The Sanger Method

By Sarah Obenrader, Davidson College
http://www.bio.davidson.edu/Courses/Molbio/MolStudents/spring2003/Obenrader/sanger_method_page.htm

Background Information

DNA sequencing enables us to perform a thorough analysis of DNA because it provides us with the most basic information of all: the sequence of nucleotides. With this knowledge, for example, we can locate regulatory and gene sequences, make comparisons between homologous genes across species and identify mutations. Scientists recognized that this could potentially be a very powerful tool, and so there was competition to create a method that would sequence DNA. Then in 1974, two methods were independently developed by an American team and an English team to do exactly this. The Americans, lead by Maxam and Gilbert, used a "chemical cleavage protocol", while the English, lead by Sanger, designed a procedure similar to the natural process of DNA replication. Even though both teams shared the 1980 Nobel Prize, Frederick Sanger's method became the standard because of its practicality (Speed, 1992).

Sanger's method, which is also referred to as dideoxy sequencing or chain termination, is based on the use of dideoxynucleotides (ddNTP's) in addition to the normal nucleotides (NTP's) found in DNA. Dideoxynucleotides are essentially the same as nucleotides except they contain a hydrogen group on the 3' carbon instead of a hydroxyl group (OH). These modified nucleotides, when integrated into a sequence, prevent the addition of further nucleotides. This occurs because a phosphodiester bond cannot form between the dideoxynucleotide and the next incoming nucleotide, and thus the DNA chain is terminated.

The Method

Before the DNA can be sequenced, it has to be denatured into single strands using heat. Next a primer is annealed to one of the template strands. This primer is specifically constructed so that its 3' end is located next to the DNA sequence of interest. Either this primer or one of the nucleotides should be radioactively or fluorescently labeled so that the final product can be detected on a gel (Russell, 2002). Once the primer is attached to the DNA, the solution is divided into four tubes labeled "G", "A", "T" and "C". Then reagents are added to these samples as follows:

guanine "G" tube: all four dNTP's, ddGTP and DNA polymerase
adenine "A" tube: all four dNTP's, ddATP and DNA polymerase
thymine "T" tube: all four dNTP's, ddTTP and DNA polymerase
cytosine "C" tube: all four dNTP's, ddCTP and DNA polymerase

As shown above, all of the tubes contain a different ddNTP present, and each at about one-hundreth the concentration of the the normal precursors (Russell, 2002). As the DNA is synthesized, nucleotides are added on to the growing chain by the DNA polymerase. However, on occasion a dideoxynucleotide is incorporated into the chain in place of a normal nucleotide, which results in a chain-terminating event. For example if we looked at only the **"G" tube**, we might find a mixture of the following products:

Figure 1: An example of the potential fragments that could be produced in the "G" tube. The fragments are all different lengths due to the random integration of the ddGTP's (Metzenberg).

The key to this method, is that all the reactions start from the same nucleotide and end with a specific base. Thus in a solution where the same chain of DNA is being synthesized over and over again, the new chain will terminate at all positions where the nucleotide has the potential to be added because of the integration of the dideoxynucleotides (Russell, 2002). In this way, bands of all different lengths are produced. Once these reactions are completed, the DNA is once again denatured in preparation for electrophoresis. The contents of each of the four tubes are run in separate lanes on a polyacrylmide gel in order to separate the different sized bands from one another. After the contents have been run across the gel, the gel is then exposed to either UV light or X-Ray, depending on the method used for labeling the DNA.

