Polymerase Chain Reaction (PCR) and Forensic Investigation

(Adapted from Carolina Scientific student manual)

a) Introduction

You may be aware that crime scene investigators are able to identify individuals by a single fleck of dried blood, one hair, or even one drop of saliva left on a glass. But how is this done? Molecular biology techniques have revolutionized crime scene investigations. In this laboratory, you will perform a forensic analysis of DNA using the polymerase chain reaction (PCR) method of DNA amplification.

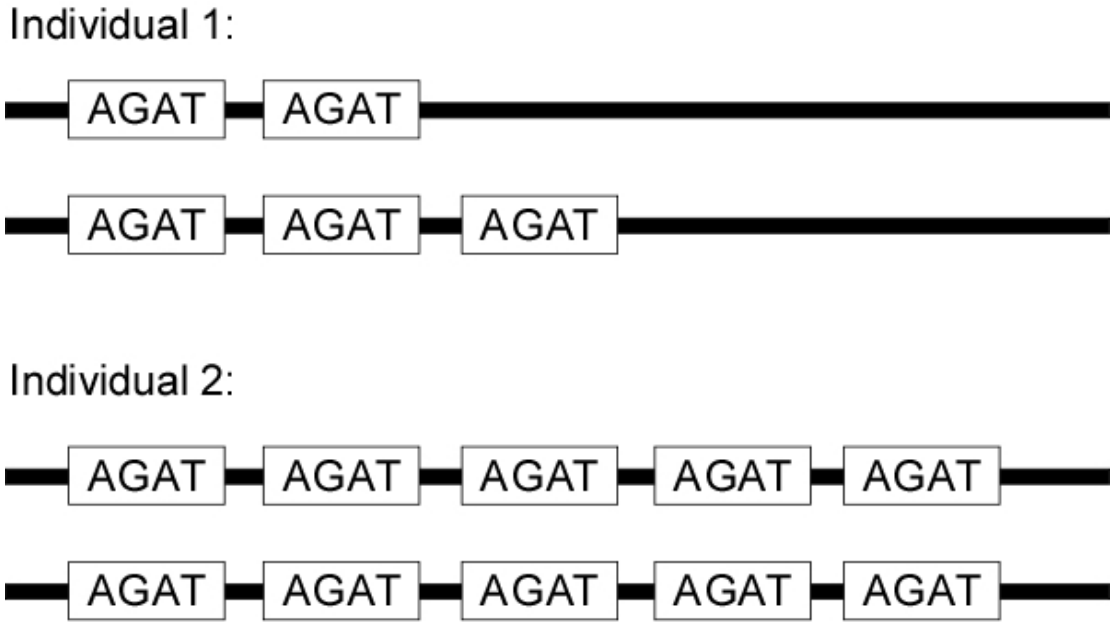
Tandem repeats

At many loci (locations) in the genome there are regions of DNA that contain repeated base-pair sequences, where a simple DNA sequence is repeated many times in a row. At any such locus, the number of tandem repeats of the simple sequence varies between individuals. Furthermore, an individual often has a certain number of tandem repeats at that locus on one member of their homologous pair of chromosomes, and a different number of tandem repeats at the same locus on the other chromosome of their homologous pair. In other words, individuals are often heterozygous for the number of tandem repeats they have at a given locus in their genome.

Because tandem repeat sequences are so variable in number, analysis of even a few of these sites can serve as a unique identifier of an individual (called a "DNA profile" or a "DNA fingerprint" of the individual). Because these sequences are inherited, they can be used to determine how closely two individuals are related (such as to determine who is the father of a baby). Another use for analysis of tandem repeat sequences is to match a crime suspect's DNA to the DNA in tissue samples found at a crime scene.

The diagram on the next page shows example tandem repeat sequences of two individuals. Each white box represent one repeat of the sequence AATG occurring at a certain locus on a chromosome (the long black line). Each person has two of each chromosome (they have a "homologous pair" of each

chromosome), one inherited from their mother and one inherited from their father. In this example, the repeated AGAT represents a tandem repeat sequence. Individual 1 has two repeats of the sequence on one chromosome and three repeats of the sequence on the other chromosome. Individual 2 has five repeats of the sequence on both chromosomes.



