miniPCR™ Forensics Lab:
Analysis of the D1S80 VNTR

Student's Guide Contents

1. Scenario overview

The crime:

Yesterday, after class, a half-consumed bottle of water was found at one of the lab benches in your classroom. As eating and drinking in the lab are strictly prohibited, your teacher stepped into crime solving mode. Your teacher swabbed the mouth of the bottle to collect a DNA sample. They then purified the DNA and used PCR to amplify the D1S80 region of chromosome 1.

To eliminate yourself as a suspect, you must now amplify your own D1S80 region. If your D1S80 alleles do not match the criminal's, you may be eliminated as a suspect, but if your DNA matches the criminal you must be considered as a possible perpetrator.

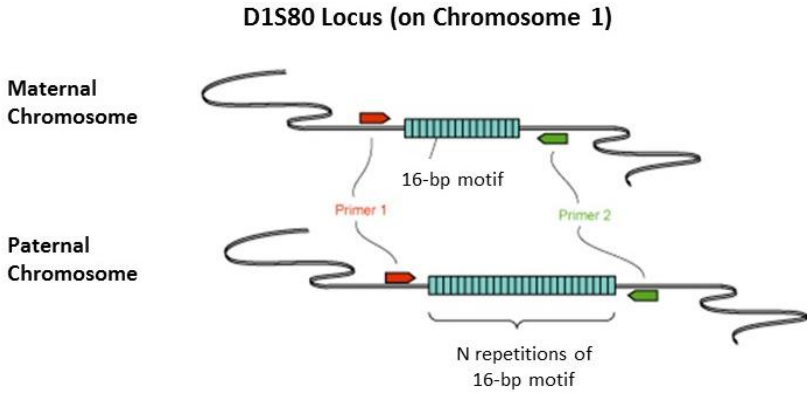
Background:

D1S80, a Variable Number Tandem Repeat Region

Variable number tandem repeats (VNTRs) are regions in a genome that contain short stretches of DNA (6 to about 100 base pairs) repeated a number of times. The number of repeats in a particular VNTR can vary from individual to individual, and between chromosomes within an individual. In this way, VNTRs are one source of genetic polymorphism (variation), and can be used as markers for personal identification as well as in the study of inheritance, genetic diversity, population genetics, and genetic disorders.

VNTRs are predominantly found in non-coding regions of DNA, and therefore their length does not usually bear any functional importance (though there are reported exceptions). The highly repetitive nature of VNTRs can lead to a relatively high frequency of errors in DNA replication or crossing over. When these mutations occur, they add or remove repeats thus changing the length of the VNTR. The result is that there are many different versions (alleles) of any one VNTR in the population, each with a different number of repeats. This variation means that we usually inherit different alleles from each parent, making most people heterozygous for a given VNTR. This variation also means that any two randomly selected people will rarely have matching alleles for their two copies of a VNTR. When we analyze two people for more than one VNTR, the likelihood of a random match decreases multiplicatively.

D1S80 is a VNTR region located on human chromosome 1 and consists of a 16-base pair-long repeat unit. Most people have between 14 and 41 copies of this repeat, resulting in D1S80 alleles with repeat regions of 224 to 656 base pairs in length. The most common alleles are comprised of 18 and 24 repeat units (288 and 384 base pairs).



PCR amplification and gel electrophoresis can be used to establish the length of a person's D1S80 alleles. To amplify the D1S80 allele, primers are used that bind just outside the region of repeats. PCR will then copy the entire set of repeats. When the PCR products are run on an electrophoresis gel they will differ in size depending on how many repeats are present in the D1S80 allele. The actual length of product analyzed on the gel will include the repeats plus the additional length of the PCR primers and any DNA region flanking the repeats. The additional length of the PCR product in this lab is about 180 base pairs.

Mutation vs Genetic Polymorphism:

In this lab, we use *mutation* to refer to a change in a DNA sequence. Mutations can occur in individual cells due to copying errors in replication, errors in crossing over during meiosis, or even rearrangements of chromosomes. The changes in a DNA sequence that are introduced by a mutation can vary significantly in their effect on an organism. Often, they are deleterious and will quickly be removed from a population by natural selection. Other changes to the DNA are neutral and have no effect on the organism; they may remain in a population and spread through stochastic processes like genetic drift. Occasionally, a change in DNA sequence will be beneficial; these mutations will tend to spread in a population propelled by natural selection.