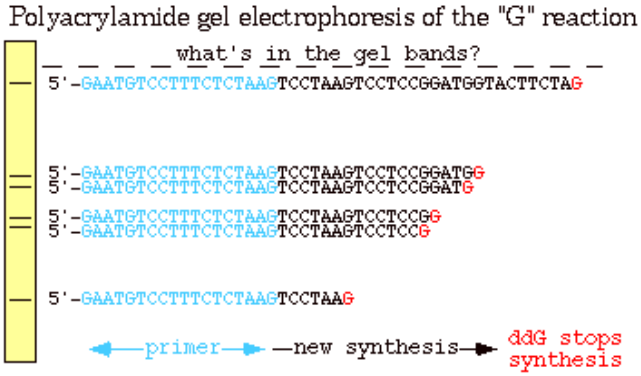


Figure 2: This is a polyacrylmide gel of the reactions in the "G" tube (the same sequences seen in figure 1). The longer fragments of DNA traveled shorter distances than the smaller fragments because of their heavier molecular weight.The blue section indicates the primer, the black section indicates the newly synthesized strand and the red denotes a ddGTP, which terminated the chain (Metzenberg).

As shown in Figure 2, smaller fragments are produced when the ddNTP is added closer to the primer because the chains are smaller and therefore migrate faster across the gel. If all of the reactions from the four tubes are combined on one gel, the actual DNA sequence in the 5' to 3' direction can be determined by reading the banding pattern from the bottom of the gel up. It is important to remember though that this sequence is complementary to the template strand from the beginning.

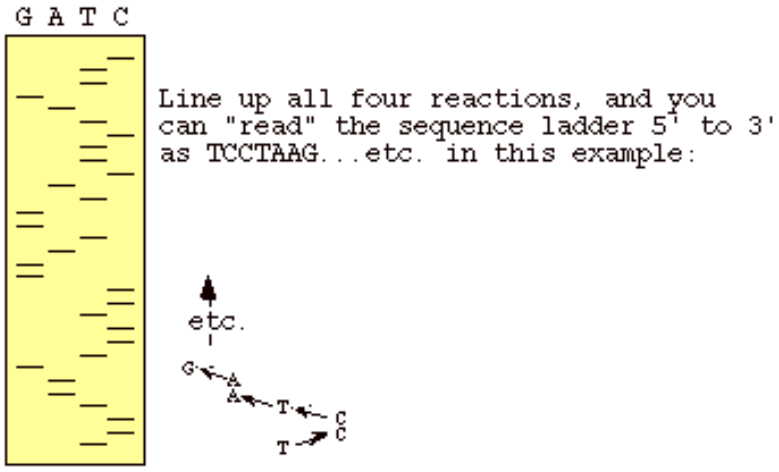


Figure 3: This is an autoradiogram of a dideoxy sequencing gel. The letters over the lanes indicate which dideoxy nucleotide was used in the sample being represented by that lane. When you read from the bottom up, you are reading the *complementary* sequence of the template strand (Metzenberg).

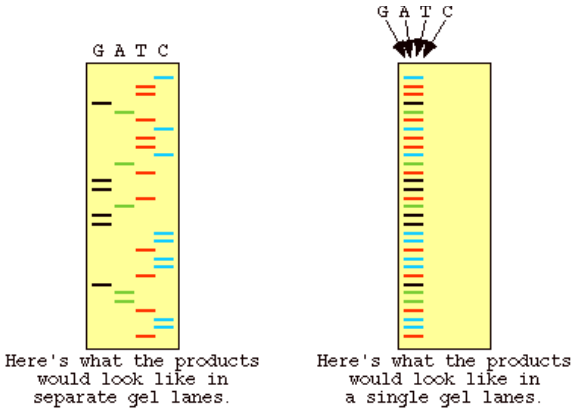
Automated Sequencing

With the many advancements in technology that we have achieved since 1974, it is no surprise that the Sanger method has become outdated. However, the new technology that has emerged to replace this method is based on the same principles of Sanger's method. Automated sequencing has been developed so that more DNA can be sequenced in a shorter period of time. With the automated procedures the reactions are performed in a single tube containing all four ddNTP's, each labeled with a different color dye (Russell, 2002).

Figure 4: In automated se-quencing, the oligonucleotide primers can be "end-labeled" with different color dyes, one for each ddNTP. These dyes fluoresce at different wave-lengths, which are read via a machine (Metzenberg).

