

EXPERIMENT 2: DETECTION OF HUMAN POLYMORPHIC ALLELES

DAY ONE: INTRODUCTION TO HUMAN DNA IDENTIFICATION THEORY

OBJECTIVES:

Today's laboratory will introduce you to the theory behind human DNA identification. By the end of the period you should be able to

- ❑ Describe the various applications of human DNA identification testing
- ❑ List the qualities of a good DNA identification "marker"
- ❑ Draw out a simple diagram of the polymerase chain reaction
- ❑ Describe the nature of short tandem repeats (STRs)
- ❑ Define "PRM" and calculate PRM values given allele frequency tables

INTRODUCTION:

Applications

Human DNA identification testing is used in a variety of settings. Perhaps the most widely recognized and publicized is its use in solving crimes (**DNA Forensics**), either to convict guilty individuals or exonerate innocent ones. In the past several years, over 100 people have been released from prison based on a reexamination of DNA evidence linked to their cases. Several in the United States have been released from Death Row. In addition, hundreds of violent crimes that would have previously gone unsolved have resulted in convictions.

The use of DNA has revolutionized the field of forensics. Prior to the discovery that subtle individual differences in DNA sequence could be exploited to identify individuals from one another, forensic scientists were limited to **eye-witness reports**, **fingerprinting**, and **blood protein markers**. Unfortunately, eye-witness reports are notoriously unreliable. Many of the people who were wrongly accused of crimes and later exonerated on the basis of DNA evidence were originally convicted on the basis of faulty eye-witness reports, including the highly suspect testimony of "jailhouse snitches" who were given reduced sentences for their testimony. In addition, many crimes do not have eye-witnesses, particularly when the victim has been killed or incapacitated. Moreover, if an assailant wears a disguise, the victim may not be able to pick him out of a police line-up or may identify the wrong person.

Fingerprints are excellent evidence when they can be found. However, many criminals have learned to wear gloves and/or wipe away their prints. Additionally, fingerprints can often be explained away by acquaintance, particularly when the suspect is a family member, co-worker, or close friend of the victim. For example, if a husband kills his wife, the presence of his fingerprints in the house cannot be used as evidence of his involvement in the murder.

Blood typing became possible in the early 1960s and was largely used to exclude suspects. If the victim was Type A and there was also Type B blood at the scene, a suspect with Type A, AB, or O could be excluded as the source of the second blood type. However, even when several blood protein systems were used together, the probability of a random match (**PRM**) in the population as a whole was about 1 in 1,000. While this might seem to be powerful evidence, a quick calculation reveals that 250 persons in a city of 250,000 would share this same suite of markers. Most juries were not willing to convict on the

basis of this kind of evidence alone. Another problem with blood protein markers is that proteins in general are not very stable molecules; they degrade quickly and cannot be retrieved from decomposed remains. Blood typing was therefore limited to those cases where evidence was found soon after the crime was committed. In addition, the typing technology required that a significant amount of blood be available for testing. Trace evidence could not be analyzed, and this further reduced the utility of the technique in forensics.

One of the main reasons that DNA identification testing has had such a huge impact is that it can be performed on all types of body fluids and tissues. While only certain types of body fluids contain the blood protein markers, all cells contain the same DNA. Moreover, DNA identification testing can be performed on **trace amounts** of such samples, even when they are **decomposed**. DNA profiles can be generated from bones exposed to the elements for over a year (e.g. Chandra Levy remains found in a Washington DC park in 2002) from the saliva left on a toothbrush (e.g. Xiana Fairchild child abduction case in August of 2000), from spittle (e.g. Peterson "Duck Robber", 1998), dried semen (e.g. Monica Lewinsky's dress), and from hair, sweat, and even dandruff.

Paternity testing is another area where DNA testing has been extremely useful. Maternal self-reports of paternity are highly suspect. Recent studies by genetic counselors reveal that about 10% of all men who think they are the fathers of a particular child are not, including husbands and live-in boyfriends. DNA evidence has allowed paternity to be established with a high degree of accuracy, and is widely accepted in court to establish child custody and support, immigration claims, and to resolve insurance and inheritance disputes. The testing can also be used to confirm family relationships after adoption or separation. American servicemen who fathered children with native women during the Vietnam War have been successful in finding and confirming their relationships with their offspring even though they had never seen them before. To perform a paternity test, blood samples are usually taken from the alleged father, the mother, and the child. Since each child receives one allele at each DNA locus from each parent, the mother's contribution can be easily seen in the child's profile. The other allele carried by the child must have come from the father, and the alleged father's profile is checked to see if he could have contributed the second allele at each locus. When enough markers are used, paternity can be established with a high degree of certainty (greater than 99%).

An extension of paternity testing is **relationship testing**, which can establish familial connections through grandparents, mothers, or siblings. This type of testing was used to reunite family members after Argentina's "Dirty War", in El Salvador in the 1990s, and is currently being used in Africa to return children enslaved on coffee plantations to their parents and/or extended families. Relationship testing is also used to identify bodies or body parts after man-made or natural disasters that render bodies unrecognizable. Examples include the remains of passengers on the TWA Flight 800 flight that mysteriously exploded off New York's Long Island in 1996, the victims of the 911 attack on the World Trade Center, and the victims of numerous fires, floods, mudslides, hurricanes, and earthquakes. Generally, to perform relationship testing, a DNA profile is performed on the tissue or bone of the body or body part and compared to the profile of close relatives, using an approach similar to that of paternity testing.

The Polymerase Chain Reaction

The breakthrough technology that allowed the analysis of trace and decomposed DNA evidence is the **polymerase chain reaction (PCR)**. The PCR is a technique for copying or *amplifying* DNA in a test tube. Discovered by Kari Mullis at Cetus Corporation in the 1980s, it revolutionized the field of

molecular biology and won Mullis a Nobel Prize in Chemistry in 1993.

The power of PCR lies in its ability to make billions of copies of a *specific* DNA sequence of interest. Consider the following analogy. Imagine that you have a pile of a thousand pennies. Among that stack is a particular penny of interest that is very valuable and that you want to study further. However, to study that penny you first need to find it among all the other pennies in the stack. That could be a very tedious task without some type of tool to help you out. Now imagine that you have a "penny copying machine" that can somehow seek out the special penny and make a billion copies of it. The machine doesn't isolate the special penny you're after from all the other pennies in the stack. However, it makes so many copies of it that you now have a huge stack of special pennies to study and the stack contains only a very small percent of pennies that are ordinary.

