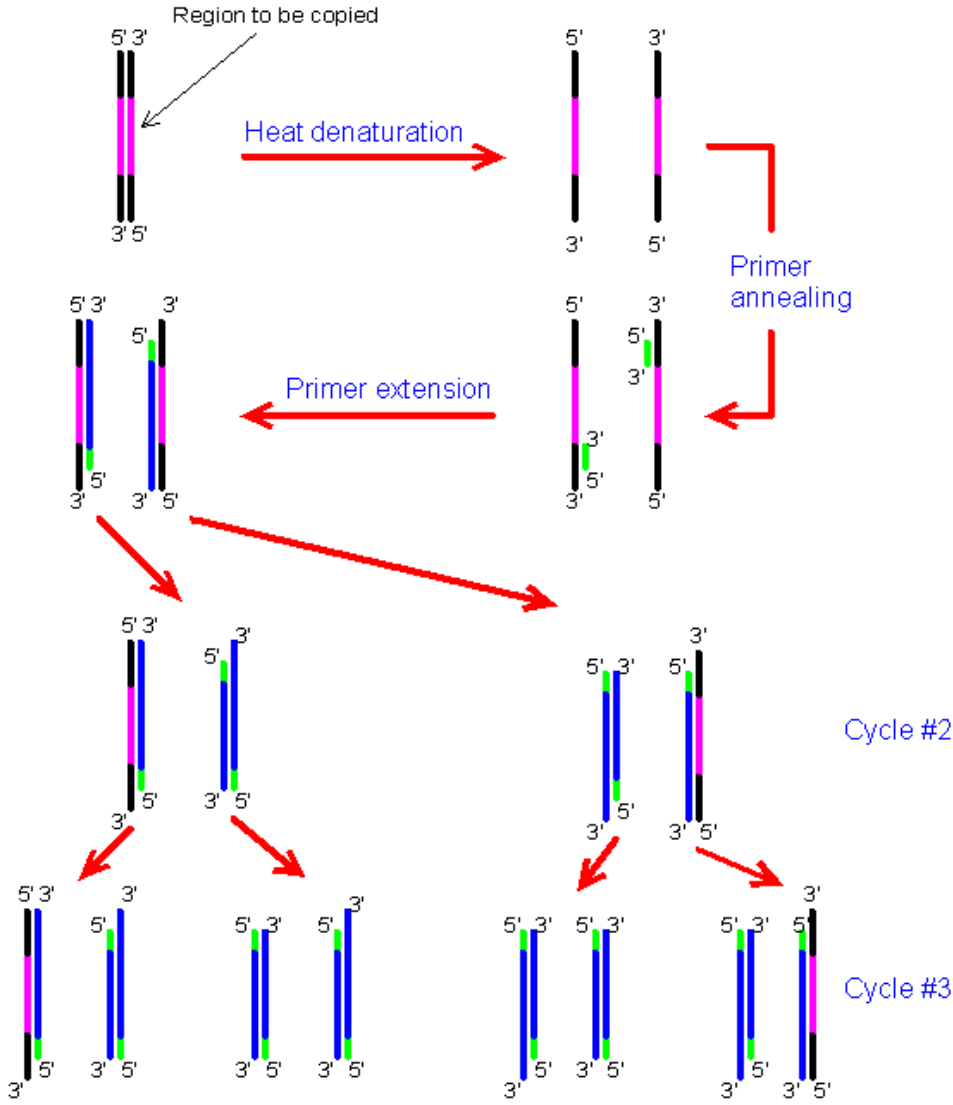
3. Denaturing



- Step 3: the DNA is removed from the bacterial cells (by lysing and centrifuging) and purified.
- The DNA is then **heated** so that it can be <u>denatured</u> (become single stranded)

