Background Information

VNTR HUMAN DNA TYPING

Polymorphic DNA refers to chromosomal regions that vary widely from individual to individual. By examining several of these regions within the genomic DNA obtained from an individual, one may determine a "DNA Fingerprint" for that individual. DNA polymorphisms are now widely used for determining paternity/maternity, kinship, identification of human remains, and the genetic basis of various diseases. The most far-reaching application, however, has been in the field of criminal forensics. DNA from crime victims and offenders can be matched to crime scenes, often affecting the outcome of criminal and civil trials.

The beginning of DNA fingerprinting occurred in the United Kingdom in 1984, following the pioneering work of Dr. Alex Jeffreys at the University of Leicester. Analysis by Jeffreys led to the apprehension of a murderer in the first DNA fingerprinting case in September 1987. The first U.S. conviction occurred on November 6, 1987 in Orlando, FL. Since then, DNA analysis has been used in thousands of convictions. Additionally, over 100 convicted prison inmates have been exonerated from their crimes, including several death row inmates.

In 1990, the Federal Bureau of Investigation (FBI) established the Combined DNA Index System (CODIS), a system which allows comparison of crime scene DNA to DNA profiles in a convicted offender and a forensic (crime scene) index. A match of crime scene DNA to a profile in the convicted offender index indicates a suspect for the crime, whereas a match of crime scene DNA to the forensic index (a different crime scene) indicates a serial offender. CODIS has now been used to solve dozens of cases where authorities had not been able to identify a suspect for the crime under investigation.

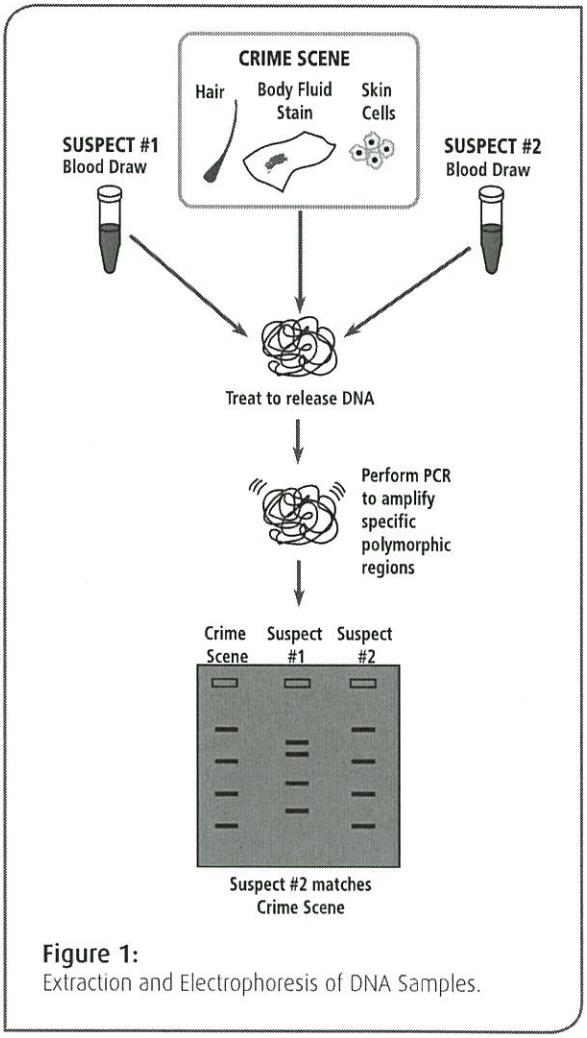


Figure 1:
Extraction and Electrophoresis of DNA Samples.