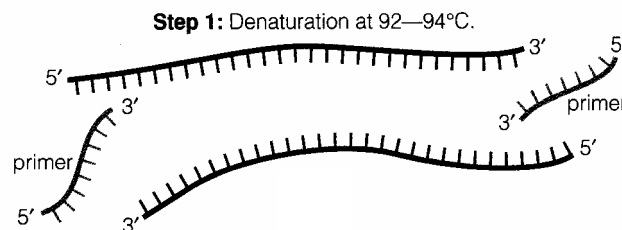
Genomes are like stacks of pennies. The "pennies" are DNA sequences. However, a researcher is usually interested in studying only one specific DNA sequence at a time (e.g. the human sickle cell anemia gene or the regulatory region of a bacterial gene involved in glucose metabolism). PCR is like the "penny copying machine" because it makes copies of a specific DNA sequence. The PCR reaction doesn't eliminate the rest of the DNA but it overwhelms it by making billions of copies of the sequence of interest. Thus, the researcher has plenty of his or her specific DNA sequence to study and encounters only minimal interference from the rest of the DNA (the "ordinary" pennies).

PCR works by mimicking the replication of DNA that takes place in cells. However, rather than producing two identical copies of the DNA (as a cell does prior to mitosis and subsequent cell division), PCR produces literally billions of copies of the DNA. In addition, a cell replicates its *entire* genome (which can be millions or even billions of base pairs in length) whereas PCR amplifies short, highly specific stretches of DNA (usually 2,000 base pairs or less).

The PCR reaction involves 4 steps: **(1) denaturation**, **(2) primer annealing**, **(3) primer extension**, and **(4) cycling**. The molecular events happening at each step are described below and the entire process is illustrated in the diagram on the next page. Note that the end result is many copies of a specific small segment of DNA defined by the sequences of the primers that are used in the reaction.

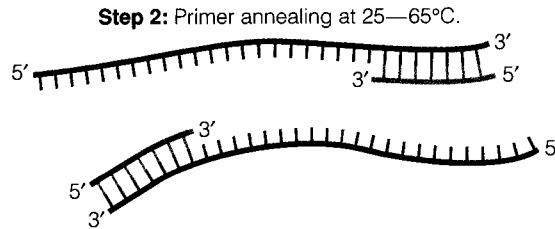
STEP 1: **Denature Template DNA.** In order to replicate the DNA, it must first be *denatured* (made single stranded) to expose the single strands to DNA polymerase. In cells, helicases unwind the double-stranded DNA. In PCR, **heat** is used to accomplish this task.

The reaction tubes are raised to about 95 degrees C for about one minute to denature the DNA. It is critical that the template DNA be completely denatured prior to step 2.



STEP 2: **Anneal Primers.** Once the DNA has been denatured, the primers will find their single-stranded complementary sequences and *hybridize*. The primers are required to specify to the polymerase which region of the DNA is to be copied and to provide the free 3'-

OH groups required by the polymerase to start synthesizing a new DNA strand. (Remember that in cells, this function is served by short RNA "primers" laid down by primase.)



The annealing temperature chosen depends largely on the sequence of the primers. For greatest specificity, primers are annealed at or near their **melting temperature** - the temperature at which the primer will just barely "sit down" on its complementary sequence without melting off. Since A-T base pairs are less stable than G-C base pairs, A-T rich primers tend to melt off the DNA more easily and have lower melting temperatures than do G-C rich primers. As a rule of thumb, assign 2 degrees C for each A or T in the primer and 4 degrees C for each G or C to determine a primer's melting temperature. Although, in practice, computer programs are used to predict annealing temperatures and search known sequences for ideal primer sites.

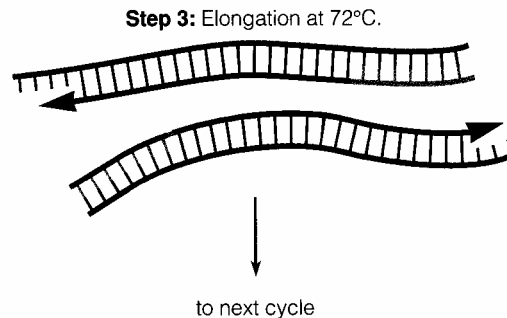
For example, what melting temperature would you assign to the following primer?

5'- aac tta agc taa cac cgt agg -3'

Estimated Melting Temperature (T_m): _____ **degrees C**

Standard annealing temperatures range from **50 - 70 degrees C**.

STEP 3: Extend Primers. After the primers have annealed in their correct locations, DNA polymerase recognizes their 3' ends and begins extending both primers in the 5' to 3' direction. Bases are added according to the standard base-pairing rules, where A pairs with T and G with C. **Note that as this happens, the primers on the two strands are extended toward one another across the target sequence.** They will not "meet" however, since the primers are being extended on two different strands of the original parent DNA molecule. Since the most favorable temperature for heat stable DNA polymerases is about **72 degrees C**, the extension step is usually run at about this temperature.



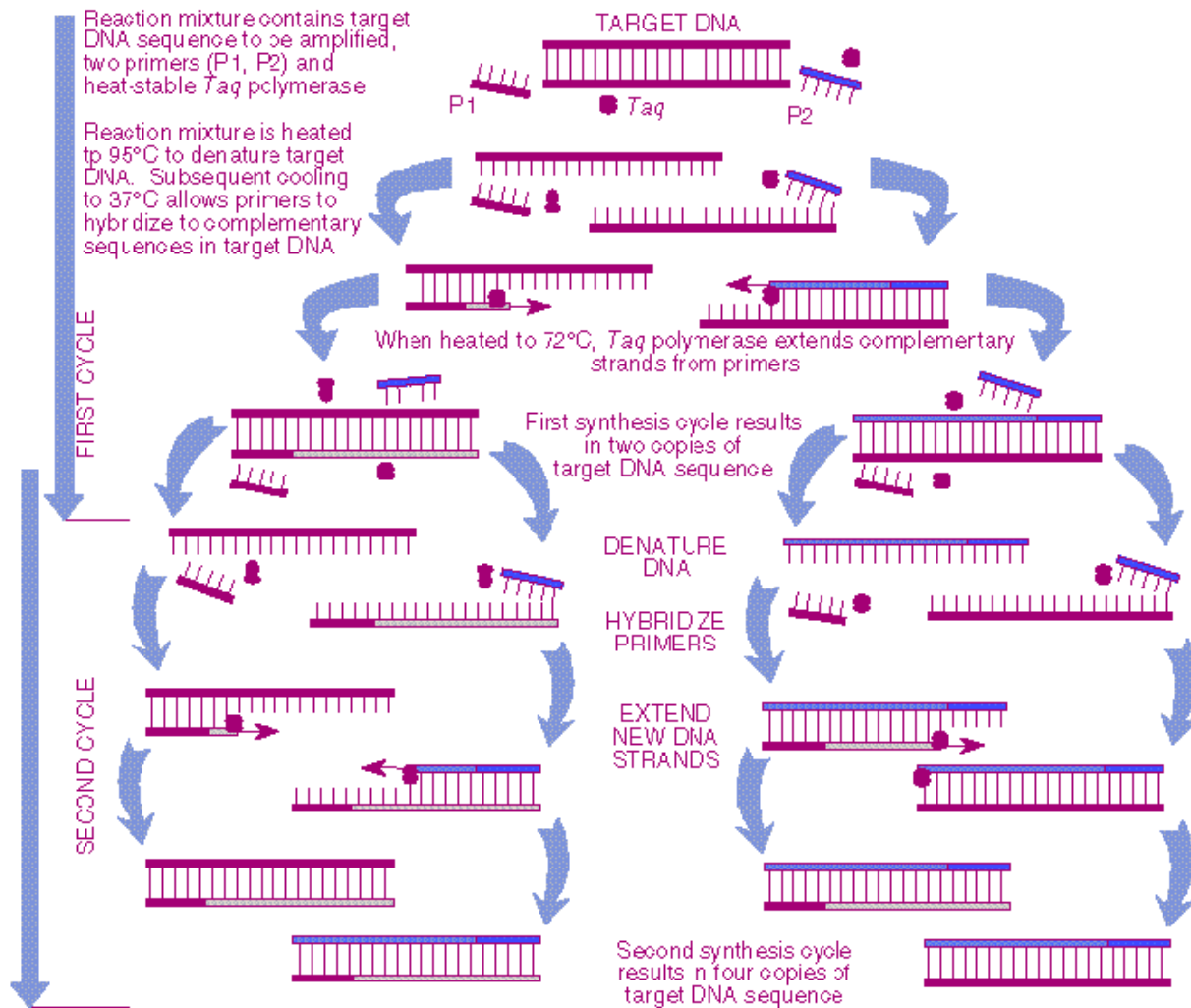
STEP 4: Cycling. At the conclusion of step 3, the DNA target sequence has been copied one time

