Forensic DNA Analysis

JESSICA MCDONALD, DONALD C. LEHMAN

LEARNING OBJECTIVES:

1. Discuss the important developments in the history of DNA profiling.
2. Compare and contrast restriction fragment length polymorphism and short tandem repeat analyses in the area of DNA profiling.
3. Describe the structure of short tandem repeats and their alleles.
4. Identify the source of DNA in a blood sample.
5. Discuss the importance of the amelogenin gene in DNA profiling.
6. Describe the advantages and disadvantages of mitochondrial DNA analysis in DNA profiling.
7. Describe the type of DNA profiles used in the Combined DNA Index System.
8. Compare the discriminating power of DNA profiling and blood typing.

ABBREVIATIONS: AFLP - amplified fragment length polymorphism; CODIS - Combined DNA Index System; DNA - deoxyribose nucleic acid; HLA - human leukocyte antigen; mtDNA - mitochondrial DNA; PCR - polymerase chain reaction; RFLP - restriction fragment length polymorphism; RMP - random match probability; STR - short tandem repeats; VNTR - variable number of tandem repeats.

INDEX TERMS: Allele, Amelogenin, Amplicon, DNA profile, Electropherogram, Genotype

Clin Lab Sci 2012;25(2):109

Jessica McDonald, *Senior Student, University of Delaware, Newark, DE*

Donald C. Lehman, EdD, MT(ASCP), SM(NRM), *University of Delaware, Newark, DE.*

Address for Correspondence: Donald C. Lehman, EdD, MT(ASCP), SM(NRM), *Department of Medical Technology. University of Delaware, Newark, DE 19716, (302) 831-8432, dlehman@udel.edu*

Television shows such as CSI: Crime Scene Investigation, Law and Order, Criminal Minds, and many others portray DNA analysis as a quick and simple process. However, these portrayals are not accurate. Since the discovery of DNA as the genetic material in 1953, much progress has been made in the area of forensic DNA analysis. Despite how much we have learned about DNA and DNA analysis (Table 1), our knowledge of DNA profiling can be enhanced leading to better and faster results. This article will discuss the history of forensic DNA testing, the current science, and what the future might hold.

Table 1. History of DNA Profiling

1953	Franklin, Watson, and Crick discover structure of DNA
1983	Kary Mullis develops PCR procedure, ultimately winning Nobel Prize in Science in 1993
1984	Sir Alec Jeffreys & "DNA fingerprinting." by RFLP
1986	First time DNA profiling was used to convict an offender, Colin Pitchfork
1988	First commercial forensic PCR kit detecting SNPs at HLA DQA1 locus
1990	PCR using STR technique adopted
1992	Amelogenin discovered First commercial PCR STR kit First case involving mtDNA
1995	First national DNA database, United Kingdom's NDNAD, established
1996	First time mtDNA was used to convict an offender, Paul Ware
1997	DNA profiling from touched objects and single cells demonstrated
1998	First United States DNA database, CODIS
2000	Databases configured to use STR instead of AFLP

AFLP, amplified fragment length polymorphisms; CODIS, Combined DNA Index System; HLA, human leukocyte antigen; NDNAD, National DNA Database; mtDNA, mitochondrial DNA; RFLP, restriction fragment polymorphism; SNPs, single nucleotide polymorphisms; STR, short tandem repeats. Data derived from Jobling et al., 2004.

History of Forensic DNA Testing

Before 1953, it was unknown what molecules living organisms used to store information and to pass traits to offspring. In 1953, Rosalind Franklin, James Watson and Francis Crick determined DNA's double helix

structure. DNA consists of two twisting strands of polymers held together by hydrogen bonds formed between the complementary base pairing of nucleotides: adenine to thymine and cytosine to guanine. This discovery led to subsequent questions on DNA's variability. How can such a simple molecule provide for the myriad of traits seen in a population? It was discovered that the differences in the arrangement of base pairs are the reason for the differences in individuals. These differences in base pair arrangements not only make individuals look different, but they also give individuals a unique DNA pattern or profile.

In 1984 in Leicester, United Kingdom, Sir Alec Jeffreys paved the way for future advances in DNA profiling by developing the restriction fragment length polymorphism (RFLP) technique.[1] He used this method to determine variations of tandem repeats, patterns of two or more repeated nucleotides, in DNA sequences. This process was referred to as RFLP of variable number of tandem repeats (VNTR)[2] or less formally "DNA fingerprinting." In this assay, enzymes called restriction endonucleases cut the double stranded DNA at specific known nucleotide sequences. The resulting DNA fragments are separated by electrophoresis producing a unique individual pattern. Unfortunately, RFLP analysis is a time consuming process.

