

DNA Profiling

Clue: The Real Thing

Background

The DNA of humans is more alike than different, but tiny differences make the DNA of each human unique. The technique of DNA profiling can be used to distinguish individual humans using regions of the genome where there is great variability among individuals. Typically these areas of the genome are not working genes and so over time, have been free to vary without selective pressures. These hypervariable parts of the genome are targeted to characterize individuals by their DNA.

The Scenario

Miss Scarlet was found murdered in the library in the middle of the night by Mrs. White, the maid. A candlestick lay near the body in a pool of blood. The list of suspects includes wealthy Mrs. Peacock, Professor Plum and even sweet Mrs. White. Mrs. Peacock, a wealthy middle-aged heiress, was known to hate Miss Scarlet for stealing the attentions of her ex-suitor, Colonel Mustard. Professor Plum, the noted con artist, may have had a reason to silence Miss Scarlet, as only she knew his true identity. Perhaps even Mrs. White had her own reasons. She despised Miss Scarlet and blamed Miss Scarlet for the death of her only son.

1. Make sure you have two pieces of DNA. The paper strip represents the DNA of only one strand in the molecule.

a) Why do you have two pieces of DNA?

2. Imagine that we are digesting the DNA with restriction enzyme EcoRI, whose recognition sequence is 5'- GAATTC -3'. The enzyme makes a cut after the G. You might draw a line after the "G" to represent the cut and then count the number of letters (bases) in each "fragment." Remember that the recognition sequence might appear more than once in any piece of DNA.

3. Record the number of fragments you have and the size of each fragment in the class data table (Table 1a and Table 1b as appropriate).

4. Using the data from Table 1a, create a diagram of the gel that would result from the crime scene DNA and suspects' DNA. You can now see which fragments found at the crime scene match with victim's fragments and with a suspect's fragments. You eliminate the fragments that match with the victim. Looking at the remaining fragments from the crime scene, see if you can match the fragments to the pattern of one of the suspects. Every fragment remaining in the crime scene sample must be found in the suspect's sample or you do not have a match. At this point you should have a suspect that matches the crime scene sample.

b) Why are there more bands in the crime scene lane than other lanes?

c) Which suspect's DNA matches the crime scene DNA?

d) Why is this match not sufficient to conclude guilt?

e) How can we explain the presence of the 12 and 8 letter fragments in samples from both Mrs. Peacock and Mrs. White?

5. You now have to address the question of the probability of the crime scene sample match NOT coming from the suspect. To do this we will find the probability of finding the same fragments in another individual. The probability of finding any one pattern is equal to the number of times that a specific pattern is found divided by the sample size (*i.e.*, the number of people in the population sampled).

f) What is the sample size?

Determine the probability of finding each pattern and complete Table 2.

6. Use the product rule to calculate the probability of another person having the same combination of fragments as the suspect. Remember that the probability of events occurring simultaneously is found by multiplying the probabilities for the occurrence of each individual event.

g) What does the probability do for your confidence that the crime scene DNA is from the suspect?

h) Why is this still insufficient to conclude guilt?

i) The probability you calculated is based on a group of random individuals. Is it reasonable to assume that a random group of people would have been in library in the middle of the night?

What does this suggest about the probability we calculated?

j) What could we do to increase the certainty of our conclusion?

7. In the case of Mrs. White, we might find that the restriction patterns for DNA regions, A, B, C, D, and E, found in her genome occur in the general population with the following frequencies, $1/60$, $1/20$, $1/100$, $1/50$, $1/40$.

k) What is the probability of having the same set of restriction patterns as Mrs. White at these regions?

l) What does that probability suggest about the likelihood of the DNA belonging to her?

Table 1a Fragment sizes for crime scene and suspects (Class data)

	Crime scene	Miss Scarlet	Mrs. Peacock	Prof. Plum	Mrs. White
# fragments					
Size of fragments					

Table 1b Fragment sizes for general population (Class data)

	Gen Pop	Gen Pop	Gen Pop	Gen Pop	Gen Pop	Gen Pop	Gen Pop
# fragments							
Size of fragments							

Table 2 Probability of finding given fragments in the general population

Identity of Sample	Fragment Sizes	# of people with this pattern	Probability of seeing pattern
Crime scene sample	13, 7	1	
	10, 5, 5	1	
	2, 18	1	
	8, 12	1	
Mrs. White	2, 18	1	3/30
	8, 12		5/30
General population	13, 7	2	2/30
(sample size = ?)	10, 5, 5	10	10/30
	2, 18	3	3/30
	8, 12	5	5/30
	6, 4, 10	4	4/30
	3, 10, 7	5	5/30
	1, 19	1	1/30