We refer to *genetic polymorphisms* when there is more than one sequence variant, or allele, present in a population for a particular gene or locus. A new variant first arises through mutation, and if that new sequence spreads through the population it is called a genetic polymorphism. Genetic polymorphisms exist on the population level. An individual that has more than one version of a gene or DNA sequence is called a heterozygote. Heterozygotes can only exist in a population that has a genetic polymorphism.

Use of DNA as an identifier in forensics

Every person has a unique DNA sequence; it is what makes one person different from every other person. For this reason, DNA is in many ways the best, most accurate way of identifying an individual. The problem lies in that the vast majority of DNA between any two people is identical. Approximately 99.6-99.9% of DNA between any two people is the same.

The most comprehensive way to identify the genetic differences between individuals would be to sequence large regions of an individual's genome – identifying every possible difference in their DNA. This is not practical for many reasons, most notably cost and time required. Therefore, when using DNA to identify individuals, it is important to first identify regions of DNA that hold high levels of variability between individuals. VNTRs are just such regions.

Because the variation in VNTRs affects the length of the DNA, scientists can use techniques that detect DNA length differences to identify variability between individuals relatively easily and inexpensively. Common VNTR analysis techniques are PCR and DNA electrophoresis (gel or capillary), or other methods.

Normally when trying to make a genetic match of individuals, several variable regions are compared simultaneously. This is because when looking at any one VNTR locus, there is a reasonable likelihood that any two people's alleles may match by chance. While finding two people who match in their alleles at any one VNTR locus may be somewhat likely, the combined probability of matching many different regions simultaneously is extremely unlikely, even for related individuals. When each region investigated is on a different chromosome and therefore assorts independently in meiosis (a founding principle of Mendelian genetics), the probabilities of matching at each locus are also independent. This is where the extremely low probabilities of making a false match come from. In modern crime databases the theoretical probability of getting a false match, matching many loci just by chance, is reported to be as low as one in a quintillion.

In forensics today, the use of VNTRs has generally been replaced by identification through the use of STRs (short tandem repeats). They consist of repetitive units of DNA that are 2-13 base pairs long, instead of 6-100 base pairs like VNTRs. Otherwise, the use of VNTR and STR analysis is basically the same. Both look at length variation of a highly polymorphic region and, by looking at many of these regions simultaneously, increase the likelihood that any particular DNA profile will be unique. The FBI and other law enforcement agencies standardized the use of DNA analysis in personal identification by choosing 20 STR loci that are used in all forensics investigations. Now, because all DNA analysis is done in the same way on the same variable regions, the FBI, different states, and different law enforcement agencies across the country can all compare DNA in the same uniform way. The database used to house and organize all these data is called the Combined DNA Index System, or CODIS. When you hear on the news or on a crime TV show that a DNA sequence is being compared to a criminal database to look for a match, this set of 20 STRs is what they are talking about.

It is important to remember that, while DNA evidence has become the gold standard for forensics identification, it is not perfect. While the normally reported chance of a false match in the CODIS system is less than one in a quintillion, that number only refers to two randomly selected individuals. Two related individuals or individuals from a genetically distinct ethnic group will be more likely to share alleles (though an exact match is still highly unlikely). Much more problematic is attempting to match DNA to a complex sample. While comparing DNA evidence collected from one individual to evidence that also contains a single individual (as you will do in this lab) is fairly straightforward, often DNA samples are mixed from several individuals. Imagine, for example, if in today's lab 2 people had drunk

from the same water bottle. Every locus investigated will now be represented by alleles from two people. It may be impossible to separate the DNA from the two individuals, making analysis more difficult. These problems are not insurmountable, but they can lessen the overall reliability of DNA evidence and require highly trained technicians to interpret the data. Lastly, contamination of samples will often create the biggest problems. Because of the pervasiveness of DNA in the environment and the power of PCR amplification, very small sources of contamination can lead to false positives. To combat this problem, technicians often run multiple negative controls in their analyses. In the end, as is often the case, the evidence is only as good as the technician doing the analysis.

