PCR: a background and some Uses.

Introduction

A. DNA and Genetic Identity

Every person who ever existed is a unique individual. We can easily recognize individual differences in traits such as facial appearance, size, skin and hair color, personality, and performance. More subtle characteristics such as fingerprints and patterns of vascularization in the retina of the eye can be used to precisely distinguish individuals from one another and almost all of these physical differences are genetically determined. Put simply, people are physically different from one another because they contain different genes or different combinations of genes. Physical traits are encoded by the sequence of DNA base pairs contained in a person's chromosomes. The four nucleotide bases are represented by the letters A, C, G, and T, which stand for their chemical names, adenosine, cytosine, guanine, and thymine. The A in one strand always pairs with the T in the opposite strand, whereas C always pairs with G. The strands are said to be complementary to each other, that is if you know the sequence of one strand you can determine the exact sequence of the other strand. The sequence of base pairs that make up a gene encodes for proteins. These are the molecules that determine the biochemical properties and physical form of an organism. Most physical traits are determined by multiple genes which act together during embryonic development to give form to an individual. Each gene exists in two copies in every person, since each of our 23 chromosomes is paired. The physical location of a particular gene on a chromosome is called a genetic locus (the plural is loci). In a population, many different forms of a single gene can exist at any locus. Each unique form of an individual gene is known as an allele. When one considers that thousands of gene pairs act together during development and that multiple alleles can exist for nearly every gene, the amount of diversity we see within the human species is easy to understand.

There are approximately 3 billion DNA base pairs in the human genome. (The base pair count for a species represents the sum of genetic information for that species.) About 99.9% of this DNA sequence is identical in all humans. The 0.1% difference is responsible for our genetic identity as individuals. The differences between one person's DNA and another person's DNA are called polymorphisms (Greek: poly = many, morph = form). There are 3 different types of polymorphisms: In one form of polymorphism, the sequence of the DNA base pairs are rearranged in a way that changes the meaning of the DNA code. For example, if you change the order of the letters in the word CAT to ACT, the meaning of the word is different. A second way to change DNA is to insert additional base pairs within a coding sequence. This is like adding letters to a word and changing the meaning, as in adding 2 additional letters to the word ACT can change the meaning to ACTOR. The third way that polymorphisms occur in DNA is by subtraction (or deletion) of nucleotides. By deletion ACT could become AT, a different word with a different meaning.

B. DNA Profiling

DNA is a rather stable molecule in chemical terms. DNA released from dead or dying cells or tissues can be digested by microorganisms, but if the sample is dried, the DNA will remain intact for many years (or even thousands of years in mummified remains). Thus, material containing DNA from victims and perpetrators of a crime is often present at the scene. Since the sequence

of base pairs in a DNA sample collected at a crime scene can contain the genetic identity of the individuals present, forensic scientists have borrowed genetic techniques to develop "DNA fingerprints" of suspects.

While it is technically possible to determine the base sequence of DNA fragments collected at a crime scene, this would be difficult, expensive, and time consuming. Therefore, rapid methods that type or characterize DNA from different individuals have been developed.