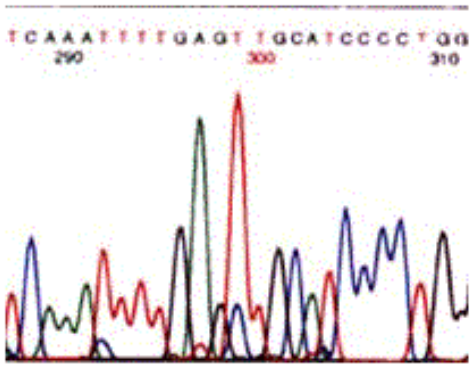
As in Sanger's method, the DNA is separated on a gel, but they are all run on the same lane as opposed to four different ones.

Figure 5: Results of gel electrophoresis for the dye labeled DNA in automated sequencing. The image on the left shows what the gel looks like if the four reactions are run in different lanes, as opposed to the image on the right which shows a gel where all the DNA is run in one lane (Metzenberg).



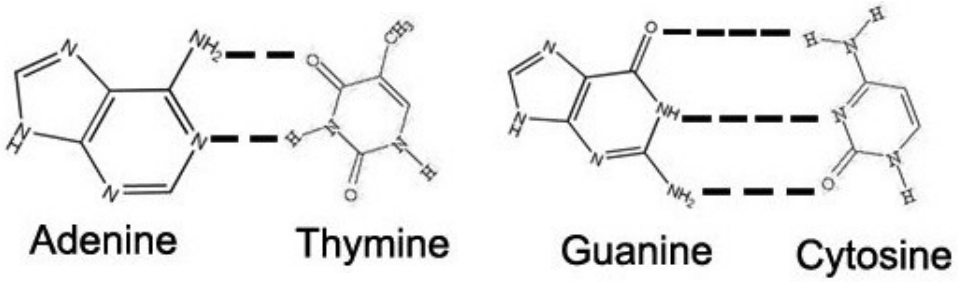
Since the four dyes fluoresce at different wavelengths, a laser then reads the gel to determine the identity of each band according to the wavelengths at which it fluoresces. The results are then depicted in the form of a chromatogram, which is a diagram of colored peaks that correspond to the nucleotide in that location in the sequence (Russell, 2002).

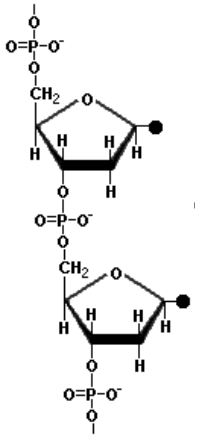
Figure 6: Results from an automated sequence shown in the form of a chromatogram. The colors represent the four bases: blue is C, green is A, black is G and red is T (Metzenberg).



References

1. Russell, Peter. 2002. *iGenetics.* Pearson Education, Inc., San Francisco, pp. 187-189.
2. Metzenberg, Stan. Sanger Method- Dideoxynucleotide Chain Termination. http://www.csun.edu/~hcbio027/biotechnology/lec3/sanger.html (Feb. 10th, 2003).
3. Speed, Terrence. Sanger's Method. http://statwww.berkeley.edu/users/terry/Classes/s260.1998/Week8b/week8b/node9.html (Feb. 10th, 2003).





Next-generation Methods (from Wikipedia)

The high demand for low-cost sequencing has driven the development of high-throughput sequencing (or next-generation sequencing) technologies that parallelize the sequencing process, producing thousands or millions of sequences at once. High-throughput sequencing technologies are intended to lower the cost of DNA sequencing beyond what is possible with standard dye-terminator methods. In ultra-high-throughput sequencing as many as 500,000 sequencing-by-synthesis operations may be run in parallel.