Tandem repeats occur at many loci in the genome. The loci that have an extremely large numbers of repeats in a row are called “satellites”. The loci that have very few repeats in a row are called “microsatellites” or “short tandem repeats” (STRs). Between these two extremes are loci that have a medium number of repeats in a row. These are called “minisatellites” or “variable number of tandem repeats” (VNTRs). Both STRs and VNTRs are used for DNA fingerprinting.

Calculating the number of possible combinations of VNTR alleles at a locus.

We can refer to each variation of the number of tandem repeats at a locus as one “allele” of that locus. As an example, consider a certain imaginary VNTR locus that we will call locus Q. Imagine that analysis of locus Q in many people has revealed that there are only six possible “alleles” at this locus: 10 repeats, 11 repeats, 12 repeats, 13 repeats, 14 repeats, or 15 repeats of the simple sequence. These are the six alleles of VNTR locus Q.

Each person has two parents, and therefore each person has two copies of locus Q. How many different possible two-allele combinations of the six alleles are there? In other words, how many different genotypes are there for locus Q? The general formula for this calculation is this:

$$\text{Number of possible genotypes (two-allele combinations)} = \frac{N \times (N + 1)}{2}$$

Where N = the number of different alleles at the locus.

For example, the Q locus discussed above has six alleles. If we substitute 6 for N in the above formula, then the formula tells us that there are 21 possible genotypes (two-allele combinations) at the Q locus.

The term "random match" means when a person happens to have the same genotype as a DNA sample but that person did not give the DNA sample. Roughly speaking, the probability that a person will be a random match for a DNA sample is the inverse of the number of genotypes. For example the chance of a person having a random match to a given genotype at VNTR locus Q is 1/21.

As another example, if another VNTR locus (let's call this locus R) has 10 different alleles, then the formula tells us that there are 55 possible genotypes at the R locus. The chance of a person having a random match to a given genotype at VNTR locus R is 1/55.

If we analyze more than one VNTR locus, then the possibility of a random match at all of the loci analyzed is the found by multiplying the probability of a random match at each individual locus. For example, the Q locus has a 1/21 probability of a random match, and the R locus has a 1/55 probability of a random match. The total possibility that a person would be a random match for given genotypes at the Q and R loci together is:

$$\frac{1}{21} \times \frac{1}{55} = \frac{1}{1155}$$

In the United States, the FBI uses 13 different VNTR loci to create the DNA fingerprint of an individual. These 13 VNTR loci are called the CODIS (combined DNA index system) 13. The probability that a person would happen to be a random match to the genotypes in all 13 VNTR is about 1 in a billion. Therefore, if a suspect's DNA matches the DNA found at a crime scene in all 13 VNTR loci, then it is almost certain that the suspect was present at the crime scene.

b) The PCR process

To perform DNA profiling, cells must first be collected from an individual. Cells from skin, hair, saliva, blood, or any other tissues or body fluids can serve. The DNA sequence of an individual is identical regardless of which cell type in the body from which the DNA was obtained. The amount of DNA isolated from a forensic crime scene sample is generally by itself not large enough to make a DNA profile. Through the process of polymerase chain reaction (PCR), however, the DNA in a small tissue sample can be amplified so that it can be used for DNA profiling.

PCR is a powerful technique that allows us to generate large quantities of a specific locus from a very small amount of starting DNA, such as the amount isolated from a single drop of blood or from a single human hair. The PCR method leads to exponential amplification of the selected locus in the DNA sample by doubling the number of copies of the locus after every replication cycle. For example, if we start with two DNA molecules (two copies of the selected locus) and we carry out 35 cycles of PCR, we could obtain 2^{35} (34,359, 738, 368) copies of the locus.

A PCR reaction requires template DNA (the sample DNA from which the locus will be amplified), DNA nucleotides (dNTPs), primers, Taq DNA Polymerase enzyme, and a buffer.

The primers are a pair of short pieces of DNA (usually about 18–22 bases in length) that are complimentary to the sequences at the end of the locus. One primer is complimentary to the sequence at the 5' end of the locus and one primer is complimentary to the sequence at the 3' end of the locus. The two primers are sometimes called the "forward" and "reverse" primers.