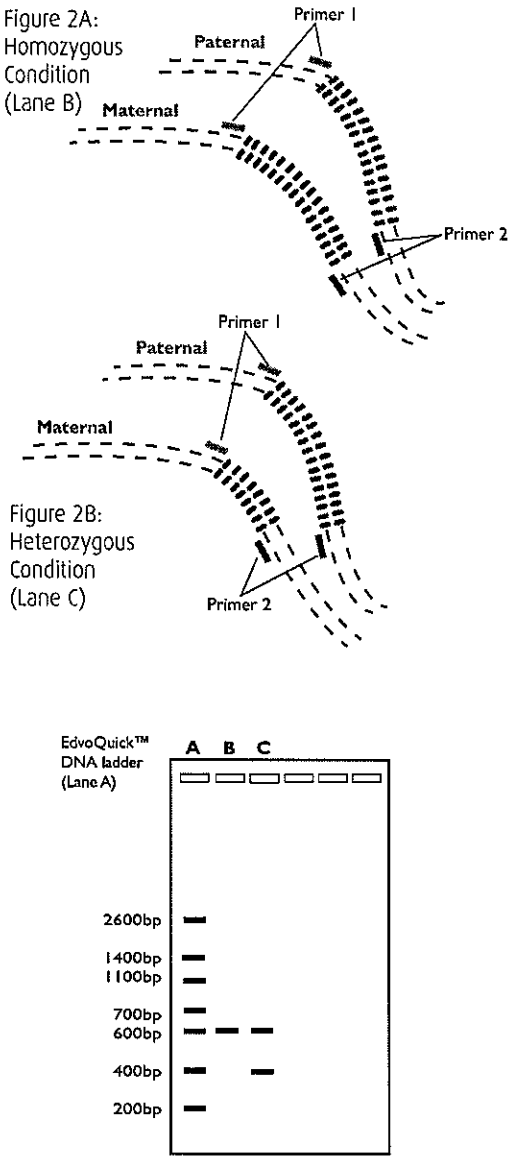
The first step in forensic DNA fingerprinting is the collection of human tissue from the crime scene or victim. These tissues include blood, hair, skin, and body fluids. The sample, often present as a stain, is treated with a detergent to rupture (lyse) cell membranes and obtain DNA for further analysis (Figure 1). One early method, called Restriction Fragment Length Polymorphism (RFLP) analysis, involves digesting DNA with restriction enzymes, separating the fragments by agarose gel electrophoresis, transferring the DNA to a membrane, and hybridizing the membrane with probes to polymorphic regions. This method is statistically very accurate, but requires relatively large amounts of DNA and takes several days to complete. Because of the time and DNA requirements, the RFLP

method is no longer used in forensics, but remains in use in certain medical genetics-based tests.

In forensics, the polymerase chain reaction (PCR) is now used to amplify and examine highly polymorphic DNA regions. These are regions that vary in length from individual to individual and fall into two categories: 1) Variable Number of Tandem Repeats (VNTR) and 2) Short Tandem Repeats (STR). A VNTR is a region that is variably composed of a 15-70 base pair sequence, typically repeated 5-100 times. An STR is similar to a VNTR except that the repeated unit is only 2-4 nucleotides in length. By examining several different VNTRs or STRs from the same individual, investigators obtain a unique DNA fingerprint for that individual which is unlike that of any other person (except for an identical twin).

One VNTR known as D1S80, is present on chromosome 1 and contains a 16-nucleotide sequence which is variably repeated between 16 and 40 times. An individual who is homozygous for the D1S80 genotype will have equal repeat numbers on both homologues of chromosome 1, displaying a One VNTR known as D1S80, is present on chromosome 1 and contains a 16-nucleotide sequence which is variably repeated between 16 and 40 times. An individual who is homozygous for the D1S80 genotype will have equal repeat numbers on both homologues of chromosome 1, displaying a single PCR product following gel analysis (Figure 2A). More commonly, a person will be heterozygous, with differing D1S80 repeat numbers. Amplification of DNA from heterozygous individuals will result in two distinct PCR products (Figure 2B). For most applications, law enforcement agencies now use STRs as they are more easily amplified and thus require less starting DNA.

The objectives of this experiment are to isolate human DNA and compare DNA polymorphisms between individuals by PCR amplification and gel electrophoresis. In this experiment, each student will 1) extract his/her DNA from hair or cheek cells, 2) amplify DNA at the D1S80 locus by PCR, and 3) examine the PCR products on agarose gels.



Gel results are not drawn to scale.

