DECRYPTING DNA FINGERPRINTING: QUEST FOR THE TRUTH HIDDEN IN DNA

SABUJ SAMADDAR | ASSIGNMENT | GENETIC ENGINEERING | SCHOOL OF BIOTECHNOLOGY | INSTITUTE OF SCIENCE | BANARAS HINDU UNIVERSITY

From the discovery of the Deoxyribonucleic Acid (DNA) by James Watson and Francis Crick on 1953 has led the path of the journey in the genetic exploration, which has been nothing short of extraordinary, opening the "Door of El Dorado" to a new era of Scientific enquiry with sparking curiosity about the intricacies of DNA and its role in hereditary. The study of human genetics by using direct DNA-based techniques has developed exponentially ever since, giving birth of DNA fingerprinting, one of the greatest discoveries of the late 20th century.

THE JOURNEY: FROM DNA TO DNA FINGERPRINTING

DNA, the hereditary material within all the living organism, primarily residing inside the cell nucleus, organised into thread-like structure called chromosomes, where the DNA is associated with protein called histones, maintaining a highly compact structure without disrupting the native fundamental structure of DNA. The Diversity among all the individual who have lived and currently living on the face of the earth arises from the unique combination of four nucleotides: Adenine(A), Guanine(G), Cytosine(C), Thymine(T), these units are playing a fundamental role in determination of genetic differences. Human Genomic DNA contains approx. 3 billion nucleotide positions, coding for 50000-100000genes, these genes are situated at discrete location called 'loci' on chromosomes, exist in multiple forms known as allele, giving birth of 'Genetic Polymorphism'. Investigators have identified >2500 polymorphic loci in the human genome. Two main types of DNA polymorphism are studied in forensic science: **Single base changes or single nucleotide polymorphisms (SNP)** and **minisatellite sequences with VNTRs (variable number of tandem repeats).** Both forms of polymorphism are typically identified by variations in DNA fragment size after being digested by restriction enzymes.

Single nucleotide polymorphisms (SNPs) have a small number of alleles at these specific locations, making them less useful for forensic identification using DNA.

Whereas, **Variable Number of Tandem Repeats (VNTRs)** are a class of polymorphic loci that exhibit high variability in humans. They possess numerous alleles, high heterozygosity rates, and are widely distributed throughout the genome, making them invaluable for human gene mapping, individual identification, and assessing biological relatedness. VNTRs, also known as length polymorphisms, feature core elements of nucleotide repeats arranged in tandem. The number of repeat units varies between individuals, and the size of the core repeat differs among loci, rendering each genetic profile unique. VNTR polymorphisms, typically ranging from 9 to 80 base pair repeats, offer stability across generations and serve as distinctive genetic markers for individual identification. The human genome harbours over a hundred VNTR or hypervariable minisatellite (HVR) loci, each characterized by its unique pattern of repeat variation. Only a small amount of DNA nucleotide leads to the genetic variability among the individuals, 0.3% (10million nucleotides) of whole gene are the main players behind the exhibiting feature of genetic uniqueness. This small amount of variation leads to the development of DNA fingerprinting, the use of information embeds in the DNA for identification purpose.

SERENDIPITY IN SCIENCE: UNEXPECTED ORIGINS OF DNA FINGERPRINTING

The discovery of DNA fingerprinting was bit unintentional! In 1984, at Leicester University Alec Jeffreys and his group of students were working in the study of the evolution of the globin genes present in the sea-dwelling mammals, observed a specific 33-nucleotide sequence recurring in various repetitive arrangements across nearly all globin molecules. Jeffreys and a graduate pupil entered the lab on a Monday morning to review the trial's results; Intrigued by the potential relatedness of these hypervariable regions, Jeffreys decided to isolate the minisatellite sequence from the globin gene and utilize it as a probe for studying human DNA. To his surprise, Jeffreys found that the minisatellite probe yielded a complex pattern of bands for each individual, reminiscent of a barcode. This pattern, in fact, represented numerous hypervariable regions within the human DNA that were interconnected, with each band corresponding to one of these regions. Further refinement of the DNA sequences and methodologies led to the realization that this barcode-like pattern provided a distinct identification signature for each individual, hence coined as DNA 'fingerprint'. Following the publication of this discovery in Nature in early 1985, Jeffreys and

associates delved the frequence of colourful mini-satellite fractions in 20 unconnected British Caucasians. In a posterior Nature paper latterly that time, they determined that the probability of two individualities having the same DNA point was lower than 1 in 33 billion.