to yield a total of two target DNA molecules. By itself, this is not very interesting since the DNA has only been amplified 2-fold. However, notice what happens when the steps are repeated a second time: 4 target molecules result. After cycle 3, there are 8 molecules; after cycle 4 there are 16, etc. In other words, the number of copies of the target DNA sequence increases *exponentially*, so that after 30 cycles there are about 2^{30} copies of the sequence - a *billion-fold* amplification!

Cycling is made possible in part because a vast excess of primers and dNTPs are added at the beginning of the reaction. Therefore, the reaction does not stall due to lack of raw materials. In addition, the thermal stable polymerase eliminates the need to add fresh DNA polymerase at the beginning of each extension. The heat stable polymerase can survive the cyclical denaturation steps (95 degrees) that would permanently destroy the activity of most DNA polymerases.

ORNL-DWG 91M-17476

Figure 4-1. The Polymerase Chain Reaction. (From <http://aidshistory.nih.gov/imgarchive/pcr.html>)



A standard PCR reaction requires the following ingredients:

- 1. Template DNA** - This DNA is the *source* of the DNA sequence that you seek to amplify. For example, if you want to amplify part of the human β globin gene your template DNA would need to be human. You would obtain human cells from a cheek swab, a blood sample, etc. and then apply one of a variety of techniques that are routinely used to isolate and purify DNA from intact animal cells. This would then be your DNA template for the PCR reaction.
- 2. DNA primers** - Within the template DNA is the *target sequence* you wish to amplify. In the above example, the target sequence would be a portion of the human β globin gene. In order to amplify that region of the DNA specifically, you need to design short, single-stranded DNA molecules called *primers*. These primers are made synthetically (they are usually ordered from a company for about \$1 per base) and are designed so that they will base pair with the upper and lower strands of the region immediately *flanking* the target DNA sequence. For example, suppose that you wanted to amplify the target sequence shown below.

5'-AGGACTTAGCCCTTAAATCG-portion of β globin gene-TTCCATCAGAGACTTAGACA-3'
3'-TCCTGAATCGGGAATTTAGC-portion of β globin gene-AAGGTAGTCTCTGAATCTGT-5'

To amplify this sequence by PCR, one of your primers would need to be complementary to the 3' end of the bottom strand and the other to the 3' end of the top strand, as shown at the top of the next page.

↓ Melt DNA and add primers

5'-AGGACTTAGCCCTTAAATCG-portion of β globin gene-TTCCATCAGAGACTTAGACA-3'
3'- aaggtagtctctgaatctgt-5'

5'-aggacttagcccttaaatacg-3'
3'-TCCTGAATCGGGAATTTAGC-portion of β globin gene-AAGGTAGTCTCTGAATCTGT-5'

- 3. dNTPs** - Nucleotide triphosphates are the raw ingredients for DNA replication. DNA polymerase cleaves off two of the phosphates to provide energy for the polymerization reaction and then adds the appropriate nucleotide monophosphate to the 3' end of the growing DNA daughter chain. (The "N" stands for either A, G, C, or T.)
- 4. DNA polymerase** - This enzyme catalyzes the formation of a new strand of DNA using dNTPs and an existing DNA strand as a template. As mentioned above, in PCR, the polymerase used is usually able to tolerate high temperatures (up to 100 degrees C) without denaturing. Thermal stable polymerases are isolated from heat tolerant bacteria that live in thermal vents.
- 5. Salts/buffer** - DNA polymerases require magnesium so PCR reaction buffers always contain magnesium as well as a buffering agent.

STANDARD THERMOCYCLER PROGRAM FOR PCR:

Step 1	95 degrees C	60 seconds	Step 4	Cycle Steps 1-3, 30 times
Step 2	55 degrees C	60 seconds	Step 5	End

Step 3 72 degrees C 60 seconds

What Makes a Good DNA Marker?

Polymorphism

There are three requirements for a good DNA identity marker. First it must be **polymorphic** - that is, there must be many (*poly*) forms (*morphs*) of the marker found in the population. As you should remember, different forms of a gene are called **alleles**. So, a good DNA marker must have many different alleles. Although each person can only have two alleles for a gene, there must be many alleles in the population *as a whole*.

One example of a polymorphic DNA marker is the **ABO blood group gene**. This gene codes for an enzyme that adds carbohydrate groups to a protein located on the surface of red blood cells. If the "A" allele is present, one type of carbohydrate group is added, and if the "B" allele is present, a slightly different carbohydrate group is added. If *both* A and B are present, *both* types of carbohydrate groups are added (creating the rare "AB" blood type). "O" is a null allele - that is, it codes for a non-functional enzyme and no carbohydrate groups are added.