Shortly after the development of RFLP analysis, the DNA profiling method was first put to use in a criminal case in 1986. Colin Pitchfork, a United Kingdom resident, was convicted of a double rape and murder because his DNA profile matched DNA found at both crime scenes.[1] RFLP of VNTR using single locus probes was also used in this case to exclude an individual. A suspect, who confessed to the murders despite his innocence, was exonerated because his single locus probe profile did not match DNA found at either crime scenes. The Pitchfork case was monumental because it was the first time DNA profiling was used in a criminal case to convict an offender.

In 1983, Kary Mullis developed the polymerase chain reaction (PCR) technique that became a landmark for DNA testing—an accomplishment that led to a Nobel Prize in science in 1993. PCR allowed for the rapid in vitro replication of specific nucleotide sequences in DNA. The amplified products are called amplicons. PCR has sometimes been referred to as "molecular Xeroxing."[3] The PCR assay has several advantages over the RFLP method. One advantage is that PCR is more sensitive, therefore less starting material is required. With PCR, a profile can be made from a DNA sample that is 100 times less than that of a sample needed for RFLP.[4] Also with PCR, a DNA molecule that is degraded (randomly broken into smaller fragments) can be used and analyzed for DNA typing.[1] This is extremely important due to the fact that much DNA found at crime scenes is degraded. Because of these advantages, the PCR method is now the main system used for DNA profiling.

The application of PCR in forensic science was based initially on amplified fragment length polymorphisms (AFLPs). The PCR of AFLP system used the specific locus D1S80.[4] The D1S80 analysis method was useful because small and degraded samples could be analyzed. In addition, this assay was valuable because the length of the locus, 16 base pairs (bp), allowed for greater variation within the population.[4] Greater variation gave the assay method stronger discriminating power when matching DNA samples.

Despite the success of PCR of AFLP's, in the 1990s DNA profilers switched to PCR of short tandem repeats (STRs); a method that uses much smaller repeat units, only 2 to 7 bp long.[2] STRs are also referred to as microsatellites or simple sequence repeats. The short length of the repeat unit gives them the ability to be amplified more easily, and they are less prone to problems with degraded DNA.[3] Smaller quantities of DNA are required for PCR of STRs. This allows the use of small, partially degraded DNA for analysis.[4] Another advantage is that PCR of STRs allows for multiplexing. With multiplexing several different loci can be analyzed at the same time.[2] Not only does this save time in the laboratory, but it also saves materials and uses a smaller sample size.

STRs are classified by the length of their repeat: mono-, di-, tri-, tetra-, penta- and hexa- nucleotides. Tetra-nucleotides are the most applicable due to the fact that they have a smaller probability of stutter products, amplicons that are one repeat less than the true allele.[5] STR repeat sequences are named by the base composition of the repeat unit (in parenthesis) followed by the number of times it is repeated in subscript, e.g.,

$(GAAT)_3$, a tetranucleotide STR with the sequence GAAT repeated 3 times.

Forensic laboratories in the United States analyze 13 different STR loci. At each locus there are several possible alleles. The alleles have a different number of repeated tandem sequences. Population studies have determined the frequencies of each allele at the individual loci. Once a DNA profile is determined on an unknown sample, it is possible to statistically predict the likelihood that an individual would have a particular allele at a certain locus. Multiplying the probabilities for all 13 loci produces a probability that an individual would have that profile. The probability is on the magnitude of about 1×10^{-15}. This means that the likelihood of two unrelated individuals having the same DNA profile is the reciprocal of the probability or about 1 in 594 trillion individuals.

Principle of PCR

Nuclear DNA is the preferred material for forensic studies because of its large discriminating power. It can be collected from a number of sources. In dried, pooled blood from a crime scene, the white blood cells contain enough DNA for testing. Because red blood cells and platelets lack nuclei, they cannot be used. Buccal cells collected by swabbing the oral cavity are useful for human DNA profiling. Because DNA testing is so sensitive, it is often possible to perform DNA profiling on a used drinking glass and even from fingerprints and perspiration. Only a few nucleated cells are necessary. Because the PCR method uses primers specific for human DNA sequences, the presence of bacterial DNA does not interfere with the results.

End-point PCR is commonly used in forensic laboratories for DNA analysis. After the extraction of DNA from the physical evidence, the first step of the PCR process is amplification of a target DNA sequence. This step generally takes a few hours.[3] The DNA is denatured, or separated into single strands. DNA primers, nucleotides, and DNA polymerase are added. The mixture is placed into a thermocycler which rapidly cycles the reaction vessel through a series of temperature changes. The number of amplicons is doubled after each cycle, producing a geometric increase in the target sequence.