Today's lab

While STRs are currently used in most forensics analyses, they present some problems for classroom investigation. Because STRs consist of such short, repeated segments, they are usually analyzed by capillary electrophoresis, a procedure that uses highly specialized equipment. Capillary electrophoresis operates on the same principle as traditional gel electrophoresis, but can resolve size differences of just a few base pairs. A major downside, however, can be the cost and complexity of the equipment. This tends to relegate the use of such equipment to highly specialized laboratories.

Today we will use the D1S80 locus, a VNTR found on chromosome 1. Because the size differences regularly found in the D1S80 locus are on the order of dozens or hundreds of base pairs, regular classroom gel electrophoresis equipment can be used to analyze your data. Students should note that the VNTR analysis done in this lab is based on the exact same principles as STR analysis, and can be used for DNA identification in the same way. The main divergences from actual forensic analysis are that your size differences may be larger than some STR differences, and that we will only analyze one locus instead of the 20 usually used for criminal forensics data.

Read how miniPCR thermocyclers are used by real forensic scientists in the field using this miniPCR Case Study.
https://www.minipcr.com/case-studies/forensic-dna-analysis-evidence/

Learn more about DNA fingerprinting using DNAdots – a free resource from miniPCR.
https://dnadots.minipcr.com/dnadots/dna-fingerprinting

2. Laboratory guide

A. Cheek cell collection and DNA Extraction

1. **Label one 200 µL thin-walled PCR tube per student** on the side wall
   - Label with your initials
2. **Add 50 µL of X-Tract DNA Extraction Buffer** to your tube
3. **Scrape the inside of your cheek** multiple times with a flat-end toothpick to saturate the end of the toothpick with cells
   - Rub *gently* along cheek, taking care not to perforate skin. It shouldn't hurt!
4. **Dip the toothpick in the tube** swirling it in X-Tract™ DNA Extraction Buffer
   - Swirl toothpick thoroughly in the buffer to release cells
5. **Tightly cap the tubes**
6. **Incubate** the tubes for 10 minutes at 95˚C
   - Use miniPCR™ machine in Heat Block mode, a heat block, or water bath
   - Ensure the tubes remain steady and in vertical position, allowing cell debris to decant to the bottom of the tube by gravity
7. **Remove tubes from heat block** and let them rest in a tube rack
   - DNA extract must be used immediately for PCR

DNA extracted by this method is not stable for long-term storage

B. PCR set up

1. **Label 1 clean PCR tube** (200 µL thin-walled tubes) per student (4 per lab group) <u>on the side wall</u>
    - With student initials

2. **Add PCR reagents** to each labeled PCR tube

	Per Tube
D1S80 Primer Mix	12 µL
2X Blue PCR Master Mix	12 µL
DNA extract sample	2 µL
Final volume	***26 µL***

Remember to change tips at each step!

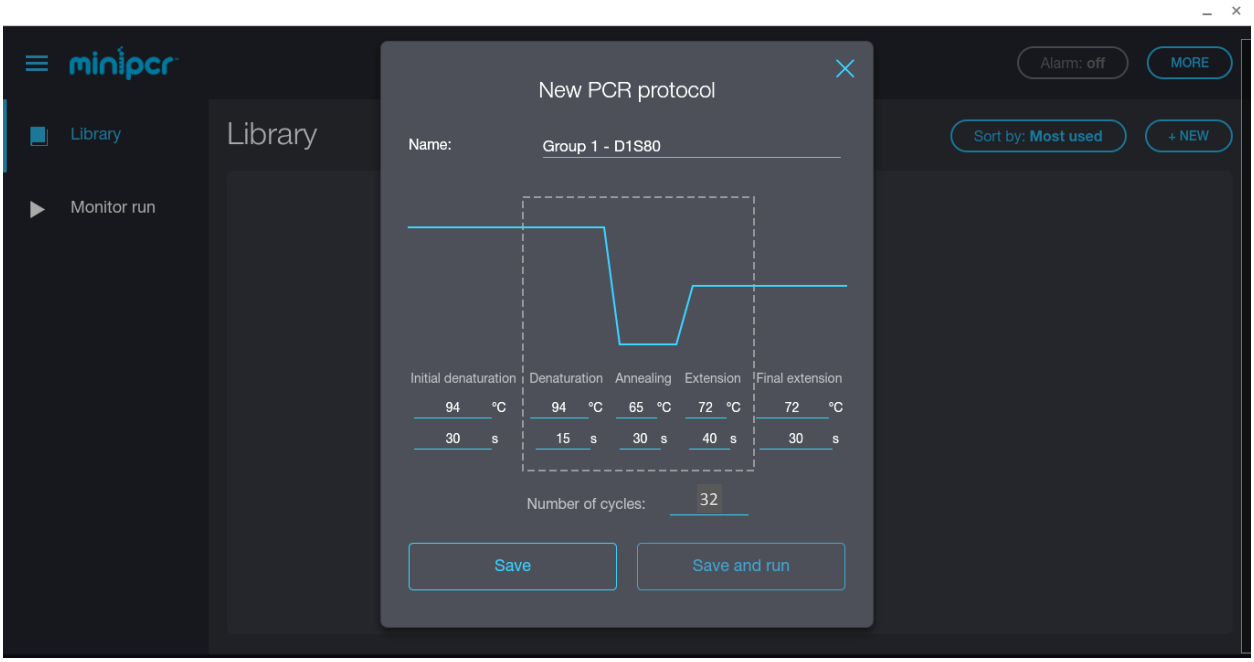
3. **Gently mix the reagents by pipetting up and down 3-4 times, cap the tubes**
    - Make sure all the liquid volume collects at the bottom of the tube
    - If necessary, spin the tubes briefly using a microcentrifuge