The polymerase chain reaction was developed by Kary P. Mullis, who was awarded a Nobel Prize for the discovery. The PCR reaction uses a DNA polymerase enzyme that is isolated from the bacteria *Thermus aquaticus* (Taq) that lives in hot springs in Yellowstone National Park. This organism normally lives at temperatures exceeding 70° C. In nature Taq DNA polymerase replicates the bacterial chromosome during cell division; however, in a test tube reaction, Taq polymerase can be used as a thermo-stable reagent to allow multiple rounds of replication of relatively short, well-defined DNA segments from any organism. In the PCR reaction, a segment of a complex genome is "targeted" for amplification by Taq polymerase by defining the region with 2 short synthetic DNA molecules complementary to sequences flanking the region of interest. These short DNA molecules are known as primers. PCR primers are made on a special machine called an oligonucleotide synthesizer and they can contain any desired sequence. Therefore, any gene or DNA region where the sequence is already known can be amplified by PCR. The regions of DNA we will be amplifying are called microsatellites. Microsatellites are non-coding regions of the DNA (they do not make proteins and have no function as far as we know) consisting of long repeats of specific base pairs. For example the three microsatellites we will be amplifying consist of ATA repeats, that is they are series of ATA ATA ATA repeated various numbers of times. Yes, in this case all three of the microsatellites have the same repeat unit. However, this no the case with all microsatellites, which, can have repeats using any combination of the four base pairs. Because microsatellites are non-coding there is a high mutation rate because it doesn't matter if you have 20 repeats or 50 repeats. Given this, every individual will have a unique combination of microsatellites and a unique DNA fingerprint using these markers. These markers are also used to help diagnose genetic diseases (none of our primers are associated with any know disease) and determine parentage. Each of you has two copies of each chromosome, one from each parent and thus two copies of each microsatellite. When both of your copies of the microsatellite are the same size (have the same number of repeat units) we will see only one band on the gel and you are called homozygous for that locus. If the alleles you got from your parents are different sizes we will see two bands on the gel and you are called heterozygous for that locus. We are working with three human microsatellite loci, one from chromosome 11, one from chromosome 12, and one on the Y.

A PCR reaction contains the target DNA and all the necessary components for DNA duplication: large quantities of all 4 deoxynucleotide bases, large quantities of the two primer DNA fragments, salts, and Taq polymerase. The polymerase chain reaction is carried out in 3 steps:

1) The 2 strands of the target DNA double helix are separated by heating to 94° C for a few seconds.

2) The reaction is then cooled to about 58° C. This allows the 2 primers to bind or "anneal" to their complementary strands in the target DNA. This takes 20 - 30 sec.

3) The reaction is then heated up to 72° C, the temperature at which Taq polymerase works best. The Taq polymerase begins adding deoxynucleotide bases to the ends of each primer and eventually makes a complementary copy of the template DNA. This completes one PCR cycle. A diagram of the PCR cycle is shown in Figure 1.

Theoretically, PCR can amplify a single molecule of target DNA. Thus, at the end of the first cycle of PCR, we would now have 2 molecules of the targeted DNA segment. The 3 steps in the PCR process - strand separation, annealing of the primers to the template DNA, and synthesis of the new DNA strands - takes less than 2 minutes. But the cycle can be repeated at least 30 more times before the Taq polymerase loses activity from the repeated heating. Each newly synthesized DNA strand can act as a new template, so after 30 cycles, 1 billion copies of the original DNA segment can be produced! Taking into account the time it takes to change temperature of the reaction vial, 1 billion copies can be ready in about 3 hours.

The speed and extreme sensitivity of the PCR reaction has revolutionized the use of DNA evidence in criminal investigations. A few million copies of an amplified DNA fragment is sufficient material to be detected by staining with fluorescent dyes following separation of the reaction products by electrophoresis, or via laser excitation of fluorescent markers attached to the fragments. Thus, it is possible for forensic scientists to analyze the genetic patterns found in only a few molecules of DNA collected at a crime scene. A large number of DNA primers flanking different regions of the human genome that are known to be polymorphic are now available. Most of these polymorphisms involve deletions or insertions; thus it is relatively easy to detect the various forms of the genes that exist in different individuals. As with Southern blotting, if enough different genetic loci are examined by PCR, the probability that any two people have exactly the same PCR products becomes vanishingly small. Genetic "fingerprints" yielded by PCR can be just as effective as traditional fingerprints in distinguishing identity. The simplicity of the reaction now makes PCR the method of choice for many applications that require DNA typing. These applications include criminal investigations, paternity cases, screening for genetic disorders, and the identification of human remains following disasters. To facilitate identification of military personnel, samples of DNA from all U.S. servicemen are now kept on file in case DNA fingerprint analysis is ever required in the future.

Normally, prior to performing a PCR reaction we would quantify and amount and purity of the DNA you extracted. However, since we are doing the scat samples the majority of the DNA you would be measuring will be bacterial. We will quantify the other DNA you extracted (Badger and your hair samples) in lab. The following protocol is for amplifying up the scat DNA only. Tissue DNA will follow a slightly different protocol.