Figure 2:
PCR Amplification Products of D1S80

ALU-HUMAN TYPING

The human genome consists of 2.9 billion base pairs of DNA. Of this total, only about 5% consists of exons which code for protein. Introns and other noncoding sequences make up the remainder; although some of these sequences may possess undiscovered functions, most appear to have none. Many of these non coding sequences appear to be self-replicating and are repeated hundreds or thousands of times throughout the genome. These repetitive sequences have been termed "selfish" or "parasitic" DNA, as they often

appear to possess no function except for their own reproduction. These repetitive elements account for more than 20 percent of the human genome.

In 1979, it was discovered that human DNA contains a 300 base pair repetitive element. Copies of this element contain a recognition site for the restriction enzyme Alu I, and were subsequently named Alu elements. Although Alu elements have been found in exons, most exist in introns and other non-coding regions. When Alu sequences do insert into protein-coding regions, disruption of the gene usually results, often causing harm to the organism. Alu sequences replicate through an RNA intermediate which is copied into a double-stranded DNA segment called a retrotransposon. The details of this process are not well understood. The retrotransposon then inserts elsewhere in the genome. It is also theorized that most Alu sequences are incapable of replication and that only a small number of "master genes" are duplicated to form new elements.

Although all humans (and other primates) possess hundreds of thousands of Alu elements, variations in their placement may occur. DNA sequences which vary between individuals are known as polymorphisms. For example, one person may possess an Alu insertion at a specific DNA locus, while another individual lacks that insertion. Furthermore, the insertion may be present or absent on each homologous chromosome and are called dimorphic. One such dimorphic Alu sequence is found on a section of chromosome 16 known as the PV92 locus. This section of DNA is 700 nucleotides in length. Insertion of the 300 base pair Alu element results in a length increase to 1000 base pairs.

One may test whether a person possesses an Alu insertion at the PV92 locus by amplification of the locus using the polymerase chain reaction (PCR). If a person is homozygous for the insertion, a gel of the PCR product will result in a single band at 700 base pairs (Figure 3A). If a person is heterozygous, i.e., possesses the insertion on one chromosome 16 homologue but not the other, two bands will be present following PCR. One band will be 400 base pairs and the other will be 700 base pairs (Figure 3B). If a person lacks the insertion on either chromosome homologue, that person is said to possess the null genotype and PCR will result in only one band at 400 base pairs (Figure 3C).