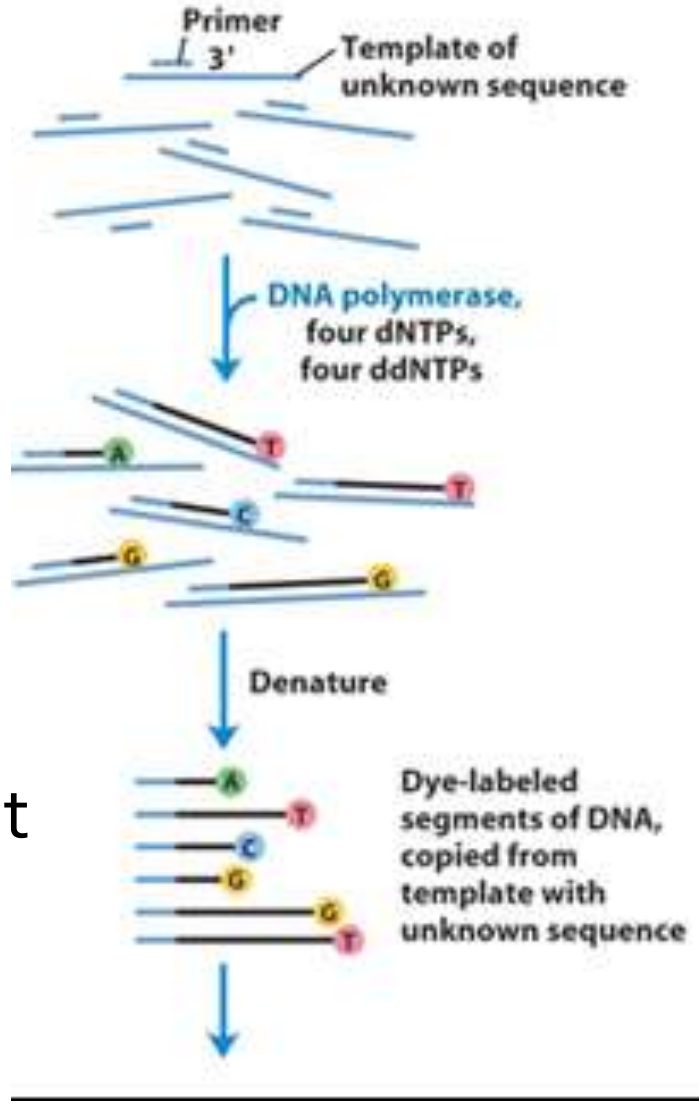
4. Addition of Primers

- Step 4: primers are added to tell the polymerase where to add bases
  - Basically, the primer is sort of like the lights on an airport runway
  - Just like those lights tell the plane where to land, primers tell polymerase where to start.



5. Addition of Polymerase and ddNTPs

- Step 5: Once the DNA has been denatured and primed, we can add the polymerase, bases, and ddNTPs
  - Reminder: *polymerase* is the enzyme that adds bases to make a copy strand of DNA
  - A ddNTP is a special nucleotide that makes this whole process work