Analysis comparing the ratio of absorbance readings at A260 and A280 (260 nm and 280 nm) was first described by Warburg and Christian (Warburg, O. and W. Christian. 1942. *Isolation and Crystallisation of Enolase*. Biochem. Z. 310:384-421) to assess protein purity in the presence of nucleic acid contaminates. Today, this method is commonly used to determine both nucleic acid purity and yield. Nucleic acids absorb UV light with maxima for the four nucleotide

components centred primarily at 260 nm. Using a 1-cm light path, the extinction coefficient for nucleotides at this wavelength is 20. Based on this extinction coefficient, the absorbance at 260 nm in a 1-cm quartz cuvette of a 50µg/ml solution of double stranded DNA or a 40µg/ml solution of single stranded RNA is equal to 1. You can calculate the concentration of the DNA or RNA in your sample as follows:

DNA concentration (µg/ml) = (OD 260) x (dilution factor) x (50 µg DNA/ml)/(1 OD260 unit)

A trough is usually observed at 228 nm. Proteins, a frequent contaminant of both DNA and RNA samples, absorb UV light at both 280 nm and 228 nm. Polypeptide chains commonly contain up to three amino acids with aromatic moieties that absorb significantly at 280 nm. Some proteins, such as histones or protamines, contain few or no aromatic residues and have little or no absorbance at 280 nm.

These variations in amino acid composition can have a tremendous impact on a protein's absorbance at 280 nm. Peptide bonds absorb at 228 nm and are a more constant indicator of the presence of protein in a sample. Thus, absorbance readings measured both at 228 nm and at 280 nm provide a more accurate estimate of proteins or peptides that may be present in nucleic acid samples. Carbohydrates, sulphydryls, phenolics and other aromatic compounds can also absorb at these wavelengths.

By comparing the ratio of the A260 reading to both the A280 and A228 measurements, the presence of contaminants in a nucleic acid sample can be better evaluated than by determining just the A260/A280 ratio alone. This type of UV absorbance analysis requires a sensitive instrument capable of detection far into the UV portion of the electromagnetic spectrum.
The ratio of the absorbance at 260 nm/ absorbance at 280 nm is a measure of the purity of a DNA sample; it should be between 1.65 and 1.85, with 1.9 being pure DNA. The ratio of the 260/230 should be >2.0 since nucleic acids have an abosrobance minima at 230nm.

Creating the PCR mix

We are going to be running standard PCR reactions of 20µl (0.00002 L, or 0.000005 gallons). The PCR mix will contain the following amounts for each reaction you will perform. As you add the specific reagent put a check mark in the column under its heading and fill in the column with the appropriate values for doing 4.4reactions. Why 4.4reactions when you will only be running 4 reactions? The extra 0.1 reactions is to allow for pipetting errors.

Reagent	Volume 1 rxn	Total for 4.4 rxn	MM d-Loop
Ultra Pure Water	11.55 µl		
10x PCR Buffer with $MgCl_2$	2 µl		
5x TaqMaster	1µl		
BSA (2mg/ml)	2µl		
dNTP's	1µl		
Forward Primer	0.6 µl		
Reverse Primer	0.6µl		
Taq polymerase	0.25µl		
Template	1 µl	------	
Total	20µl		

Table 1
Amount of reagents need for each primer. Place an X in the cell as the reagent is added.

All reagents must be completely thawed prior to use and thoroughly vortexed.

1. Label a 1.5 ml tubes as the master mixes for each primer (MM d-Loop) and place it in a tube rack in an ice bucket.

ALL REAGENTS AND DNA MUST BE KEPT ON ICE AT ALL TIMES!!!!

2. Label a set of 0.2 ml PCR tubes with your intitials.
3. Add the required amount of each reagent from your calculations in table 1.

MAKE SURE THAT YOU ADD THE REAGENTS IN THE ORDER LISTED ABOVE!!!!
Note: we will not add the Taq until just before we are ready to dispense the master mix

4. Vortex this mixture and place 19 µl of the appropriate master mix in the appropriate 0.2µl tubes.
5. Add 1µl of each individual's DNA to the appropriate tube.
6. Keep your samples on ice until everyone is done, then place them in PCR machine.