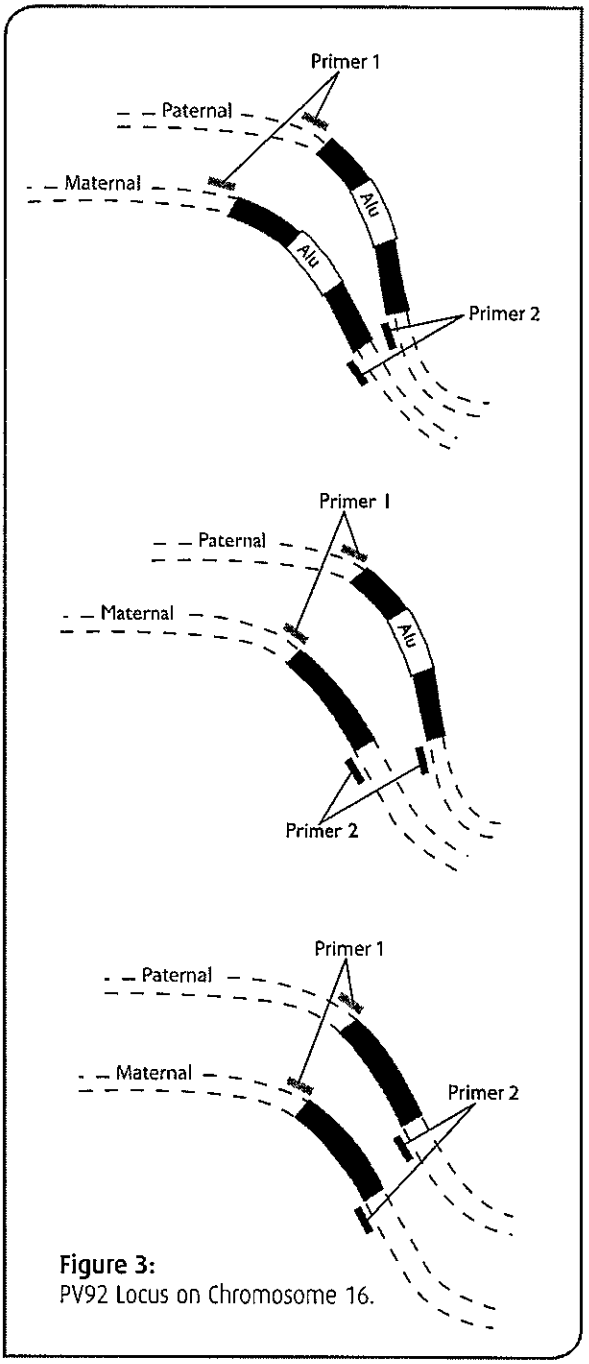


Figure 3:
PV92 Locus on Chromosome 16.

To purify DNA for this type of analysis, almost any tissue or body fluid (except urine) may be used. The most common sources of human DNA are hair and buccal cells (cheek cells). Cells to be tested must be treated (lysed) to release their DNA into solution. Following lysis, the cells are resuspended in a chelating agent, which removes cellular cations that inhibit PCR.

In this experiment, each student will extract his/her DNA from hair or cheek cells and amplify DNA at the PV92 locus by PCR. As a control, DNA purified from a cultured human cell line may be used. The PCR product(s) will then be examined on agarose gels to determine whether the student is homozygous (+/+), heterozygous (+/-), or null (-/-) for an Alu insertion at the locus. Objectives of this experiment are the isolation of human DNA and the comparison of DNA polymorphisms between individuals by PCR amplification and gel electrophoresis.

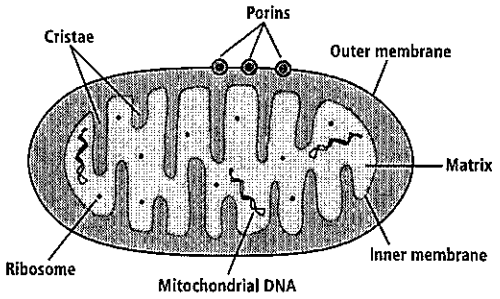


Figure 4:
Structure of a Mitochondrion.

MITOCHONDRIAL DNA TYPING

Mitochondria (plural for mitochondrion) are the energy-producing organelles of the cell. Mitochondria are generally oblong or egg-shaped. Both plant and animal cells possess mitochondria. The number of mitochondria per cell varies depending on the cell type, ranging from only a few in skin cells to thousands in skeletal muscle cells.

Unlike other organelles, mitochondria have two separate membranes. The outer membrane is fairly porous, possessing a protein called porin. The inner membrane, however, is highly impermeable to ions and is enriched in a rare, negatively charged phospholipid known as cardiolipin. The inner membrane is highly convoluted, with infoldings called cristae (Figure 4) that greatly increase the total membrane surface area. The inner membrane also contains the enzymes that catalyze cellular respiration, the process whereby energy is produced for the cell.

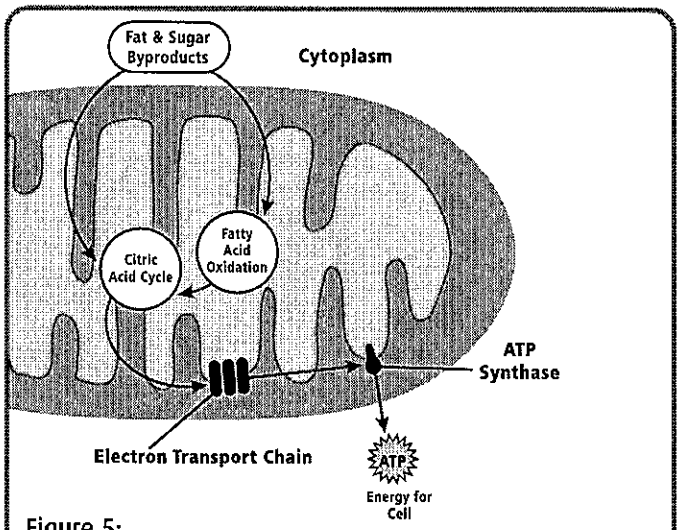


Figure 5:
Site of metabolic pathways that convert sugars and fats to energy.