Comparison of next-generation sequencing methods

Method	Single-molecule real-time sequencing (Pacific Bio)	Ion semiconductor (Ion Torrent sequencing)	Pyro-sequencing (454)	Sequencing by synthesis (Illumina)	Sequencing by ligation (SOLiD sequencing)	Chain termination (Sanger sequencing)
Read length	2900 bp average	200 bp	700 bp	50 to 250 bp	50+35 or 50+50 bp	400 to 900 bp
Accuracy	87% (read length mode), 99% (accuracy mode)	98%	99.9%	98%	99.9%	99.9%
Reads per run	35–75 thousand	up to 5 million	1 million	up to 3 billion	1.2 to 1.4 billion	N/A
Time per run	30 minutes to 2 hours	2 hours	24 hours	1 to 10 days, depending upon sequencer and specified read length	1 to 2 weeks	20 minutes to 3 hours
Cost per 1 million bases (in US$)	$2	$1	$10	$0.05 to $0.15	$0.13	$2400
Advantages	Longest read length. Fast. Detects 4mC, 5mC, 6mA.	Less expensive equipment. Fast.	Long read size. Fast.	Potential for high sequence yield, depending upon sequencer model and desired application.	Low cost per base.	Long individual reads. Useful for many applications.
Disadvant-ages	Low yield at high accuracy. Equipment can be very expensive.	Homopoly-mer errors.	Runs are expensive. Homopoly-mer errors.	Equipment can be very expensive.	Slower than other methods.	More expensive and impractical for larger sequencing projects.

Polymerase Chain Reaction (PCR)
<https://www.genome.gov/10000207/polymerase-chain-reaction-pcr-fact-sheet/>

What is PCR?
Sometimes called "molecular photocopying," the polymerase chain reaction (PCR) is a fast and inexpensive technique used to "amplify" - copy - small segments of DNA. Because significant amounts of a sample of DNA are necessary for molecular and genetic analyses, studies of isolated pieces of DNA are nearly impossible without PCR amplification. Often heralded as one of the most important scientific advances in molecular biology, PCR revolutionized the study of DNA to such an extent that its creator, Kary B. Mullis, was awarded the Nobel Prize for Chemistry in 1993.

What is PCR used for?
Once amplified, the DNA produced by PCR can be used in many different laboratory procedures. For example, most mapping techniques in the Human Genome Project (HGP) relied on PCR. PCR is also valuable in a number of laboratory and clinical techniques, including DNA fingerprinting, detection of bacteria or viruses (particularly AIDS), and diagnosis of genetic disorders.

How does PCR work?
To amplify a segment of DNA using PCR, the sample is first heated so the DNA denatures, or separates into two pieces of single-stranded DNA. Next, an enzyme called "Taq polymerase" synthesizes - builds - two new strands of DNA, using the original strands as templates. This process results in the duplication of the original DNA, with each of the new molecules containing one old and one new strand of DNA. Then each of these strands can be used to create two new copies, and so on, and so on. The cycle of denaturing and synthesizing new DNA is repeated as many as 30 or 40 times, leading to more than one billion exact copies of the original DNA segment.

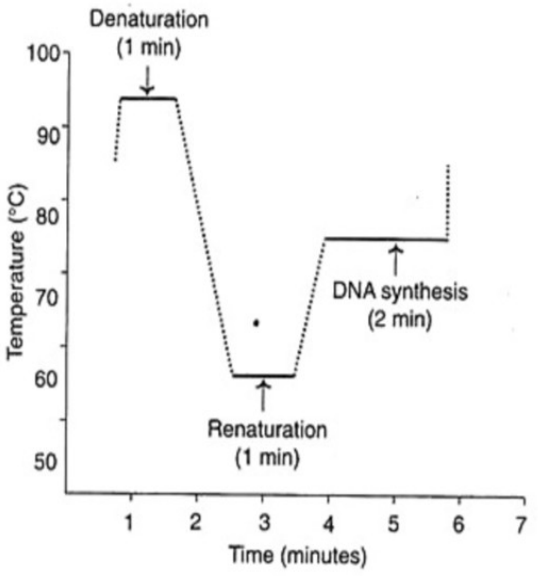


Fig. 8.1 : The three stages in each cycle of PCR in relation to temperature and time (Each cycle takes approximately 3-5 minutes).