However, the ABO blood system only has three different alleles so it is not very polymorphic. With three alleles, there are only six possible genotypes: AB, AA, AO, BO, BB, and OO. As a general rule, if the number all alleles = x, the number of genotypes = $[(x+1)^2 - (x+1)]/2$ as shown in the table below. From this table, it is easy to see that the more polymorphic a locus is, the more different kinds of genotypes are possible.

# alleles in the population (x)	# genotypes in the population $[(x+1)^2 - (x+1)]/2$
1	1
2	3
3	6
4	10
5	15
6	21
7	28

Equal Allele Frequencies

Second, the **alleles must have roughly equal frequencies**. If there are several alleles at a locus but only one is common, the marker will be of little use. For example, there are over 200 different alleles at the cystic fibrosis locus. However, the only common allele is the normal, or wildtype form. The others cause disease and are selected against. Such a locus, while highly polymorphic, would not make a good DNA marker because almost everyone has the normal form of the gene.

Detectable by PCR

Third, the best DNA markers can be assayed by PCR. This is especially critical in forensics since PCR *amplifies* the available DNA. Since many crime scenes yield only trace amounts of biological evidence or the evidence is degraded due to exposure, DNA profiles can only be obtained if a PCR-based method is used.

The STR Marker System

STRs (Short Iandem Repeats) are ideal DNA identity markers because they are highly polymorphic, no single allele is particularly common, and they are easily amplified by PCR. As a result, STR analysis is now the system of choice in forensics and paternity laboratories around the world.

A typical STR locus is **D7S820**, located on chromosome 7. This site consists of a section of DNA in which a 4-base-pair sequence, or **tetramer**, is repeated over and over again, as shown below.

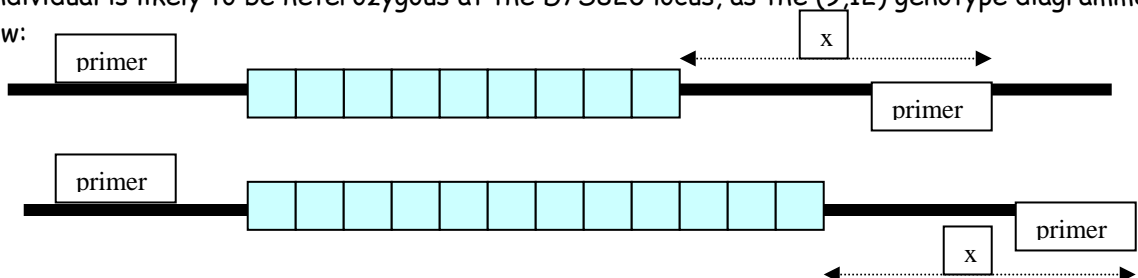
```
aatTTTTgta tTTTTttag agacgggggtt tcaccatggt ggtcaggctg
actatggagt tattttaagg ttaatataata taaagggtat gatagaacac
ttgtcatagt ttagaacgaa ctaacgatag atagatagat agatagatag
atagatagat agatagatag atagatagat agtttt tttttatctc
actaaatagt ctatagtaaa catttaatta ccaatatttg gtgcaattct
gtcaatgagg ataaatgtgg aatcgttata attcttaaga atatatattc
cctctgagtt tttgatacct cagattttaa ggcc
```

In this particular example, the tetramer "**gata**" is repeated 14 times. However, this is not always the case. In fact, it can be repeated anywhere from 6 to 15 times at this locus. Each of these repeat numbers represents a different D7S820 allele. However, having 14 repeats is not particularly common or particularly rare. In U.S. Caucasians, the frequencies of the different alleles range from 0 (alleles 6 and 15 are not routinely found in U.S. Caucasians) to 26.2% (allele 10), as shown in the table below.

ALLELE FREQUENCIES AT D7S820

ALLELE (# of repeats)	FREQUENCY (U.S. Caucasians)
6	0.000
7	0.018
8	0.162
9	0.131
10	0.262
11	0.250
12	0.128
13	0.046
14	0.003
15	0.000

Note that in the general human population, where all the alleles are present, the total number of possible genotypes is $[(15+1)^2 - (15+1)]/2 = 120$. In fact, most alleles are so relatively uncommon that an individual is likely to be heterozygous at the D7S820 locus, as the (9,12) genotype diagrammed below:

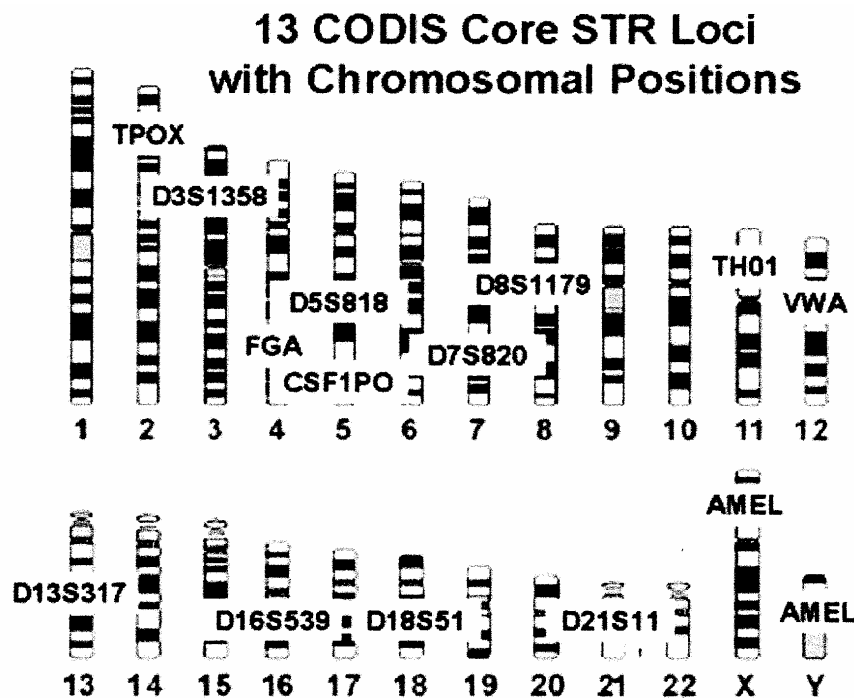


Thirteen polymorphic STR loci have been developed for human identity testing. However, there are lots more loci to choose from if need be - over 500,000 STR sites are thought to exist, spread throughout the human genome!

The 13 STR loci are distributed across several different chromosomes, as shown below. In a typical DNA testing protocol, all 13 of these loci are amplified by PCR, using primers that flank the STR repeat sites. Note, in the figure above, the distance between the end of the repeat and primer is the same (x) in both alleles. The smaller the number of repeats at a particular locus, the smaller the PCR product that is generated. Thus, the alleles are detected by **PCR product size differences due to differences in the number of repeats**. However, between STR loci, both the size of the repeats and the distance to primer will vary, therefore it is possible that some alleles at two different loci will be the same size.