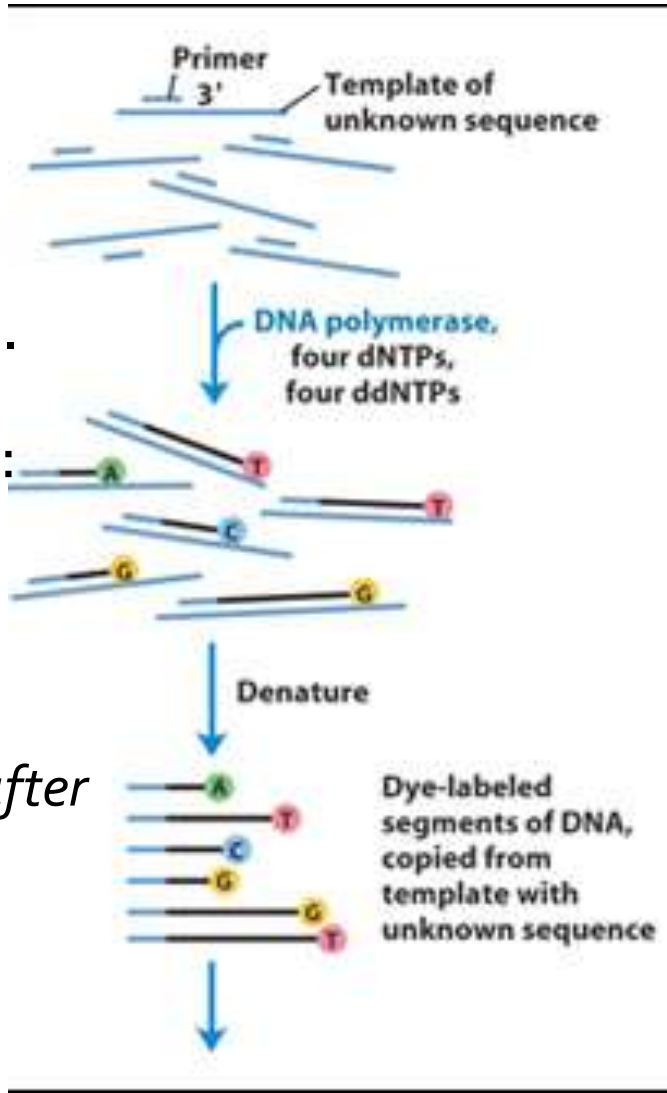
ddNTPs

- ddNTPs (or dideoxyribonucleotides) are really 'fake' nucleotides.
- They are just like regular nucleotides, but with one crucial difference – no additional bases can be added afterwards.

- For example, you might see the following:

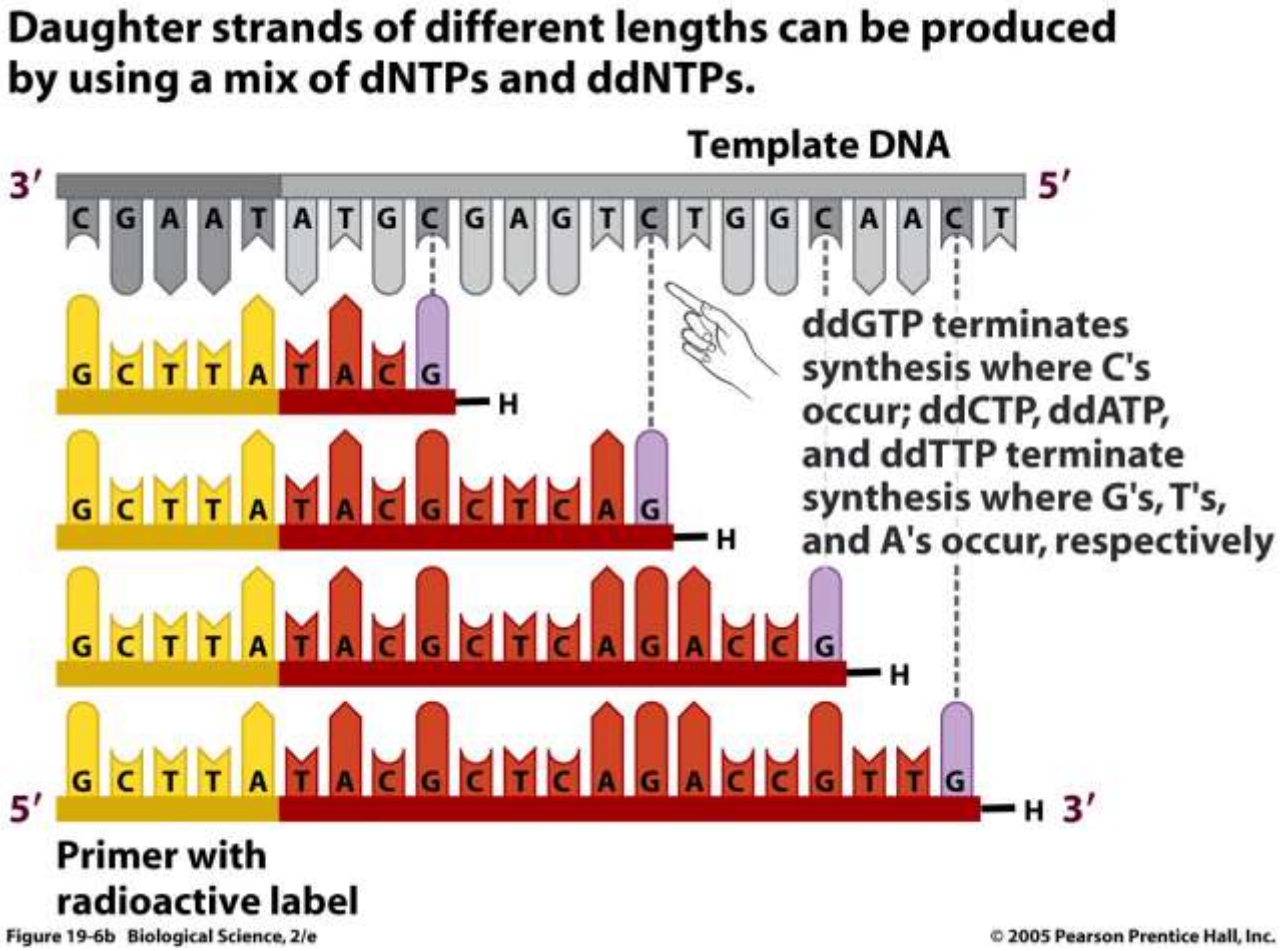
  5'-ATCGTAGTGACGTTATTCCCAT-3'
- 3'-TAGCTAGACTGCAATAA<u>G</u>

