Undergraduate Journal of Mathematical Modeling: One + Two

Volume 12 | 2021 Fall 2021 Article 1

2021

Evaluating the Improvement in DNA Fingerprinting

Dani Dray

Advisors:
Arcadii Grinshpan, Mathematics and Statistics
John Angell, Integrative Biology

Problem Suggested By: John Angell

Recommended Citation
Dray, Dani (2021) "Evaluating the Improvement in DNA Fingerprinting," *Undergraduate Journal of Mathematical Modeling: One + Two*: Vol. 12: Iss. 1, Article 1.
DOI: https://doi.org/10.5038/2326-3652.12.1.4937
Available at: https://digitalcommons.usf.edu/ujmm/vol12/iss1/1

Evaluating the Improvement in DNA Fingerprinting

Abstract

DNA fingerprinting is a forensic technique used to create patterns that are unique to a person's DNA. Previously, these fingerprints were made from 13 different segments of DNA, but today they are made from the 20 ones. The fundamental principle of counting is used to determine how much of an improvement was made after adding the 7 additional DNA segments. It is found that this addition greatly reduces the likelihood of two people having the same exact fingerprint, therefore improving the accuracy and reliability of DNA fingerprinting.

Keywords

DNA, fingerprinting, minisatellite, fundamental principle of counting

PROBLEM STATEMENT

In 2017 the FBI added 7 more sites to how a DNA fingerprint is generated. How much did it improve our ability to differentiate between DNA samples?

MOTIVATION

From paternity tests in criminal court cases, DNA fingerprinting is a crucial part of DNA analysis. A DNA fingerprint is composed of sections of human DNA that show more variation from one person to another. These sections of DNA, known as minisatellites, are used to create unique patterns known as a DNA fingerprint. A DNA fingerprint can be used to link DNA left at a crime scene to a suspect. If the patterns made from both sets of DNA are identical, it is safe to say that they came from the same source. This technique can also be used for paternity testing. A child's DNA fingerprint will consist of patterns from the mother as well as the father. Analyzing the patterns between the mother, child, and a few potential fathers provides insight on a possible match. A DNA fingerprint was made up of 13 different minisatellites, or short tandem repeats (STRs), however, in 2017 the FBI increased the amount of minisatellites used to 20. This increase rather improved the accuracy of DNA fingerprinting, but by how much.

MATHEMATICAL DESCRIPTION AND SOLUTION APPROACH

Each minisatellite can be anywhere from 10-60 base pairs long. To simplify things, we will assume that each site is 10 base pairs long. This means that a DNA fingerprint with 13 sites (minisatellites) consists of 130 base pairs, and the one with 20 sites consists of 200 base pairs. Half of each base pair can be any one of the 4 bases: adenine, cytosine, thymine, or guanine. DNA will be made up of these bases in any combination possible. The fundamental principle of counting can be used to calculate the total number of possible combinations and it is given by:

$$n_1 * n_2 * \ldots * n_k , \qquad (1)$$

where n_l ($l=1,\ldots, k$) is equal to the number of possible choices for the corresponding base pair, and k is equal to the total number of base pairs.

DISCUSSION

DNA Fingerprint with 13 Sites:

In this case, each $n_l = 4$ in (1), i.e., it is equal to the number of possible choices for each base pair, and k =130 is equal to the total number of base pairs. This leads to $4^{130} \approx 1.8527*10^{78}$ different possible combinations. The probability of a person having any one of these combinations is $1/(1.8527 * 10^{78})$. Thus, the possibility of two people having the exact same DNA fingerprint is $(1/1.8527 * 10^{78}) * (1/1.8527 * 10^{78})$, or $1/(1.8527 * 10^{78})^2 \approx 2.9 * 10^{-157}$.

DNA Fingerprints with 20 Sites:

As in the previous case, the number of possible choices for each base pair n_l=4, but the total number of base pairs k =200 in (1). This results in $4^{200} \approx 2.5823*10^{120}$ different possible

DOI: https://doi.org/10.5038/2326-3652.12.1.4937

combinations. The probability of a person having any one of these combinations is $1/(2.5823 * 10^{120})$. Thus, the possibility of two people having exactly the same DNA fingerprint is

$1/(2.5823 * 10^{120})^2 \approx 1.5 * 10^{-241}$.

CONCLUSION

Adding 7 new sites to the composition of a DNA fingerprint greatly reduces the chance of two people to have the same fingerprint, thus improving its overall reliability in paternity testing and court cases. The possibility is decreased from $2.9 * 10^{-157}$ to $1.5 * 10^{-241}$ that is, at least, in 10^{84} times. This, however, is an overestimation. In the considered case each minisatellite was composed of 10 base pairs that could appear in any combination. In real life, each minisatellite is composed of a pattern of base pairs that repeat any number of times. The results of the previous calculations do not take into account the repeating patterns, or that different sites can have different numbers of repeats. For a more accurate answer, the author recommends to take these variables into consideration.

REFERENCES

Base pair. (n.d.). Retrieved May 06, 2021, from https://www.genome.gov/genetics-glossary/Base-Pair.

DNA fingerprinting. (n.d.). Retrieved May 06, 2021, from https://www.genome.gov/genetics-glossary/DNA-Fingerprinting.

STEWART, J. (2019). BIOCALCULUS: Calculus probability and statistics for life sciences. In BIOCALCULUS: Calculus probability and statistics for life sciences (p. 729). NEW YORK: CENGAGE LEARNING.

What is a DNA fingerprint? (2016, June 02). Retrieved May 06, 2021, from https://www.yourgenome.org/facts/what-is-a-dna-fingerprint.

DOI: https://doi.org/10.5038/2326-3652.12.1.4937