The space inside the inner membrane is known as the matrix (Figure 4). Within the matrix and inner membrane, the chemical reactions that produce energy for the cell take place. As shown in Figure 5, sugars and fatty acids, broken down to two carbon units, enter a series of reactions known as the citric acid or Krebs cycle. Sugars are broken down in the cytoplasm while fatty acids are broken down in the mitochondria by a process known as ß (beta) oxidation. The citric acid cycle generates electrons that enter the electron transport chain, a cluster of protein complexes that reside in the inner membrane of the mitochondria. In the final step of energy production, protons generated by the electron transport chain flow through a pump known as ATP synthase, driving the production of ATP, the primary energy-containing molecule used in biological systems. This final energy-producing process is known as oxidative phosphorylation.

The DNA present in the matrix is distinct from the DNA found in the cell's nucleus. Mitochondrial DNA (mtDNA) is contained in a single circular chromosome, shown in Figure 6, that has been completely sequenced. The mitochondrial chromosome contains 16,569 base pairs of DNA and 37 genes. MtDNA encodes 13 polypeptides, all of which are sub-

units of the respiratory chain complex. This is shown on the map in Figure 6, illustrating the locations of genes that encode proteins in complexes I and IV of the electron transport chain. As shown, mtDNA also encodes mitochondrial ribosomal RNA and the ATP synthase, in addition to cytochrome B, another constituent of the electron transport chain. MtDNA encodes only part of the electron transport chain; nuclear DNA encodes the remaining complex subunits. One peculiarity is that mitochondrial protein synthesis uses a slightly different genetic code than cytoplasmic translation. As all cells possess only one nucleus but several hundred or thousand mitochondria, mtDNA is present in great excess over nuclear DNA in most cells. This relative abundance of mtDNA is taken advantage of by forensic investigators after obtaining crime scene specimens that are degraded or otherwise insufficient for nuclear DNA PCR analysis. The D-loop (Figure 6) has a high degree of variability between individuals and can be sequenced to demonstrate variations. MtDNA typing, however, cannot be used to conclusively link suspects to crime scenes; rather, it can be used to include or exclude suspects from further scrutiny.

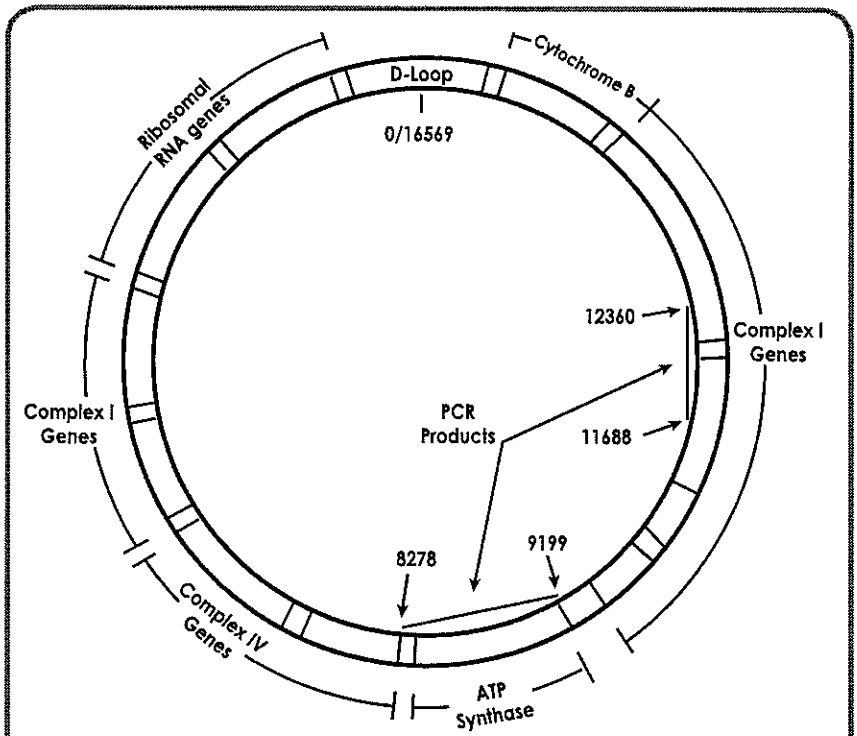


Figure 6: Genetic map of mitochondrial DNA.