The primers used in the procedure carry **fluorescent tags** that aid in the detection of the PCR products during their separation by electrophoresis. The detection process is complex, but is described briefly on the next two pages.

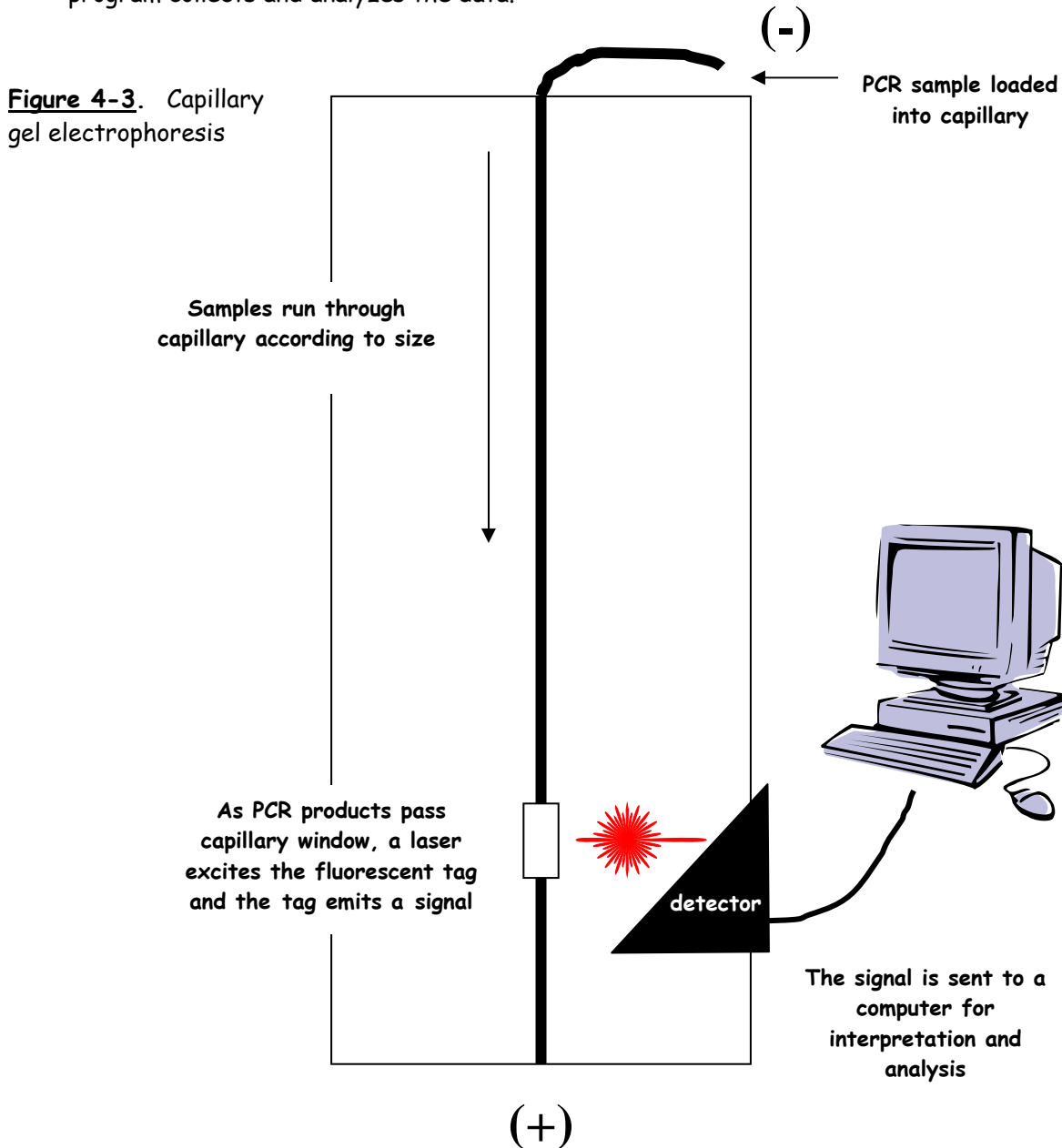
Figure 4-2. Locations of the 13 human DNA identification loci in the human genome. From <http://www.cstl.nist.gov/biotech/strbase/ppt/intro.pdf>



STR PCR products are separated by a specialized form of gel electrophoresis called **capillary gel electrophoresis**. Rather than loading the samples into a slab gel, the samples are run through a thin tube that is filled with a polymer that works like agarose to sieve the larger products from the smaller ones. Thus, the smaller PCR products run through the gel more quickly than the larger ones - just like in agarose slab gel electrophoresis.

Capillary gel electrophoresis has the advantage that diffusion and convection forces are reduced, allowing the samples to be run at high voltages and thus in shorter amounts of time. Capillary gels can also be loaded automatically, saving time and money.

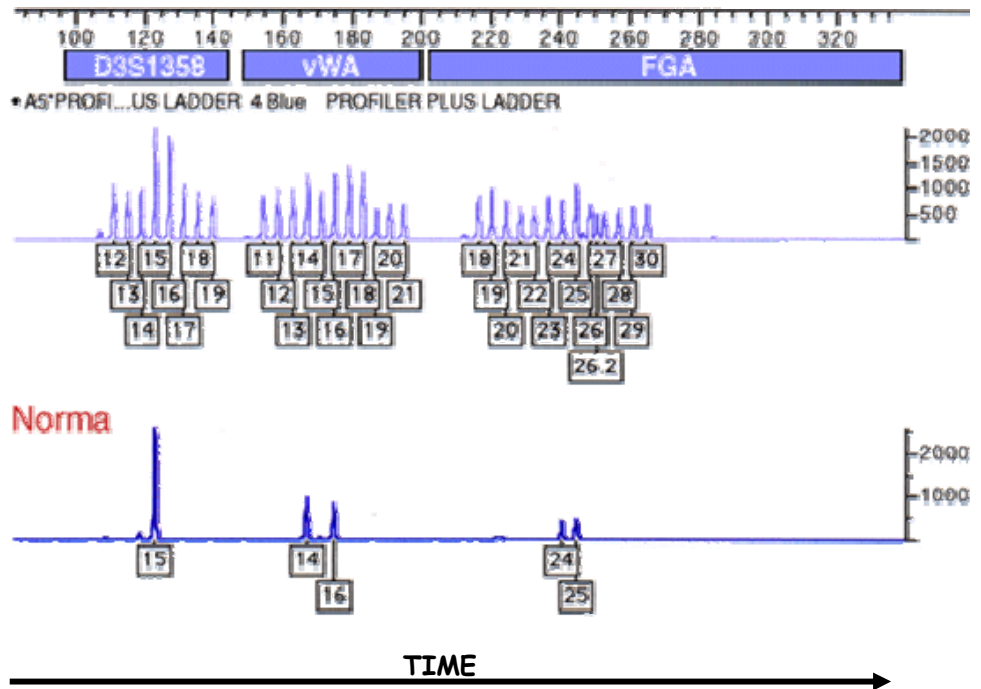
The process is diagrammed below. Of particular interest is the detection apparatus, which uses a laser to excite the fluorescent tag on the PCR products as they pass by a "window" near the end of the capillary. A computer then picks up the wavelength and amplitude of the signal digitally and a software program collects and analyzes the data.