THE INTRICATE PROCESS OF DNA FINGERPRINTING

DNA fingerprinting is a method used to analyse hypervariable sequences within DNA molecules. It starts with the extraction of DNA from various sources, ranging from blood samples in paternity disputes to minute tissue samples in forensic investigations. The extracted DNA is then treated with restriction enzymes to generate numerous small DNA fragments containing minisatellite regions. These

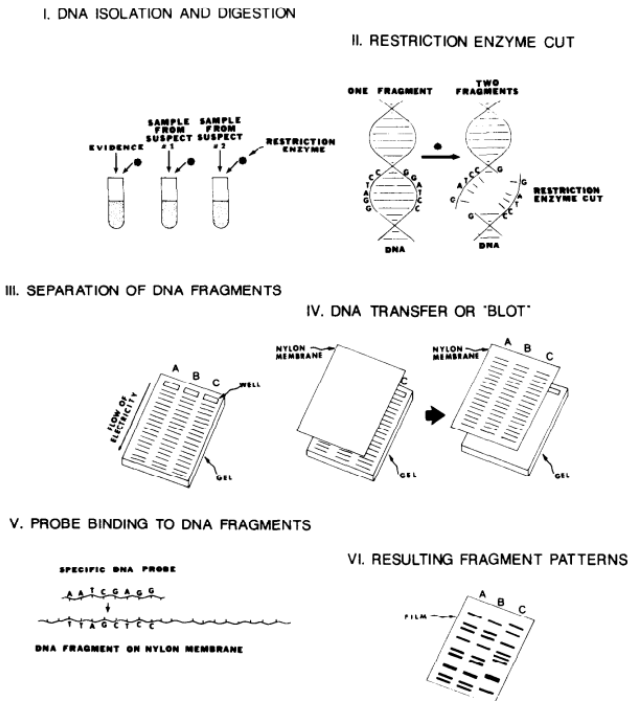

Figure:1 Standard Procedure of DNA Fingerprinting:1. Collection of Tissue Samples, 2. Fragment Generation, 3. Size-Based Separation, 4. DNA Fragment Treatment, 5. Hybridization with DNA Probe, 6. Match Identification.

fragments are separated by size using electrophoresis and transferred onto a nylon membrane.

Once on the membrane, the DNA fragments are permanently bound and subjected to detection using synthetic radioactive probes. These probes bind to specific hypervariable DNA sequences, creating a distinct pattern. Detection involves placing X-ray film against the membrane, which darkens wherever hypervariable sequences are present, forming characteristic bands.

In single locus probe analysis, radioactive probes detect only two bands instead of multiple bands, focusing on a single hypervariable region. The resulting banded pattern serves as a unique genetic fingerprint, ready for analysis. This technique, derived from medical genetics, relies on polymorphisms within DNA, specifically variable numbers of tandemly repeated (VNTR) units. The DNA is digested with restriction enzymes, fragmenting the VNTR regions into various lengths. These fragments are then transferred to a nylon membrane and exposed to radioactive probes, allowing for the identification of specific VNTR bands. The correspondence in position between VNTR bands from different samples is crucial for identification. By comparing DNA patterns between forensic specimens and reference samples, such as blood from suspects, individuals can be identified or excluded from consideration. Thus, DNA fingerprinting involves isolating DNA, fragmenting it, transferring fragments onto a membrane, detecting hypervariable sequences using radioactive probes, and comparing patterns to determine identity or relatedness.

EVOLUTION OF DNA FINGERPRINTING

The advancement of forensic DNA profiling has undergone significant technological evolution. Traditionally, in DNA fingerprinting, radio-labelled DNA probes containing specific sequences are hybridized to DNA fragments digested with restriction enzymes. These fragments, varying in size due to repeat units, are separated and visualized through autoradiography. However, this method faced limitations such as difficulty in precise matching between fragments and the need for substantial amounts of high-quality DNA. To address these challenges, single-locus profiling emerged which is able in detecting hypervariable loci using specific probes with high stringency hybridization. Despite its success, the method still relied on the quality and quantity of DNA samples, making comparisons between different sources challenging. In the early 1990s, PCR-based methods began replacing RFLP analysis due to improved sensitivity, speed, and precision. Short tandem repeats (STRs), known as microsatellites, became favoured for their sensitivity, low allelic dropout rates, and high discrimination.