After amplification is the separation stage. One multiplex PCR can produce over 20 different sized, randomly scattered DNA fragments.[5] In order to separate the amplicons to distinguish one from another, electrophoresis is used. The two main types of electrophoresis are slab gel and capillary electrophoresis.[6] Although both methods are effective, capillary electrophoresis is often the preferred method because it is an automated and quicker process.[5] Alleles with more tandem repeats will produce amplicons of larger molecular weight. The larger the molecular weight, the longer it takes the amplicons to move through the electrophoretic field.

The third step in PCR of STRs analysis is DNA detection. The most commonly used method is fluorescence detection. In this process, fluorescent dyes attached to the PCR primers integrate into the amplicons of the STRs. These amplified STR alleles are seen as bands on a slab gel. With capillary electrophoresis, an automated instrument detects the fluorescence and displays them as peaks on an electropherogram.[5] The electropherogram is displayed on a computer monitor and can be printed. Fluorescence detection is favored due to its unique property of multicolor analysis. Several different colored fluorescent dyes are used, meaning different STR loci can be analyzed at the same time.

The peaks or bands are used to create an STR genotype. An STR genotype, or locus genotype, identifies the allele(s) present at a particular locus in a sample. When all of the locus genotypes for an individual are combined, a STR genotype or profile is completed. This profile is what is entered into a DNA database for comparison.[5]

The first commercial kit for PCR of AFLPs was manufactured by the Cetus Corporation in California in 1990. The human leukocyte antigen (HLA) DQA1 kit amplified the HLA DAQ1 locus, an antigen system that plays a role in immune responses. Through a "reverse dot blot" process performed on nylon paper strips, the HLA DQA1 genotype of different DNA samples could be compared to see if they came from the same source.[2] The AFLP commercial kit was replaced by the STR commercial kit in the late 1990's. The STR multiplexing commercial kits combine the process of amplification and labeling using fluorescent primers.[3] The availability of commercial kits lead to the routine

use of DNA profiling in forensic laboratories.

Determining the Gender of DNA Contributors

In 1992, DNA profiling advanced further to include sex based and maternal linked DNA profiling. The amelogenin locus was found to contain a length variation (different number of nucleotides) between males and females.[4] Because of this variation, PCR analysis of amelogenin can be used to determine if DNA is from a male or female contributor.[4] The male amelogenin locus (AMELX) has a six base pair deletion compared to the locus in females (AMELY). Aplicons from males are therefore six base pairs shorter. The discovery of amelogenin was helpful in determining the gender and possible identity of suspects. It was also useful in cases of sexual assault crimes where sexual stains from unknown contributors are found.

It was later discovered that mitochondrial DNA (mtDNA) could be useful in forensic DNA testing. mtDNA has a maternal inheritance which means it is only inherited from the mother.[3] This gives mtDNA the unique quality of being able to track families and people from similar populations. Sequence analysis of mtDNA is frequently used for comparison of questioned and reference samples. mtDNA is also useful because it has a much greater likelihood of survival than that of nuclear DNA, the type of DNA more often used in DNA analysis procedures. The reason for this increased survival is due to mtDNA's high copy number.[1] mtDNA is especially useful in determining the source of hairs because mtDNA is found in the hair shaft not only in the root/bulb, the location of nuclear DNA. This is important because hairs often found at crime scenes do not contain the bulb.[4] mtDNA was first used in a criminal trial in 1996. Paul Ware was convicted of the rape and murder of a 4-year-old child after his hair was found on the victim's body.[1]

DNA Databases

The establishment of DNA databases was another significant advancement in DNA profiling. The first national DNA database was created in the United Kingdom on April 10, 1995.[2] The introduction of a national database allowed forensic scientists to enter unmatched DNA evidence found at crime scenes into a computerized system to make DNA matches. In the United Kingdom's system, suspect profiles are entered into the database as well as convicted offender/crime scene profiles. There are three different types of "hits" categorized by the United Kingdom system: suspect-case cold hits (sample match, no name), case-suspect cold hits (sample match and a name), and case-case cold hits (sample match, no name).[6]

It was only 3 years after the establishment of the United Kingdom's database that the United States would follow. On October 3, 1998, the United States Federal Bureau of Investigation officially launched a nation-wide DNA database.[2] This system, which was being worked on since 1990, was called Combined DNA Index System (CODIS). Before 2000, CODIS was configured to support any RFLP or PCR marker. However, after 2000 only STR data were added.[2] The United States' CODIS accepts 13 tetrameric STR loci and the amelogenin locus.[1]

To find a "hit" in CODIS, a forensic scientist enters a DNA sample's STR profile into the database. The sample could be an unidentified crime scene sample or an individual convicted of a violent crime. A random match probability (RMP) is then calculated for a matching database entry. The RMP is the estimated frequency in which a certain STR profile is expected to occur in a population.[7] In order to calculate RMP, the STR genotype frequency for each locus in the population is used. The allele frequencies for the matching samples at all 13 loci are then multiplied together to determine the DNA profile frequency estimates.