As the computer receives the data, it generates an **electropherogram** that charts the wavelength and amplitude of the signal against the time it was received.

The sample profile below was generated across 3 STR loci from a woman named "Norma". At the top is an internal ladder that shows the gel mobilities (and thus the sizes) of the common alleles in the human population. Norma is homozygous for the 15 allele at the D3S1358 locus, is a (14,16) at the vWA locus, and is a (24,25) at the FGA locus. The numbers at the top of the figure are the actual gel mobilities in DNA base pairs. Although alleles at other loci (not shown) may have mobilities that overlap with loci shown below, they will have been assigned primers of a different color. Because the computer is able to distinguish *size and color*, each allele can be unambiguously identified.

Figure 4-4. Typical electropherogram generated during human STR analysis. From http://www.biology.arizona.edu/human_bio/activities/blackett2/str_analysis.html



Generating PRM Values for STR Profiles

Once an STR profile has been generated, the next step is to assign a PRM (probability of a random match) value to the overall genotype. This number tells you the frequency of the STR profile in the population as a whole. This is accomplished by calculating the PRM values at each locus *individually* and then multiplying those frequencies together.

Genotype frequencies at individual loci are calculated using the following two equations:

Homozygotes: $PRM = p^2 + [(p)(1-p)(0.01)]$

Heterozygotes: $PRM = 2 pq$

(where *p* and *q* are the frequencies of the alleles)

In Experiment 2, we will be analyzing a paternity case to determine which of two men is the father of a little boy. We will also analyze the data to determine if either or both men can be excluded as the father of the child and, if possible, to identify the true father of the child. Finally we will be applying the formulae above to generate a PRM value for the child, such as would be generated in a criminal case.

DAY TWO: ISOLATION OF CHEEK CELL DNA FROM SWABS

OBJECTIVES:

Today you will extract the DNA from cheek cell samples that have been donated for a paternity test from a mother, her child, and two potential fathers. The photographs of the donors are displayed on the front bench of the lab. At the end of today's lab you should be able to describe:

- ❑ **The steps used to extract DNA from cheek cell swabs**
- ❑ **The importance of reducing contamination during DNA extraction for PCR**

THINGS TO DO:

1. To avoid contamination in your PCR reaction, **WEAR GLOVES THROUGHOUT TODAY'S LAB. Put them on now.** (Boxes of gloves are located on the front bench.)
2. After putting on gloves and before beginning the experiment, each lab pair should gather the following items from the front bench.
 - one sterile **scalpel**
 - one tube of **PBS***
 - one tube of **Protease Stock Solution***
 - one tube of **AL Solution***
 - one tube of **100% Ethanol***
 - one tube of **Buffer AW1***
 - one tube **Buffer AW2***
 - one tube of **Buffer AE***
 - one **QIAMP Spin Column***
 - one sterile **2-ml collection tubes**
 - one sterile **1.5-ml collection tube**
 - a **cheek cell swab** from the donor you have been assigned to test
3. Carefully open the package containing the sterile scalpel. The inside surface of this package is sterile and should be used as a working surface to remove the cotton tip from your donor's Q-tip in Step 4.
4. Open the tube containing your donor's Q-tip. Gently grip the wooden stick and withdraw the Q-tip from the tube. **DO NOT TOUCH THE TIP OF THE Q-TIP TO ANYTHING BUT THE INSIDE OF THE SCALPEL PACKAGING.** Place the cotton end of the Q-tip onto the inside of the scalpel packaging and use the scalpel to scrape the cotton off the wood. Once you have completed this task, you can discard the wooden stick in your Biohazardous Waste Container.
5. Using a sterile yellow micropipettor tip, carefully transfer the cotton swab back into the 1.5-ml microcentrifuge tube that carried the original donor Q-tip.
6. Add **400 µl of PBS** to the tube containing the cotton swab.
7. Before proceeding, make sure that you know where the vortexers are located and how to use them. Then add **20 µl of Protease Stock Solution** and **400 µl of Buffer AL** to the sample. **CLOSE THE LID OF THE TUBE AND MIX IMMEDIATELY BY VORTEXING FOR 15 SECONDS.**

8. Incubate the tube at 56°C for 10 minutes. After incubation, briefly centrifuge the tube to remove any drops from the inside of the lid.
9. Add **400 µl of 100% Ethanol** to the sample and mix again by vortexing. Briefly centrifuge to remove any drops from the inside of the lid.
10. Place the **QIAMP Spin Column** into the sterile 2-ml collection tube. Then carefully apply **700 µl** of the mixture from Step 9 to the column without wetting the rim. Close the cap and centrifuge the tube at **8000 rpm for 1 minute**.
11. After centrifugation, empty the collection tube into the collection beaker at your lab bench and place the spin column back into the emptied 2 ml collection tube. **DO NOT DISPOSE OF THIS FLUID DOWN THE SINK. IT IS HIGHLY TOXIC.**
12. Repeat Steps 10 and 11, applying up to another **700 µl** to the column, closing the cap, centrifuging, and discarding the filtrate in the collection beaker. Once again, place the spin column back into the emptied 2 ml collection tube.
13. Carefully open the QIAMP spin column lid and add **500 µl Buffer AW1** without wetting the rim. Close the cap and centrifuge at **8000 rpm for 1 minute**. Discard the filtrate as before and return the column to the emptied 2-ml collection tube.
14. Carefully open the QIAMP spin column and add **500 µl Buffer AW2** without wetting the rim. Close the cap and centrifuge at full speed (**14000 rpm**) for **3 minutes**.
15. Discard the filtrate into the collection beaker. Replace the spin column into the collection tube. Centrifuge the column for **1 minute at 14000 rpm**.
16. Discard the collection tube and filtrate and place the QIAMP spin column in a **NEW 1.5-ml collection tube**. Carefully open the QIAMP spin column and add **150 µl Buffer AE**. Close the lid and incubate at room temperature for **1 minute**. Then centrifuge at **8000 rpm for 1 minute**.
17. The DNA has now been eluted from the column and is in the 1.5-ml collection tube. **DO NOT DISCARD THE FILTRATE OR YOU WILL LOSE YOUR DNA!** Remove the column from the 1.5-ml collection tube and close the lid. Label the tube with the **donor's first name, your initials, and your lab section number**. The QIAMP spin column can now be discarded in the Biohazard Waste Container on your bench. **SAVE THE LABELED 1.5-ML TUBE** and put it in the rack on the front bench that has been provided to collect your DNA samples.