Taq DNA Polymerase is a heat-resistant enzyme that will make a complimentary DNA strand on a single-stranded template DNA strand. This enzyme was isolated from archaea that live in hot springs. Like all DNA polymerase enzymes, Taq DNA polymerase requires a primed area (a short region of double stranded DNA) to begin making the complimentary strand.

The PCR amplification reaction is done in repetitive cycles. The steps of each PCR cycle are the following (see diagram on the next page):

Step 1 = **Denaturation:** The DNA is heated to near boiling (about 95 degrees C). The heat separates the double stranded template DNA into two single stands.

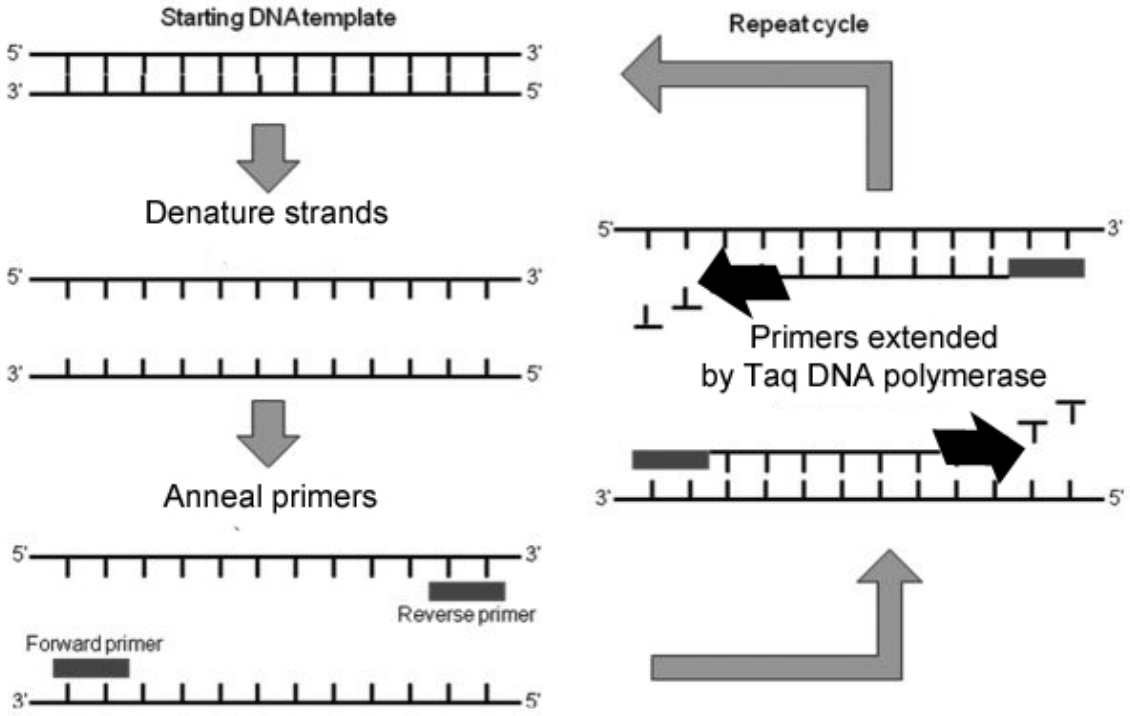
Step 2 = **Annealing:** After the temperature is lowered (usually to 55-72 degrees C) the primers bind to the ends of the locus on the DNA stands. One primer binds to the 5' end of the locus on one strand and the other primer binds to the 3' end of the locus on the other stand.

Step 3 = **Extension:** The Taq DNA Polymerase enzyme makes a complimentary DNA strand to the sequence that is flanked by each primer. Since the primers flank both strands of the locus that we wish to amplify, the net result is that the Taq enzyme makes copies of the locus.

After step 3, that cycle is complete. Next, the three steps are repeated (in other words, a new cycle begins). Each completed cycle doubles the number of copies of the locus from the number of copies in the previous cycle.