4. **Place the tubes inside the PCR machine**
    - Press firmly on the tube caps to ensure a tight fit
    - Close the PCR machine lid and tighten the lid gently

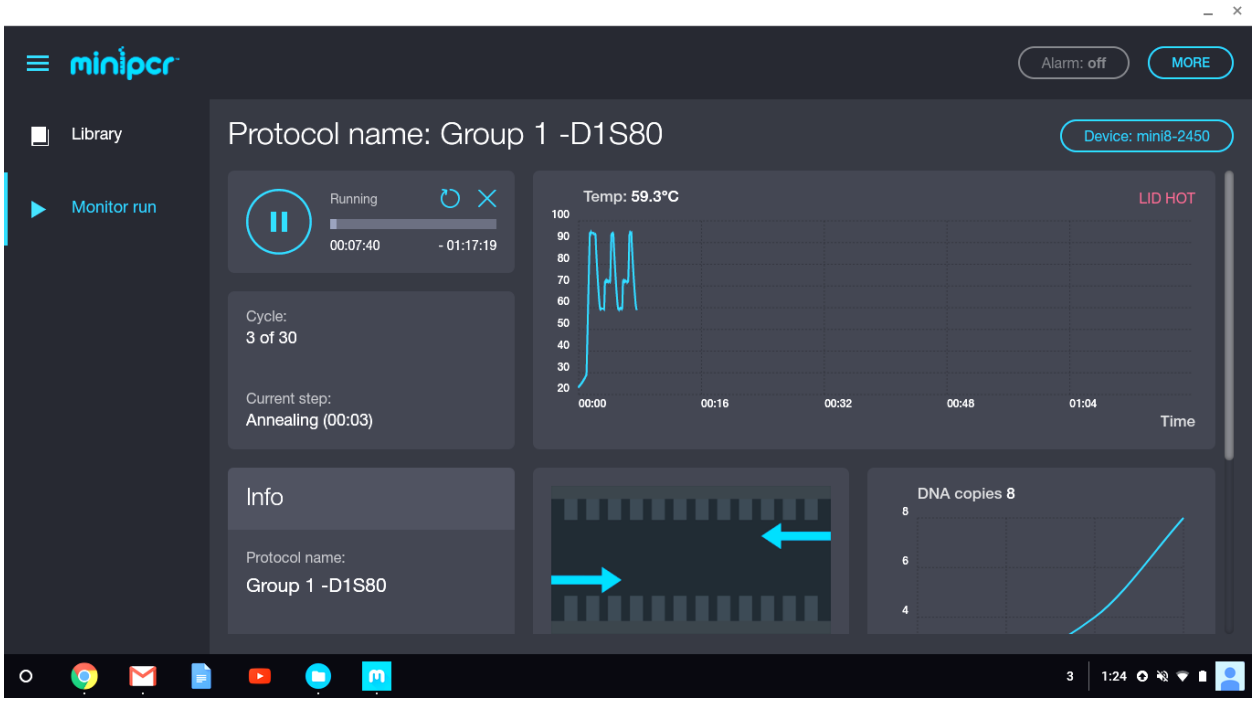
C. PCR programming and monitoring (illustrated using miniPCR® software)