***DNA Isolation Reagents:**

- PBS - phosphate buffered saline; standard biological buffer.
- AL buffer - lysozyme to break open cells.
- Protease - break down all protein in amino acids (chromatin is mostly protein).
- Ethanol - dissolves lipid and causes DNA to bind to spin column matrix.
- Spin column - contains silica salt (thin opaque layer at bottom of column) to which the DNA and other organic molecules reversibly bind (i.e., in the presence of the appropriate buffer).
- AW1/AW2 - elute (allow to pass through) non-DNA organics from spin column (i.e., clean up your sample)
- AE - elute purified DNA from spin column.

DAY THREE: SET UP PCR REACTIONS

OBJECTIVES:

In today's lab, you and your lab partner will set up a PCR reaction to amplify the human identification STR loci from the cheek cell DNA you isolated last time. Half of the lab pairs will amplify the 9 Profiler loci and the other half will amplify the 6 Cofiler loci. Your instructor will assign your lab pair to one of these reactions. We will also be taking a "field trip" to the basement of Sequoia Hall to view the capillary gel electrophoresis instrument (ABI Genetic Analyzer) and discuss the theory behind its use. At the end of today's lab you should be able to:

- ❑ Describe the steps involved in setting up a PCR reaction
- ❑ Describe the purpose of a thermocycler during PCR
- ❑ Appreciate the importance of clean technique when setting up PCR reactions.
- ❑ Describe the theory behind CGE separation of STR PCR fragments

THINGS TO DO:

1. **PUT ON GLOVES TO PREVENT CONTAMINATION OF YOUR DONOR DNA SAMPLE WITH YOUR OWN DNA.**
2. Before starting, each lab pair should gather the following items from the front bench:
 - one **0.2-mL thin-walled PCR tubes** (FRAGILE; do not squeeze!)
 - one tube **PCR Reaction Mix** (Profiler or Cofiler as assigned by your instructor)
 - one tube **Primer Set** (Profiler or Cofiler as assigned by your instructor)
 - one sterile **1.5 mL microcentrifuge tube**
 - the tube of **extracted DNA from your donor**
3. Vortex the PCR Reaction Mix, Primer Set for 5 seconds. Spin the tubes briefly at 14 rpm to remove any liquid from the caps.
4. Prepare the PCR Master Mix by adding the following volumes of reagents to a sterile 1.5-ml microcentrifuge tube:
 - 21 μ L PCR Reaction Mix
 - 1 μ L Amplitaq Gold DNA Polymerase (add this at the front bench)
 - 11 μ L Primer Set
5. Close the cap and mix thoroughly by vortexing for 5 seconds.
6. Spin the tube briefly in a microcentrifuge tube to remove any liquid from the cap.
7. Dispense 30 μ L of Master Mix into 0.2-mL thin-walled PCR tube.
8. Label the **sides of the tube** and the **hinge region of the cap of the tube** with your **lab bench number**.
9. Add 20 μ L of your donor's extracted DNA to the PCR tube.
10. Take your PCR tube to the thermocycler, where your instructor will load it into the instrument.

The thermocycler has been set to run the following program:

Step 1:	95 °C	11 minutes
Step 2:	94 °C	1 minute
Step 3:	59 °C	1 minute
Step 4:	72 °C	1 minute
Step 5:		Cycle steps 2-4 28 times
Step 6:	60 °C	45 minutes
Step 7:	10 °C	Hold indefinitely

When all the PCR reactions have been placed in the thermocycler, we will go on a "field trip" to the MBIG (Molecular Biology Interdisciplinary Group) Facility in the basement of Sequoia Hall to view the ABI 310 Genetic Analyzer and discuss its use. **PLEASE LEAVE YOUR BACKPACK AND PERSONAL ITEMS IN THE GENETICS LAB, WHICH WILL BE LOCKED DURING THE FIELD TRIP.** At the end of the field trip, you can retrieve your items from the lab.

DAY FOUR: ANALYSIS OF CGE RUN AND PROFILING RESULTS

OBJECTIVES:

After the PCR reactions were finished running the in thermocycler after last week's laboratory, they were quantified and then separated on the 310 Genetic Analyzer. During the run, ABI Collection and GeneScan software were used to gather the data, and following the run Genotyper software was used to generate DNA profiles. Today, you will be given a copy of the pooled class data for the paternity analysis and will learn how to interpret the results. After today's lab you should be able to:

- Exclude a man from the paternity of a child
- Qualitatively include a man as a possible father of a child
- Calculate a PRM value for a complete human STR profile
- Read a Genotyper print-out and understand how to interpret it

THINGS TO DO:

1. Your instructor will lead you through how to read the Genotyper printout. When you are clear about how to record the data, fill in the table below.

	Mother's Alleles	Child's Alleles	Obligate Paternal Allele	Alleged Father Mark's Alleles	Alleged Father Peter's Alleles	Can either be excluded?
PROFILER LOCI						
D3S1358						
VWA						
FGA						
D8S1179						
D21S11						
D18S51						
D5S818						
D13S317						
D7S820						
ADDITIONAL COFILER LOCI						
D16S539						
THO1						
TPOX						
CSF1PO						
Amelogenin						

- Circle the allele at each locus that the child received from the mother.
- Check to see whether either of the alleged fathers could have provided the child with the allele that is *not* circled. **This is the obligate paternal allele.** Record the obligate paternal allele in the appropriate column. Is it possible to eliminate one or both of the fathers? If so, indicate on the table where the exclusion(s) occur.
- Does either man *fail* to be excluded? Explain.
- Are your amelogenin results consistent with your expectations? Explain.
- We will now calculate a paternity index to determine the likelihood that the alleged father who was *not* excluded is the actual father of the child. The **paternity index** is a ratio of the likelihood that the alleged father contributed the obligate paternal allele versus a random man of the same ethnicity. In order the "call" (positive ID) a paternity, a **combined paternity index (CPI)** of 100 must be reached.

At each locus, the likelihood that the alleged father could have contributed the obligate paternal allele is either 0.5 (if he is a heterozygote) or 1.0 (if he is a homozygote). This makes sense, since the father has a $\frac{1}{2}$ chance of contributing the obligate paternal allele if he has two different alleles to give, whereas if he has only one, he must contribute that allele to all his offspring.

At each locus, the likelihood that a random man could have contributed the obligate paternal allele is simply the frequency of the allele in the population.