Note the temperature change between step 1 and step 2 of the PCR cycle. It would be extremely tedious to manually change the temperature of the PCR reaction by switching the test tubes between water baths of two different temperatures. Instead a laboratory device called a thermocycler carries out the temperature changes. The thermocycler can be programed for the exact temperature and duration needed for each step of the PCR cycle, and also the total number of cycles can be programed. In a typical PCR reaction, the thermocycler is programed to carry out roughly 30 cycles. The total time to complete the roughly 30 cycles is usually a few hours.

The amplified DNA from the PCR reaction is analyzed using agarose gel electrophoresis. Re-read the laboratory handout on that topic to review the major concepts of agarose gel electrophoresis and the proper use of micropipettes.



c) Overview of laboratory PCR exercise

Your lab group is a forensics crime scene investigation unit. You know that VNTRs can serve as markers to generate an individual's DNA fingerprint. You have been given the police report for a crime, along with five DNA samples from the crime scene.

On day 1 of this exercise, you will analyze the five DNA samples by PCR: The first three samples are from people involved in the crime: One sample is from the victim and two are from two suspects. So that no personal bias can affect your results, you will not know which samples are which.

The two additional samples from the crime scene are "evidence" samples that contain cells found at the scene: One evidence sample contains cells from saliva on a wine glass. The other evidence sample contains cells from under the victim's finger nails. These cells came from the attacker.

In additional to the above five samples related to the crime, you will prepare two addition test tubes that will serve as controls for your analysis: The positive control will be genomic DNA from a co-worker in your CSI station. The co-worker's alleles at this VNTR locus have been amplified previously (the co-worker has a 300 base pair allele and a 150 base pair allele). The alleles that you find in your positive control should match those the alleles. As a negative control, you will add water instead of DNA to one reaction tube. If any DNA fragments appear in this negative control, you probably have DNA contamination in your materials. Even a single flake of skin or drop of saliva that accidently falls into your test tubes can be amplified by PCR.

After your CSI team has prepared the 7 test tubes for PCR, you will insert them into the thermocycler. The laboratory technician has programed the thermocycler for 30 PCR cycles. She will start the PCR process when all test tubes have been loaded and she will store the test tubes until day 2 of the laboratory exercise.

On day 2 of the exercise, your CSI team will analyze the PCR results by agarose gel electrophoresis. Based on your findings, you will try to identify which suspect committed the crime. Good luck, investigators.

d) Police report from the crime scene

Report number: 3.14159256
Crime code: 2.71828

Victim: Carrie Meowt. Age: 22, Height: 5'5", Weight: 127 pounds.

Suspects: 1) Lida Lott. Age: 26, Height: 5'5", Weight: 130 pounds.
2) Anita Alibye. Age: 38, Height: 5'9", Weight:190 pounds.

Address where incident occurred: 123 Alda Way, San Dibeeches, CA

Ms. Meowt was found unconscious in her apartment on February 30th. Overturned tables and a spilled wine glass, along with cuts and bruises indicate a struggle. Gladys Friday (neighbor) heard a dog barking inside the victim's apartment. Bill Munthly (landlord) entered the apartment at 1900 hours and found the victim unconscious.

Victim information: Ms. Meowt left work (the Law office of Dewey, Cheetum, and Howe) early for an appointment and never returned. Status: Currently Ms. Meowt is in a coma under the medical care of Dr. Daryl B. Paine at county hospital.

Doctor Paine's medical report: Bruises on victim's hands and arms. Small laceration above left eyebrow. The victim's blood showed a blood alcohol content of 0.10 and also the presence of sodium thiopental. This drug is used in small doses as a truth serum. In larger doses it may be used as an anesthetic. Overdoses can be fatal.

Suspect 1 information: Lida Lott is the half-sister to the victim (shared father, Smokey Lott, is now deceased). Lida is a technician at Fleas Navidad Veterinary Clinic. Lida stated that she has not spoken to her sister since their father's funeral in January. She claims that she was at home all day on February 30th.

Suspect 2 information: Anita Alibye was a former co-worker of the victim at her law firm. The two were known to have quarreled often at work. Ms. Alibye claims to not have seen or spoken to the victim since she was fired from the law firm a month ago. She claims to have been hiking alone at the time of the attack.

Evidence 1: Swab taken from rim of spilled wine glass.