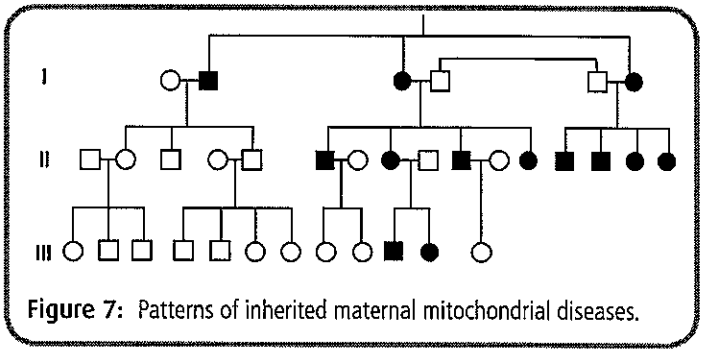


Figure 7: Patterns of inherited maternal mitochondrial diseases.

During the past twenty years, an ever-increasing number of diseases have been shown to be due to mitochondrial dysfunction. These disorders result when mitochondrial ATP generation is insufficient to meet energy needs in a particular tissue. Because muscle and nerve cells contain large numbers of mitochondria, these organ systems are most affected by mitochondrial dysfunction. Mitochondrial diseases may be due to mutations in mtDNA genes or mutations in nuclear genes that encode mitochondrial enzymes. Diseases caused by mtDNA mutations include the myopathies, diseases that affect various muscles and encephalomyopathies, which cause both muscular and neurological problems. Huntington's chorea, a devastating disease that results in dementia and loss of motor control, is caused by defects in oxidative phosphorylation and has been mapped to a mutation in nuclear DNA encoding a non-mitochondrial protein. Other diseases such as Alzheimer's and Parkinson's disease involve mitochondrial abnormalities, although it is unclear how these abnormalities relate to disease pathology. Mitochondria also appear to play roles in aging and in programmed cell death, also known as apoptosis.

Since mitochondria are present in the cytoplasm, they are inherited independently from the nucleus. A female egg cell possesses over 10,000 mitochondria, while a sperm cell has very few. Thus during fertilization, mitochondrial DNA is inherited almost exclusively from the mother. Although a small amount of paternal mtDNA is present in the fertilized egg, this DNA appears to be selectively destroyed by the newly fertilized egg. This pattern of inheritance of mtDNA is known as maternal inheritance. Maternal inheritance is indicated when all offspring, male and female, of the mother are afflicted with a specific condition (Figure 7). The severity of any particular mitochondrial disorder is highly variable, depending on the number of mutated mitochondria inherited from the mother (Figure 8).