Paternity Index (at a given locus) = $\frac{\text{Prob. Alleged Father could have contributed obligate paternal allele}}{\text{Prob. Random Man could have contributed obligate paternal allele}}$

And ... Combined PI = Product of individual PIs at each locus tested

Thus, at each locus, the PI is either $0.5/x$ or $1.0/x$, where x is the frequency of the obligate paternal allele, and the CPI is simply the product of all the individual PI values.

Based on this information, you and your lab partner can now fill in this table, calculating the PI at each locus and the CPI for the paternity test as a whole.

LOCUS	Obligate Paternal Allele	Freq. Obligate Paternal Allele	PI
D3S1358			
VWA			
FGA			
D8S1179			
D21S11			
D18S51			
D5S818			
D13S317			
D7S820			
D16S539			
THO1			
TPOX			
CSF1PO			
Combined PI			

- Based on your results, could you "call" this paternity in a court of law? *Explain*. How much more likely is it that our alleged father is the biological father of the child than a random man from the same ethnic group?
- Calculate the PRM value for the **child's** full STR genotype. Check with your instructor if you have questions about how to perform this calculation.

PRM (child's genotype) = _____

- How many people would you have to test in the human population before you found someone with this same STR genotype?

Answer: _____

Can you see why STR analysis is also called DNA *identification* testing?

APPENDIX: PROFILER PLUS AND COFILER ALLELE FREQUENCY TABLES

PROFILER PLUS								
ALLELE	Afric. Am.	Caucasian	ALLELE	Afric. Am.	Caucasian	ALLELE	Afric. Am.	Caucasian
D3S1358			D8S1179			D18S51 Cont.		
9	0.0026*	*	8	*	0.0175	16	0.1744	0.14
10	*	*	9	0.0051*	0.01*	17	0.1641	0.105
11	0.0026*	0.0025*	10	0.0308	0.08	18	0.1103	0.06
12	0.0051*	0.0025*	11	0.0436	0.0625	19	0.1026	0.0375
13	*	0.0050*	12	0.1051	0.1425	20	0.0487	0.015
14	0.118	0.1125	13	0.1974	0.3475	21	0.0103*	0.0100*
15	0.2795	0.2825	14	0.3333	0.1875	22	0.0103*	0.0025*
15.2	0.0026*	*	15	0.2103	0.13	23	0.0026*	0.0025*
16	0.3231	0.2225	16	0.0641	0.02	24	*	*
17	0.218	0.2225	17	0.0103*	0.0025*	25	*	*
18	0.0462	0.145	18	*	*	26	*	*
19	0.0026*	0.0050*	19	*	*	D5S818		
vWA			D21S11			7	0.0026*	0.0025*
11	0.0026*	*	24.2	*	0.0050*	8	0.0513	0.0050*
12	*	*	25	*	*	9	0.0205	0.0225
13	0.0154	*	26	0.0026*	*	10	0.0744	0.0675
14	0.077	0.085	27	0.059	0.0375	11	0.2539	0.3925
15	0.2205	0.0825	28	0.218	0.1625	12	0.3256	0.3325
16	0.2692	0.1975	28.2	*	*	13	0.2487	0.165
17	0.1692	0.25	29	0.2	0.2075	14	0.0205	0.0100*
18	0.1385	0.2575	29.2	*	0.0050*	15	0.0026*	*
19	0.0846	0.11	29.3	0.0026*	*	16	*	0.0025*
20	0.0205	0.015	30	0.1615	0.2625	D13S317		
21	*	0.0025*	30.2	0.0256	0.025	5	*	0.0025*
22	0.0026*	*	31	0.0897	0.055	8	0.0359	0.115
FGA			31.2	0.059	0.105	9	0.0231	0.0775
16.2	0.0026*	*	32	0.0077*	0.0125*	10	0.0231	0.0675
17	0.0026*	*	32.2	0.0692	0.0725	11	0.2718	0.3125
18	0.0051*	0.015	33	0.0051*	0.0025*	12	0.4436	0.2825
19	0.0462	0.0625	33.1	0.0026*	*	13	0.141	0.0975
19.2	0.0026*	*	33.2	*	0.0075*	14	0.0615	0.0425
20	0.0436	0.1625	34	0.0026*	*	15	*	0.0025*
20.2	*	0.0075	34.2	*	0.0075*	D7S820		
21	0.1333	0.1775	35	0.0359	*	6	0.0051*	*
21.2	0.0026*	*	35.2	*	*	6.3	*	0.0025*
22	0.1897	0.14	36	0.0154	*	7	*	0.025
22.2	0.0026*	0.0050*	38	0.0026*	*	8	0.1795	0.175
23	0.1897	0.14	D18S51			9	0.118	0.13
24	0.1641	0.1325	9	*	*	10	0.3359	0.24
25	0.1231	0.1125	10	0.0051*	0.0050*	11	0.2282	0.23
26	0.041	0.015	10.2	0.0051*	*	12	0.0949	0.16
26.2	*	*	11	0.0154	0.02	13	0.0359	0.0275
27	0.041	0.0050*	12	0.0564	0.1425	14	0.0026*	0.0075*
28	0.0077*	*	13	0.041	0.15	15	*	0.0025*
29	0.0026*	*	13.2	0.0051*	*			
30	*	*	14	0.0667	0.1675			
			14.2	0.0051*	*			
			15	0.1769	0.1425			

COFILER PLUS (does not include alleles shared with Profiler)								
ALLELE	African-American	Caucasian	ALLELE	African-American	Caucasian			
D16S539			CSF1PO					
5	0.0026*	*	6	0.0026*	*			
8	0.0385	0.015	7	0.0718	*			
9	0.1821	0.1075	8	0.0744	0.0025*			
10	0.1128	0.0425	9	0.0385	0.0275			
11	0.03	0.2975	10	0.3051	0.2675			
12	0.1846	0.3425	10.3	*	0.0025*			
13	0.1564	0.175	11	0.2128	0.2575			
14	0.0205	0.0175	12	0.2205	0.3575			
15	0.0026*	0.0025*	13	0.0692	0.07			
			14	0.0051*	0.015			
			15	*	*			

THO1		
5	0.0026*	*
6	0.1308	0.2525
7	0.4	0.16
8	0.2282	0.0925
9	0.1308	0.1375
9.3	0.1	0.35
10	0.0077*	0.0075*

TPOX		
6	0.0821	*
7	0.0359	*
8	0.3436	0.5825
9	0.2	0.1125
10	0.0923	0.0475
11	0.2154	0.225
12	0.0308	0.03
13	*	0.0025*

* A minimum allele frequency of 0.0103 is suggested by the National Research Council for forensic or paternity calculations using these databases.