Evidence 2: Tissue taken from under the victim's fingernails.

e) Day 1 procedure: PCR of the samples

1) Obtain 7 PCR test tubes. Each test tube contains a colored bead that contains a prepared mixture of Taq DNA Polymerase enzyme, buffer, and dNTPs (DNA nucleotide monomers).

2) Use a waterproof marker to label the seven PCR test tubes with numbers 1 – 7. Be sure to also mark your group's tubes with a unique group name or symbol. It is best to write the marks on the sides of the tubes because the lids may lose their ink easily.

3) Carefully open the test tubes, making sure the bead stays inside.
As you pipette in today's lab, remember that you are testing DNA samples that can easily be contaminated with your own DNA from skin, hair, or saliva. Follow these guidelines to avoid contaminating your samples and for proper pipetting.

a) Never touch the tops or the insides of any test tubes, and try not to breathe into the tubes.
b) Use a fresh pipette tip for each pipetting of each sample. Be careful that the tip does not contact the desk top, your fingers, or anything other than the solution you are pipetting with that tip.
c) Remember the first stop/second stop rule when pipetting.
d) Deliver all samples to the bottom of the tube.

4) Add the primer mix to the test tubes according to the table below. The colored bead in the tube contains the Taq DNA Polymerase enzyme, the buffer, and the dNTPs (DNA nucleotide monomers).

Test tube:	PCR bead*	Primer mix:	DNA sample:
1	1	20 ul	5 ul of sample 1 (CW, victim)
2	1	20 ul	5 ul of sample 2 (LL, suspect)
3	1	20 ul	5 ul of sample 3 (AA, suspect)
4	1	20 ul	5 ul of sample 4 (wine glass)
5	1	20 ul	5 ul of sample 5 (fingernails)
6	1	20 ul	5 ul of sample 6 (+ control)
7	1	20 ul	5 ul of sterile water (- control)

* Each PCR test tube already has a PCR bead in it. The PCR beads contain Taq DNA Polymerase enzyme, buffer for the enzyme, and dNTPs.

5) After you have loaded the primers into each test tube, come to the front desk to load in the DNA samples according to the above table. The sterile water added to tube 7 is the negative control (no DNA). Once you have added the sample, mix by gently pipetting up and down several times. Be sure the all liquids are together at the bottom of the PCR tube.
6) Make sure that each tube is securely sealed. Place your tubes into the thermocycler. The laboratory technician will run the thermocycler for 30 cycles and then store your samples until the next lab meeting.

f) Day 2 procedure: Analyzing the PCR using Agarose Gel Electrophoresis

1) The laboratory technician has prepared the agarose gels for you. Remember that this gel contains the toxic compound Nancy520. Gloves must be worn when handling it. Also obtain a p20 micropipette and tips.

2) Carefully remove the comb and the rubber dams from the gel and place the gel (still in its casting tray) in the gel box. Add approximately 250 ml of the 1X TAE buffer to the box. Pour the solution into one end of the box, allowing it to wash over the gel and into the other end. The gel should be completely, but barely, submerged in the buffer. Be sure that the wells in the gel are filled with buffer.

3) Obtain your seven PCR test tubes and a test tube containing 8 ul of the DNA size ladder. Loading dye has already been added to these tubes.

The DNA size ladder contains several DNA pieces of different sizes that differ from each other by increments of 100 bp (base pairs). The smallest DNA piece of the ladder is 100 bp, the second smallest piece is 200 bp, the third smallest piece is 300 bp, etc.

4) Gently tap the tubes to collect all liquids inside to the bottom of the tube.

5) Move your gel box (with your gel in it) next to a power supply. Load 8 ul of the DNA size ladder into the first well. Then 15 ul of samples 1 – 7 (in that order) into the other wells of your gel.

6) Set the power supply to run at a constant 60 volts. Refer to the *Electrophoresis* lab handout for directions in setting the power supply for a constant voltage. Stop the gel when the red loading dye band has moved about 3 cm (1.5 inch) from the well. This should be about 2/3 of the way to the bottom of the gel.

7) When the electrophoresis is done, turn the power supply settings to zero and turn off the power, and unplug it. Using gloves to protect yourself from the ethidium bromide in the gel, carefully remove the gel from the box (it might slide easily off the tray). Place the gel in the ultraviolet viewing box (which has a built-in digital camera). Your instructor will assist you in photographing your gel with a ruler at its side. In the results section, record your results and answer the questions.