RESENT IMPROVEMENTS IN DNA FINGERPRINTING PROCESS:

Forensic DNA profiling now utilizes multi-allelic STR markers, offering a unique genetic code for each individual. Two sets of standard markers,

European and US CODIS, are commonly used worldwide. Incorporating these markers into commercial kits has enhanced their application, allowing reproducible results even from minimal DNA samples. The establishment of government-controlled DNA databases in developed nations, starting with the UK in 1995, has facilitated criminal investigations. Cold hits, linking crime scene samples to offenders in these databases, serve as valuable investigative leads. China, the US, and the UK maintain the largest DNA databases, with increasing percentages of the population being included. However, the proportions vary significantly between countries, highlighting differences in their approach to DNA database utilization.

SOLVING THE HARDER PARADIGM IN DNA FINGERPRINTING WITH LINEAGE MARKERS:

Lineage markers are crucial in forensic genetics, particularly in cases where traditional DNA analysis methods face obstacles.

Y chromosome analysis, for example, proves invaluable when dealing with scenarios where there's a high proportion of DNA from a female victim and only minimal DNA from a male perpetrator. This is often seen in cases of sexual assaults without ejaculation or when trace amounts of male DNA are found on the victim's body or belongings.

Mitochondrial DNA (mtDNA) analysis is especially useful for examining low-level nuclear DNA samples, such as those obtained from unidentified remains or ancient specimens where DNA degradation is significant. Although the non-recombinant mode of inheritance of Y and mtDNA weakens statistical comparisons, it is highly efficient in reconstructing paternal or maternal relationships, which is crucial in cases like mass disaster investigations or historical reconstructions.

A notable example illustrating the power of mtDNA analysis is the identification of two missing children of the *Romanov family*, where mtDNA analysis provided irrefutable evidence of their identity. Additionally, the widespread acceptance of Y-STR analysis, which offers enhanced discriminative power with panels including rapidly mutating markers, has significantly advanced forensic genetics. To estimate match probabilities, databases of haplotypes sampled from diverse populations are essential. The YHRD and EMPOP databases, with their extensive collections, facilitate such analyses. Moreover, the geographic distribution of lineage markers can provide valuable intelligence, aiding investigations by establishing links between genetics, genealogy, and geography. An illustrative case demonstrates how Y-STR haplotyping helped identify suspects in a violent crime by inferring their population affiliation. Such biogeographic analyses rely on large reference databases and contribute significantly to resolving cases with complex DNA evidence. Furthermore, information on biogeographic origin can also be obtained from ancestry informative SNPs or mini-haplotypes, providing additional insights into familial or ancestral genetic pools. These markers, akin to Y and mtDNA, are instrumental in connecting individuals to their familial or ancestral lineages, thus offering valuable tools in forensic investigations.

TECHNQUES OF DNA FINGERPRINTING:

RESTRICTION FRAGMENT LENGTH POLYMORPHISM (RFLP)

RFLP analysis relies on the presence of restriction enzyme recognition sites flanking the polymorphic locus under examination. These enzymes exhibit precise cleavage patterns, resulting in the generation of distinct DNA restriction fragments.

The fundamental principle of DNA profiling revolves around restriction fragment length polymorphisms (RFLPs). Restriction enzymes recognize specific base pair sequences, enabling them to cleave the DNA molecule at precise sites termed "restriction sites." Subsequently, electrophoresis is employed to segregate DNA fragments based on their size, with smaller fragments migrating farther over time, forming distinct bands. An RFLP probe, a labelled DNA sequence, binds to specific fragments of the digested DNA sample post-electrophoresis, revealing a unique blotting pattern characteristic of a particular genotype at a given locus. These probes find utility in genome mapping and various analyses such as genotyping, forensics, paternity testing, and diagnosing hereditary diseases.

POLYMERASE CHAIN REACTION (PCR)

PCR, an in vitro enzymatic amplification technique for small DNA segments, was invented in 1985 and later adapted for forensic applications. It exploits DNA replication principles, with DNA polymerase synthesizing new complementary strands using single-stranded DNA templates and oligonucleotide primers. PCR begins with denaturing DNA strands and primer annealing to specified template points, initiating DNA synthesis. With each cycle, DNA content theoretically doubles, making PCR a potent forensic tool requiring minimal DNA quantities, enabling typing from minute

samples like single hairs or epithelial cells in cigarette butts. *Compared to RFLP, PCR-based methods offer advantages such as speed, minimal DNA requirements (ng), compatibility with degraded DNA, high discrimination power, automation, and high sample processing capacity.* Consequently, DNA extraction samples have been reduced to small tissue fragments, single hairs, or even blood droplets, facilitating applications in crime scene investigations, paternity testing, military identification, and combating animal poaching.