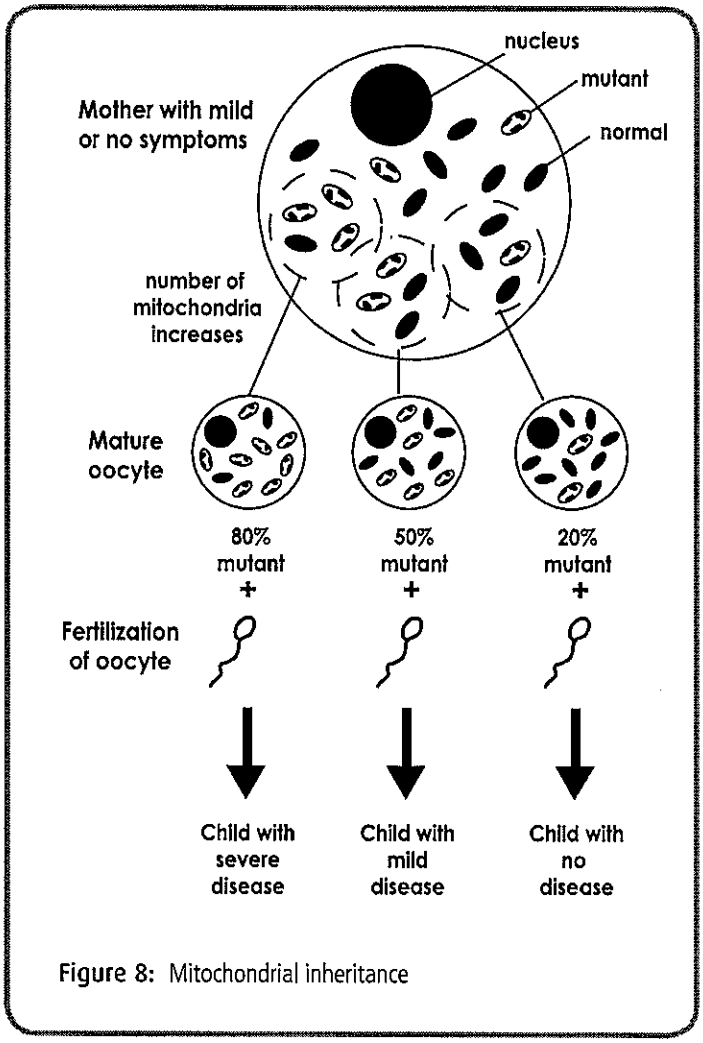


Figure 8: Mitochondrial inheritance

POLYMERASE CHAIN REACTION

To examine polymorphisms, the Polymerase Chain Reaction (PCR) is usually employed. PCR was invented in 1984 by Dr. Kary Mullis at the Cetus Corporation in California. The enormous utility of the PCR method is based on its ease of use and its ability to allow the amplification of small DNA fragments. For this groundbreaking technology, Mullis was awarded the Nobel Prize in Chemistry in 1993.

Before performing PCR, template DNA is extracted from various biological sources. Because PCR is very sensitive, only a few copies of the gene are required. Nevertheless, freshly isolated DNA will provide better amplification results than older DNA specimens that may have become degraded. In order to amplify the specific DNA or target sequence, two primers (short & synthetic DNA molecules) are designed to correspond to the ends of the target sequence.

To perform PCR, the template DNA and a molar excess of primers are mixed with the four "free" deoxynucleotides (dATP, dCTP, dGTP, and dTTP), and a thermostable DNA polymerase. The most commonly used DNA polymerase is *Taq* DNA polymerase. This enzyme, originally purified from a bacterium that inhabits hot springs, is stable at very high temperatures. These components (template DNA, primers, the four deoxynucleotides, and *Taq* DNA polymerase) are mixed with a buffer that contains Mg^{+2}, an essential cofactor for *Taq* polymerase. The PCR reaction mixture is subjected to sequential heating/cooling cycles at three different temperatures in a thermal cycler.

- In the first step, known as "denaturation", the mixture is heated to near boiling (94° C - 96° C) to "unzip" (or melt) the target DNA. The high temperature disrupts the hydrogen bonds between the two complementary DNA strands and causes their separation.