8) Clean up: All materials contain Nancy520, a toxic material. Therefore, nothing goes in the trash or down the sink. Place your gel in a special disposal tray, place the clear plastic gel tray, the comb, and the dams in another special tray. Place the gel box (with the TAE buffer still in it) back on the cart where you obtained it. The laboratory technician will clean these materials. This is the end of the day two activities.

g) Results section:

a) Determine the size (in bp) of each piece of DNA on the gel by comparing the piece to the DNA size ladder. In the table below, list the genotypes of the victim and the two suspects by writing the size of their VNTR locus allele(s). For example, an individual's genotype for the locus might be the 200 base pair allele and the 500 base pair allele. Next, state whether each person was heterozygous or homozygous for their two copies of the VNTR locus that was amplified.

	Genotype:	Heterozygous or homozygous?
Carrie Meowt:		
Lida Lott:		
Anita Alibye:		

b) Based on your findings, which crime scene samples match which people?

For results questions c - g, assume that there are 7 alleles of the VNTR locus that you amplified.

c) How many genotypes are there for this VNTR locus?

d) What portion of the population has this same genotype? _____
Do your findings prove who left body tissue at the crime scene? Explain:

e) Do your findings prove who did not leave body tissue at the crime scene? Explain:

f) Do you think there is enough evidence to reach a conclusion about who committed the crime? Explain.

g) In terms of VNTR loci, what might be done to improve the certainty that the crime scene tissue samples were left by the one of the suspects?

h) Review questions

1) Describe the activity of Taq DNA Polymerase enzyme at a molecular level. Describe exactly what the enzyme does in terms of the molecule it works on, its product molecule, and its substrate molecules. Be as specific as possible.

2) What special property does Taq DNA polymerase have that is not a property of other DNA polymerase enzymes (such as human or bacterial DNA polymerase enzymes)?

3) What sort of species was Taq DNA polymerase isolated from?

4) Why is it necessary to have a primer on each side of the DNA segment to be amplified? What would happen if you only used one primer?

5) Describe the 3 main steps of each cycle of PCR amplification and what reactions occur at each temperature.

6) If a PCR sample originally contained six copies of the suspect's genome (for example, there were six cells in the tissue sample that was PCRed), how many copies of the VNTR locus are there…

a) In the original sample: _____
b) In the sample after 1 cycle of PCR amplification: _____
c) In the sample after 2 cycles of PCR amplification: _____
d) In the sample after 3 cycles of PCR amplification: _____

7) If there are 11 alleles at a certain VNTR locus in the human population, what is the probability that a person would be a random match at that locus?

8) The human population has the following possible alleles at a certain VNTR locus:

a) 7 repeats of CGGCTGATT
b) 8 repeats of CGGCTGATT
c) 9 repeats of CGGCTGATT
d) 10 repeats of CGGCTGATT
e) 11 repeats of CGGCTGATT

How many possible genotypes are there at this locus?

9) VNTR locus R has 10 possible genotypes. VNTR locus S has 3 possible genotypes. What is the probability that a person would be a random match at both loci?

10) VNTR locus C has 10 possible genotypes. VNTR locus D also has 10 possible genotypes. VNTR locus E has 5 possible genotypes. What is the probability that a person would be a random match at all three loci?

11) In today's exercise, you analyzed just one VNTR locus to match the suspects to the crime scene tissue samples. How many VNTR loci would the FBI analyze to match suspects to crime scene tissue samples?

12) If the VNTR locus in today's exercise has 7 possible alleles, and if these 7 alleles are present in equal amounts in the human population, what is the chance that an innocent suspect would be a random match the tissue sample from the crime scene? What would your answer be if the VNTR locus had 23 possible alleles?

13) Each sample that you PCRed should have shown either just one band (if the tissue sample was homozygous for the VNTR alleles) or two bands (if the tissue sample was heterozygous for the VNTR alleles). If all of your samples (including your control tubes) showed two extra bands and those two bands extra were exactly the same in all PCR samples, what is the probable explanation for these extra bands?