Open the miniPCR software app and remain on the **"Library"** tab

1. Click the **"New"** button on the top right corner
2. Select the "**Create PCR Protocol"** from the drop-down menu



3. Enter a name for the Protocol; for example **"Group 1 – D1S80 Lab"**
4. Enter the PCR protocol parameters:
   - **Initial Denaturation** 94°C, 30 sec
   - **Denaturation** 94°C, 15 sec
   - **Annealing** 65°C, 30 sec
   - **Extension** 72°C, 40 sec
   - **Final Extension** 72°C, 30 sec
   - **Number of Cycles** 32
5. Click **"Save"** to store the protocol or **"Save and run"** to start the program immediately.
6. To start a saved protocol, in the "Library" tab, select your program. Click the ▶ ("play" symbol) to start the program. Make sure that the power switch on the miniPCR is in the ON position
7. Click on **"Monitor run"** tab to begin monitoring the PCR reaction



The miniPCR™ software allows each lab group to monitor the reaction parameters in real time, and to export the reaction data for analysis as a spreadsheet.

Once the PCR run is completed (approximately 70 min), the screen will display:
"**Finished". All LEDs on the miniPCR machine will light up.**

You can now open the miniPCR lid and remove your PCR tubes.

Be very careful not to touch the metal lid which may still be hot

PCR product is stable after amplification at room temperature for several days. For long term storage, move tubes to a freezer (up to 1 month)

D. Gel electrophoresis – Pouring agarose gels (Preparatory activity)

TIME MANAGEMENT TIP

If the lab is going to be completed in a single time block, agarose gels should be prepared during the PCR run to allow the gels to solidify.

If the lab is going to be performed over two periods, gels can be prepared up to one day ahead of the second period and stored in a refrigerator, covered in plastic wrap and protected from light.

1. Prepare a clean and dry agarose gel casting tray
    i. Seal off the ends of the tray as indicated for your apparatus
    ii. Place a well-forming comb at the top of the gel (5 lanes or more)

2. For each lab group, prepare a 1.5% agarose gel using electrophoresis buffer
    i. Adjust volumes and weights according to the size of your gel tray
    ii. For example, add 0.3 g of agarose to 20 ml of electrophoresis buffer (or scale up multiplying by the number of gels poured)
    iii. Mix reagents in glass flask or beaker and swirl to mix

3. Heat the mixture using a microwave or hot plate
    i. Solution thoroughly heated when agarose powder dissolved and solution becomes clear
    
    ii. Use caution, as the mix tends to bubble over the top and is very hot

4. Let the agarose solution cool for about 2-3 min at room temperature.
    i. Swirl the flask intermittently

5. Add gel staining dye (e.g. GelGreen™)
    i. Follow dye manufacturer instructions
    ii. Typically, 10,000X stock is supplied (1 µL of staining dye per 10 mL of agarose solution)

    
    Note: We recommend the use of safe alternatives to ethidium bromide such as GelGreen™ Plus (available at www.miniPCR.com).

6. Pour the cooled agarose solution into the gel-casting tray with comb

7. Allow gel to completely solidify (until firm to the touch) and remove the comb
    i. Typically, ~10 minutes

8. Place the gel into the electrophoresis chamber and cover it with run buffer

Gel electrophoresis – Running the gel

1. Make sure the gel is completely submerged in electrophoresis buffer
    i. Ensure that there are no air bubbles in the wells (shake the gel gently if bubbles need to be dislodged)
    ii. Fill all reservoirs of the electrophoresis chamber and add just enough buffer to cover the gel and wells

2. Load DNA samples onto the gel in the following sequence
    i. **Lane 1:** 10µL DNA ladder
    ii. **Lane 2**: 15µL PCR product from student PCR sample
    iii. **Lane 3**: 15µL PCR product from student PCR sample
    iv. **Lane 4**: 15µL PCR product from student PCR sample
    v. **Lane 5**: 15µL PCR product from student PCR sample
    vi. **Lane 6**: 15µL PCR product from "Crime Scene DNA" sample

Note: there is no need to add gel loading dye to your samples. The *miniPCR Blue PCR Master Mix, 100 bp DNA Ladder,* and *Crime Scene DNA* come premixed with loading dye, and ready to load on your gel!