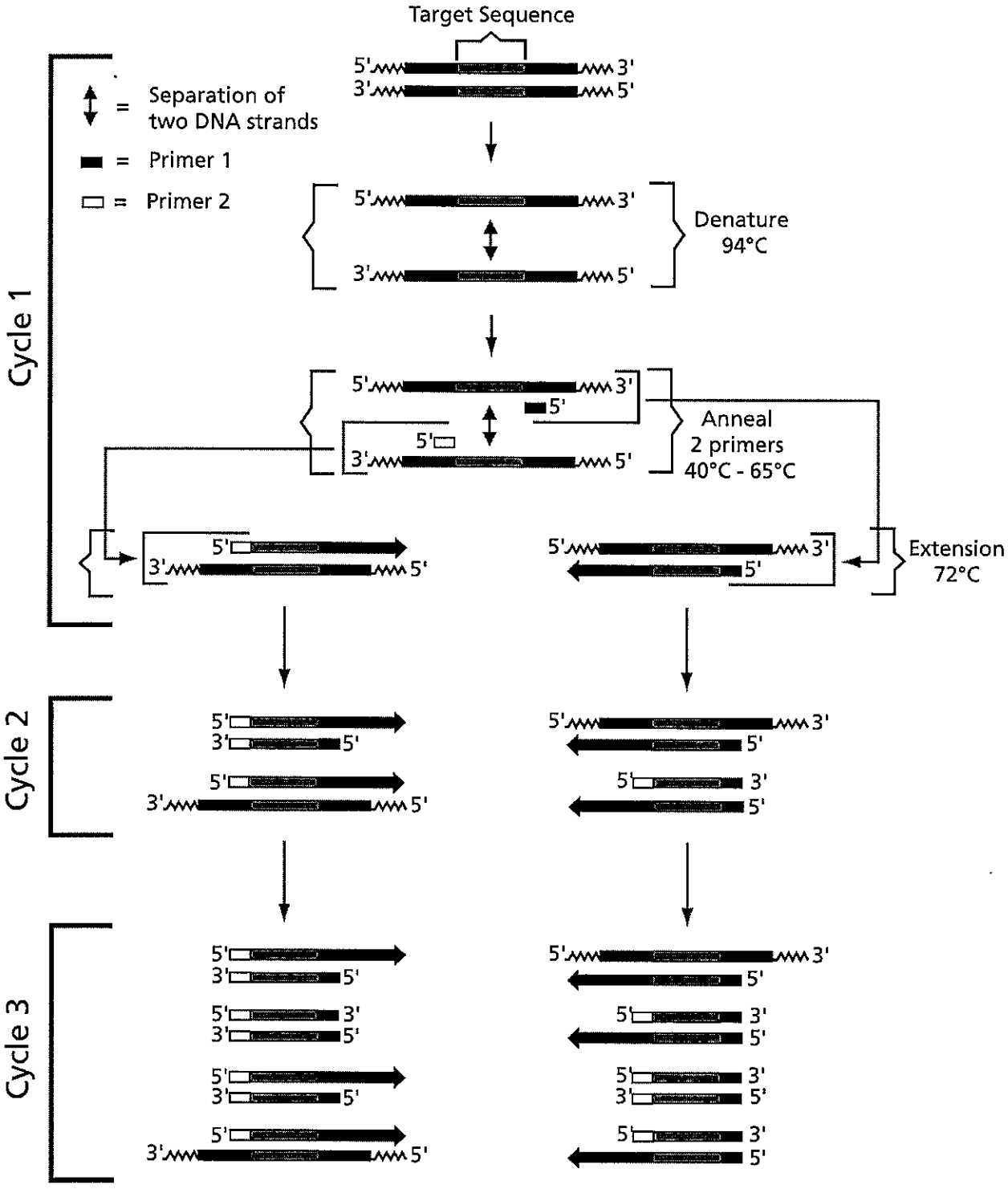


Figure 9:
Polymerase Chain Reaction

- In the second step, known as "annealing", the reaction mixture is cooled to 40 - 65° C, which allows the primers to base pair with the target DNA sequence.

- In the third step, known as "extension", the temperature is raised to 72°C. This is the optimal temperature at which *Taq* polymerase can add nucleotides to the hybridized primers to synthesize the new complementary strands.

These three steps - denaturation, annealing, and extension - constitute one PCR "cycle". Each PCR cycle doubles the amount of the target DNA in less than five minutes. In order to produce enough DNA for analysis, twenty to forty cycles may be required. To simplify this process, a specialized machine, called a "thermal cycler" or a "PCR machine", was created to rapidly heat and cool the samples.

In this experiment, students will extract their own DNA from hair or cheek cells. Using PCR, they will amplify DNA at (1) two separate regions of the mitochondrial chromosomes, (2) the PV92 locus, and (3) the D1S80 region of chromosome 1. As a control, DNA purified from a cultured human cell line may be used. The PCR product(s) will then be analyzed using agarose gel electrophoresis.