  - *This last <u>G</u> was a ddNTP – nothing can come after it!*



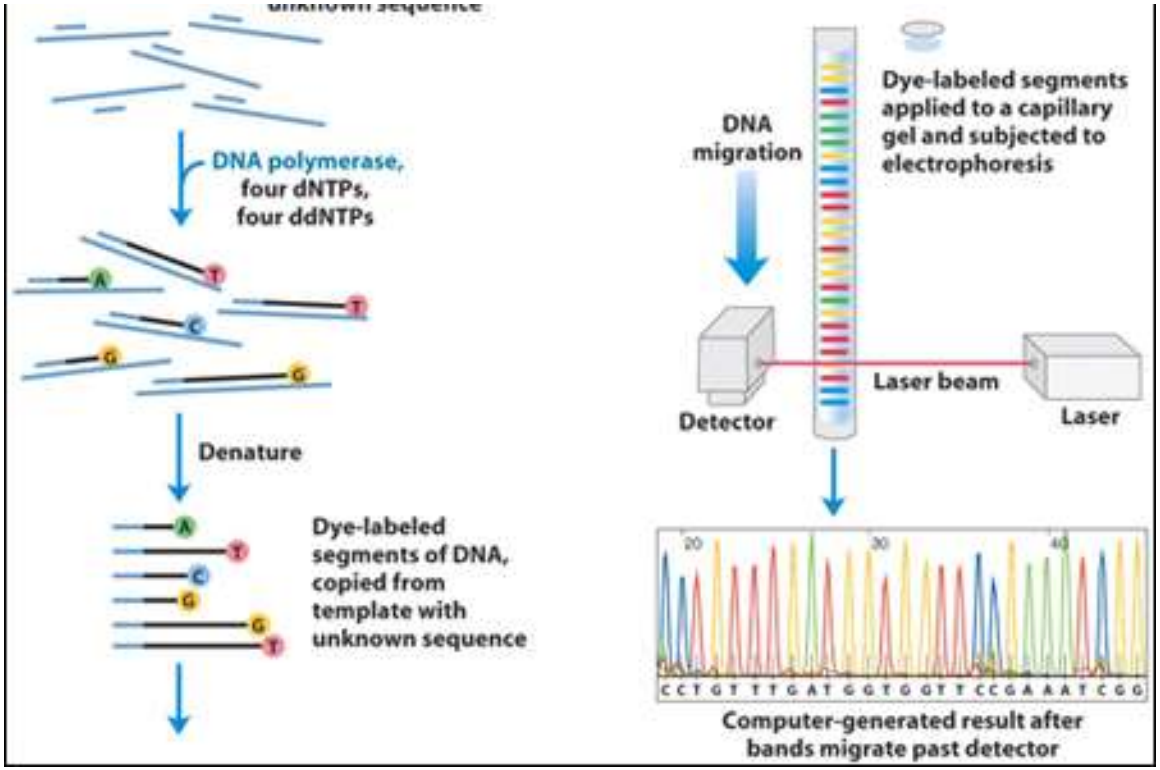
ddNTPs (cont)

- A ddNTP will also be 'tagged' with a dye.
- Each base will have a specific color.
- We will repeat this process many times, creating many copies of the same gene (but of varying lengths).