DNA databases are a valuable tool in crime scene investigations. An advancement being considered is standardizing the loci analyzed in each country's database. Currently, only 8 of the 13 loci in CODIS overlap with the 10 loci used in United Kingdom's National DNA Database.[5] The United Kingdom is considering adding perhaps as many as five more loci. Canada, Australia, New Zealand, and Japan, have also established successful DNA databases. One way to improve international DNA database comparisons could be to increase the number of loci tested or make sure the same loci are used in all systems.

Problems could, arise from changing the loci compared in established databases. Replacing old STR loci with newer ones by retesting DNA samples would be an expensive and intricate process. It could also lead to

legacy DNA profiles that are already in the system to be ruled negligible. It would be impossible to accurately match a profile using new loci to a profile that used old loci.[5] However, keeping all current loci and adding new loci would make it possible to match samples currently in the database to those with the same loci plus a larger number of new loci.

The Future of DNA Profiling

Future advancement in forensic DNA profiling could be in the area of new methodologies and instrumentation. DNA profiling is still a lengthy, expensive, and complicated process; it would be helpful to have new technologies that could quicken the process, reduce the costs, and make it easier to do DNA testing at the crime scene. Other improvements could include technology to test older, degraded samples and smaller concentrations of DNA. Another possible advancement is expanding DNA testing to more applications. For example lacing ink samples with synthetic DNA fragments in order to determine the origin on a written document.

A promising new technology is the "laboratory on a chip," or a portable DNA testing device.[1] The basis of this method is analyzing DNA in a miniature capillary electrophoresis process. If the entire capillary electrophoresis process were miniature, this would result in smaller individual capillary channels. The smaller the capillary channels, the quicker the DNA separation. It is possible that it might make DNA separation 10 to 100 times faster.[5] The development of a portable DNA testing device would not only allow for quicker DNA profiling, but also for portability to crime scenes.[1]

Other advancements that are most likely to occur in the future deal with the automation of DNA profiling. Several steps of DNA analysis, such as DNA extraction, PCR setup and amplification, involve manual pipetting. It is projected that robotic liquid-handling platforms would be faster and more accurate.[5] Robotics could reduce the time DNA analysts spend on pipetting tasks, giving forensic scientists more time to spend on important tasks such as data interpretation or completing case work.

SUMMARY

Before the routine use of DNA profiling, blood typing was an important forensic tool. However, blood typing was not very discriminating. For example, roughly 30% of the United States population has type A-positive blood.[8] Therefore, if A-positive blood were found at a crime scene, it could have come from 30% of the population. DNA profiling has a much better ability for discrimination. Forensic laboratories no longer routinely determine blood type. If blood is found at a crime scene, DNA profiling is performed.

From Jeffrey's discovery of DNA fingerprinting to the development of PCR of STRs to the formation of DNA databases, our knowledge of DNA and DNA profiling have expanded greatly. Also, the applications for which we use DNA profiling have increased. DNA profiling is not just used for criminal case work, but it has expanded to encompass paternity testing, disaster victim identification, monitoring bone marrow transplants, detecting fetal cells in a mother's blood, tracing human history, and a multitude of other areas. The future of DNA profiling looks expansive with the development of newer instrumentation and techniques.

REFERENCES

1. Jobling M, Gill P. Encoded evidence: DNA in forensic analysis. Nature Rev 2004;5:739-52.
2. Lynch M, Cole S, McNally R, Jordan K. Truth machine: the contentious history of DNA fingerprinting. Chicago: University of Chicago Press; 2008.
3. Butler JM. Forensic DNA typing: biology, technology, and genetics of STR markers. Burlington: Elsevier Academic Press; 2005.
4. Inman K, Rudin N. An introduction to forensic DNA analysis. Boca Raton: CRC Press LLC; 1997.
5. Butler JM. Fundamentals of forensic DNA typing. Burlington: Elsevier Academic Press; 2009.
6. Asplen C. The application of DNA technology in England and Wales. US Department of Justice. 2004; Doc.203971.
7. Giese H, Lam R, Selden R, Tan E. Fast multiplexed polymerase chain reaction for conventional and microfluidic short tandem repeat analysis. J Forensic Sci. 2009;54:1287-96.
8. American National Red Cross [Internet]. Washington, DC. [cited 2011 Dec 19]. Available from: http://www.redcrossblood.org/learn-about-blood/blood-types.