3. Place the cover on the gel electrophoresis box
    i. Ensure the electrode terminals make good contact

4. Conduct electrophoresis for 20-30 minutes, or until the colored dye has progressed to at least half the length of the gel
    i. Check that small bubbles are forming near the electrode terminals
    ii. Longer electrophoresis times will result in better size resolution
    iii. Lower voltages will result in longer electrophoresis times

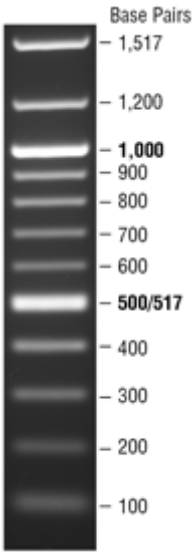
5. Once electrophoresis is completed, turn the power off and remove the gel

E. Size determination and interpretation

1. Place the gel on the transilluminator (or turn on the blueGel™ blue light)

2. Verify the presence of PCR product

3. Ensure there is sufficient DNA band resolution in the 400-700 bp range of the 100 bp DNA ladder
    i. Run the gel longer if needed to increase resolution
    ii. DNA ladder should look approximately as shown *(Source: New England Biolabs)*



4. Document the size of the PCR amplified DNA fragments by comparing the PCR products to the molecular weight reference marker (100 bp DNA ladder)

    i. DNA band size can be estimated by comparing to the DNA ladder. Sizes of bands in the DNA ladder are shown to the right.
    ii. Capture an image with a smartphone camera
    iii. If available, use a Gel Documentation system

3. Study questions

Questions before experimental set-up

1. **What are VNTRs?**

2. **Why can VNTRs be used in personal identification?**

3. **Compare and contrast STRs and VNTRs.**

4. **Why do police use twenty different regions to test someone's identity?**

5. **If you were able to analyze the D1S80 alleles from two people, what would you expect to see if the two individuals were biological siblings? You may want to use a Punnett square to justify your answer.**

6. **What would you expect to see if one individual were the parent of the other?**

7. **Looking at data like this, is it easier to rule someone out as a suspect or to determine that a person is guilty?**

8. **If you have a D1S80 allele with 32 repeats, how big a segment of DNA do you actually expect to see on your gel?**

Questions after gel electrophoresis and visualization

1. **By comparing your DNA bands to the ladder, can you estimate the size of the two fragments on your DNA sample?** (Note: Identifying the precise size of your DNA band is difficult. Use your best estimate for size by comparing your band to the DNA ladder. You should be able to resolve sizes within 30-50 base pairs.)

2. **Using the sizes you estimated in question number 1, can you estimate how many repeats you have in each of your alleles? Remember that the sizes you estimated include the length of your primers and flanking regions (180 bp) and that D1S80 repeats are 16 base pairs long.**

3. **Do you or anyone else have only one band instead of two? What does this mean?**

4. **How do your D1S80 alleles compare to the "Crime Scene DNA" alleles? Can you rule yourself out as a suspect?**

5. **Compare your alleles to other students' alleles on your gel. Does anyone match at both alleles? Does anyone match at only one allele?**

Questions after compiling Class Data

Compile the number of repeats found in each student in your class. A formatted spreadsheet for this task is available at miniPCR.com.

1. **Looking at your class data, does it appear that any two individuals have the same alleles for D1S80? Explain why this would be relevant for forensics analysis. (Remember that band size is an estimate so allow for minor variation between individuals due to lack of precision in size)**

2. **Do any alleles appear to be more common than others? What does this say about the probability of have matching alleles with another person?**

3. **We have said that mutations causing genetic polymorphisms in the D1S80 loci are relatively common compared to the coding regions of genes. Considering, however, that the 18 and 24 allele repeats tend to segregate at significantly higher frequencies than other alleles, how frequently do you expect mutations to actually occur? Nearly every generation? Spread out occasionally over human history? Some other amount?**

Calculating probabilities

1. **Assume that the 18-repeat allele represents approximately 25% of all alleles in the population. What is the probability that an individual will have at least one 18-repeat allele?**

2. **Assume that the 18-repeat represents approximately 25% of total alleles and the 24 repeat represents 35% of total alleles. What is the chance of being an 18/24 repeat heterozygote.**