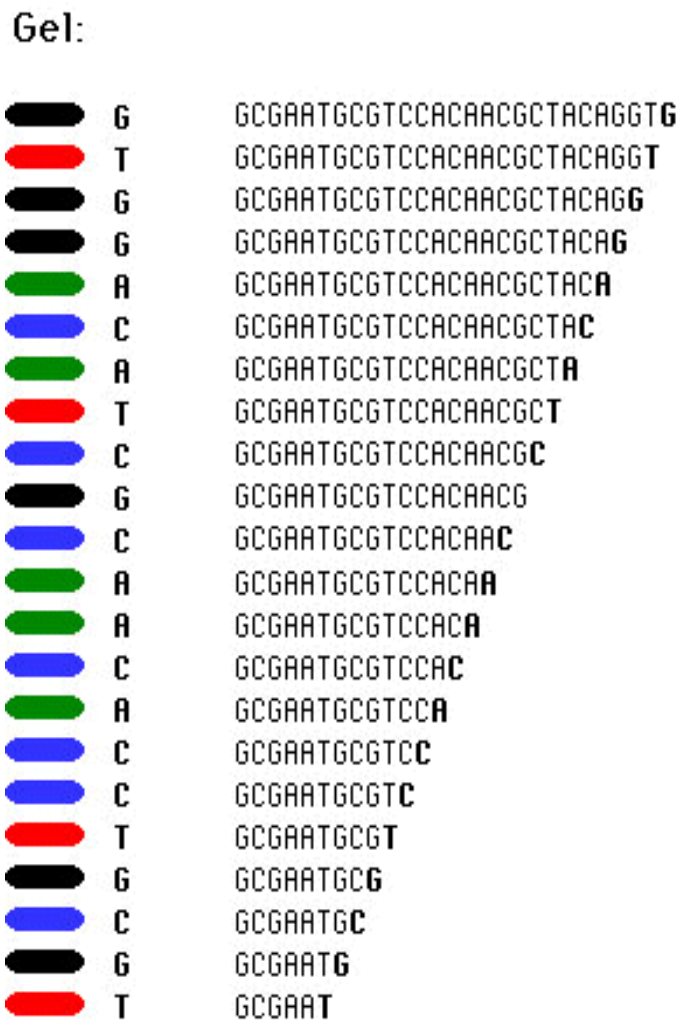
6. Gel Electrophoresis

- Next we will put all of the copies of that gene into an electrically charged gel.
  - DNA will move towards the positive end of the gel (because DNA is negatively charged.



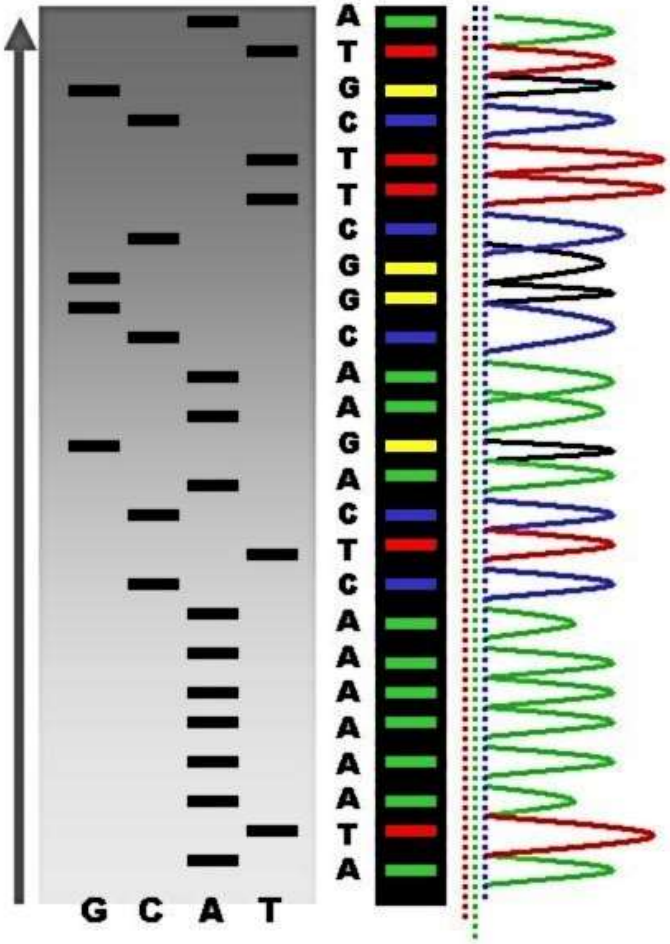
6. Gel Electrophoresis (cont)