Mrs. Peacock
5' GAATTCCCATACGAGTTCCC 3'

Mrs. Peacock
5' GGGATTCATACGAATTCCCC 3'

Professor Plum
5' GGAATTCATACGAATTCCC 3'

Professor Plum
5' GGGGAATCATACGGAATTCC3'

Miss Scarlet
5' GGGGGAATTCACGAGAATTC3'

Miss Scarlet
5' GGGATTCCATACGAATTCCC 3'

Mrs. White
5' GGGATTCATACGAATTCCCC 3'

Mrs. White
5' GGAATTCCATACGAGTTCCC 3'

crime scene
5' GGAATTCCATACGAGTTCCC 3'

crime scene
5' GGGATTCCATACGAATTCCC 3'

crime scene

5' GGGGGAATTCACGAGAATTC3'

crime scene

5' GGGATTCATACGAATTCCCC 3'

General population

5' GGAATTCCATACGAGTTCCC 3'

General population

5' GAATTCCCATACGAGTTCCC 3'

General population

5' GGAATTCCATACGAGTTCCC 3'

General population

5' GGGAATTCATACGAATTCCC 3'

General population

5' GGGGAATCATACGGAATTCCC3'

General population

5' GGGATTCATACGAATTCCC 3'

General population

5' GGGGGAATTCACGAGAATTC3'

General population

5' GGGATTCATACGAATTCCCC 3'

General population

5' GGGAATTCATACGAATTCCC 3'

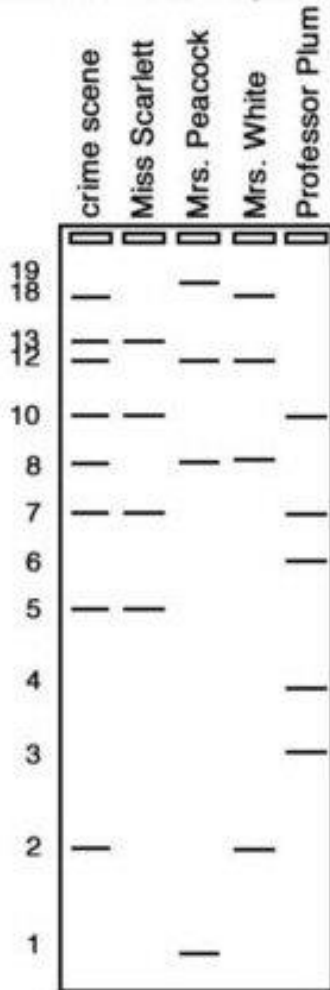
General population

5' GGGATTCATACGAATTCCCC 3'

General population

5' GAATTCCCATACGAGTTCCC 3'

A picture of the agarose gel (autoradiograph) from an EcoRI digest of crime scene DNA and DNA from the suspects.



Teacher Notes

A minimum number of DNA sequences would be two strips of paper for each of the suspects, four for the crime scene, two for the victim, and at least ten for the general population (between 18-22, depending on the number of suspects). The size of the general population sample can be increased so that every student gets their own DNA sequence.

In this case, the pattern of restriction fragments from blood found at the crime scene is compared to the patterns from the potential suspects and the victim. The table and the gel (below) show that all of the DNA fragments (13, 10, 7, 5) found in Miss Scarlett's blood are present in the sample from crime scene. The two five letter fragments show students what might happen when two DNA fragments are the same or close to the same size. Professor Plum's DNA fragments don't match any of the fragments from the crime scene DNA, so he can be ruled out as a suspect. Next, the pattern from Mrs. Peacock's blood contains two fragments (12 and 8) that are the same size as DNA fragments in those from the crime scene. However, the other fragments seen in her blood (19 and 1) are missing from the crime scene sample, allowing her to be ruled out as well. All of the fragments seen in Mrs. White's fingerprint can be found in the crime scene sample making her the most likely suspect.

Statistical analysis. The probability of finding Mrs. White's DNA fingerprint in the general population is then calculated.

The sample size will be determined by the number of students holding the "General population" sequences. This will vary for every class, so it's best to draw a table like the one shown below on an overhead transparency and fill in the numbers during the exercise.

The probability of finding a pattern that matches Mrs. White's, is $= (3/30)(5/30) = 1/60$. One person in a group of sixty people might have the same fingerprint as Mrs. White.

Analysis of multiple sites in the genome. The results of RFLP analysis become more convincing with data from additional regions of the genome. To illustrate why this is true, data from other portions of the genome can be included in this exercise.

For these analysis, it is important that the regions of DNA tested in this process must not be linked. If two sites are linked, then they would tend to occur together and the probability calculations would be invalid. Unlinked sites are often located on separate chromosomes. A typical analysis in a criminal case might examine as many as ten different sites.