- As it moves through the gel, the smaller fragments will move further than the large fragments
  - They have less resistance and will be able to go further in the same amount of time
- Each fragment will line up longest to shortest.
- Each gene will end with the color of the last base added (the ddNTP).



7. Reading the Sequence

- A computer will then read each color and record the base found at each point.
  - The computer isn't looking at just *one* individual nucleotide, but a collection of thousands of the same stretch of DNA.
- Each stretch of DNA is dyed the same color corresponding to the letter of the last ddNTP added.

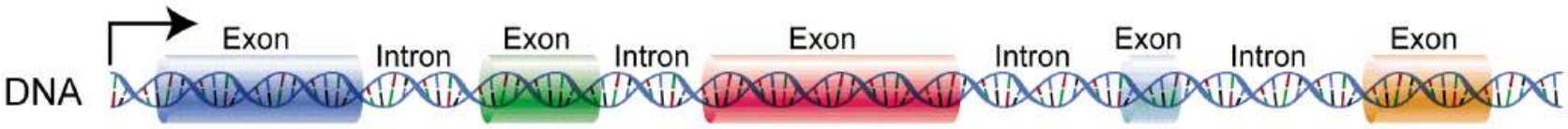


8. Analyzing the DNA

- A problem we face when reading DNA is that most of our DNA codes for nothing.
- Lengths of non-coding DNA are called introns.
  - Introns are genes for nothing – they aren't used to create proteins (they are incapable of creating proteins).
  - They are like junk DNA (*sort of*).
- On the other hand, we want to see exons, the DNA that is actually used to create a protein.
  - Exons are exceptional at creating proteins.

Introns vs. Exons

- So how do we tell the difference between junk introns and useful exons?
- Exons have an Open Reading Frame, or ORF
  - An ORF simply means that there are no STOP commands between the beginning and end.
  - *A computer can be programmed to look for multiple STOP commands*
  - *This would tell us that a gene is a useless intron (non-colored below)*
- If none are found, we know that the stretch of DNA is a useful, protein-coding exon!



Source: simple.wikipedia.org

The Human Genome Project

The Human Genome Project

- The Human Genome Project is among the greatest achievements ever in science.
- It is comparable to landing a man on the moon, creating the Internet, and eradicating small pox.
- It was created faster than expected, without going over budget, and united scientists all over the world from many nations.

Things to consider…

- 60 years ago, we didn't even know what DNA looked like or how it worked.
- In 1990 we hadn't read any genomes
- In 2000, we had only completed the genome of a few species.
- Today, we can identify every letter in your genome or of any living species.
- The progress and rate of discovery of biotechnology is still increasing, while the cost of this work is continually decreasing.

Other Genome Projects

- Since the completion of the HGP, the genomes of the cow, dog, horse, pig, and many other animals have been sequenced as well.
- So why would we want to know the *entire* genome of humans and other animals?

Benefits of the Genome Project

- Since its completion in 2003 (2 years ahead of schedule), over 1800 genes for human disease have been discovered
- Today, researchers can find the cause of a genetic disease in days. Before the HGP, it took years
- Over 1000 proven tests now exist for genetic diseases as a direct result

Benefits to Come

According to the National Institute of Health, the Human Genome Project will make it likely that…

- The genetic factors for heart disease, diabetes, mental illnesses, and many other problems will be discovered in a few years
- The causes of all 50 types of cancer can be determined in detail, contributing to a cure
- Drugs may possibly become tailor made to suit our specific genetic needs, making the more effective
- Personalized and preemptive medical procedures may replaced reactive medicine practiced today
  - I.e. we may be able to predict who will have a heart attack and who won't

http://www.nih.gov/about/researchresultsforthepublic/HumanGenomeProject.pdf