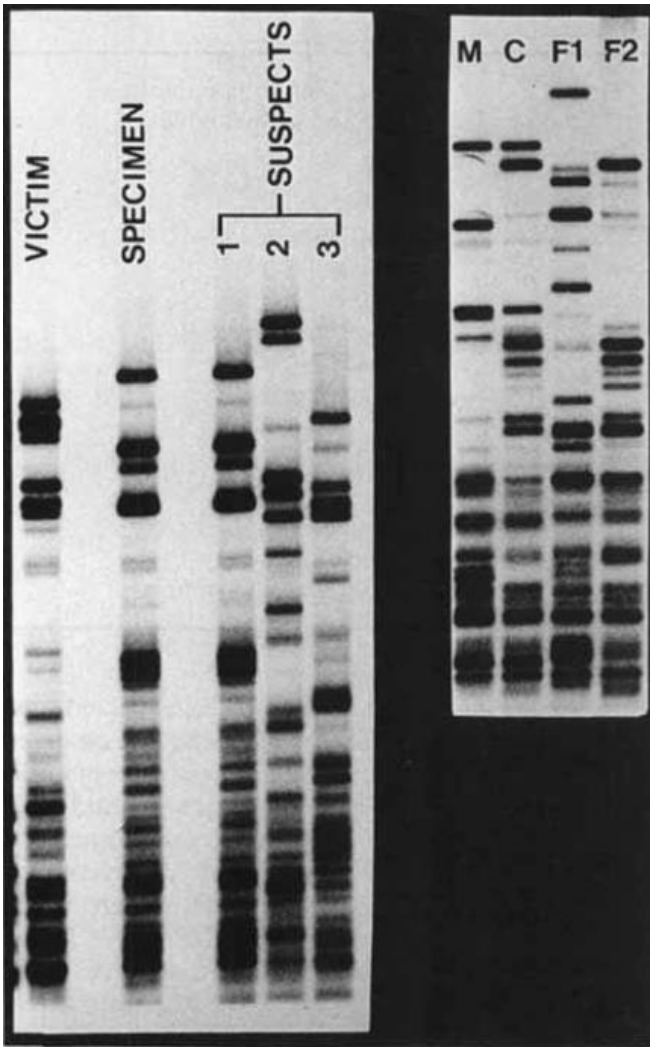

Figure 2: Left: Illustrative Case of Murder Investigation Using DNA Fingerprints. In a simulated murder case, DNA fingerprint analysis revealed that the blood stain discovered on the victim did not match the victim's DNA. Among three potential suspects, a comparison of their blood samples with that found on the victim revealed a match with suspect I. Right: Exemplary Paternity Determination In a paternity determination scenario, the DNA fingerprint pattern of the child (C) is expected to reflect a combination of genetic markers inherited from both biological parents. By comparing the child's DNA bands with those of the mother (M), it becomes possible to identify the genetic markers inherited from the biological father. Analysis of the DNA fingerprints of the potential fathers (F1 and F2) indicates that only F2 shares the genetic markers required to be the father of the child.

ETHICAL IMPLICATION IN DNA FINGERPRINTING

ESTABLISHING CLEAR BOUNDARIES FOR VNTR MARKER ANALYSIS

In the analysis of VNTR markers, precise boundaries are crucial to differentiate between DNA fragments of similar sizes. For instance, while a fragment containing 20 VNTR units can be distinguished from one with 10 units, it becomes challenging to differentiate between 20 and 21 units. To manage this, measurements of VNTR bands are categorized into "bins" comprising bands of comparable sizes. The width of each bin directly impacts probability estimates, as wider bins increase the likelihood of two distinct bands being grouped together erroneously.

To mitigate this issue, testing multiple VNTR markers is essential to enhance the likelihood of detecting clear differences in band sizes.

THE FYRE CRITERIA AND LEGAL DEBATE IN COURTROOM CONTROVERSY

The controversy surrounding the acceptance of DNA testing in courtrooms largely stems from the matching procedure and its statistical significance. While apparent matches based on conventional analysis of blood type and serum enzymes carry a chance of genetic identity due to coincidence ranging from 1 in 100 to 1 in 1000, the estimates for RFLP matches range from 1 in 100,000 to 1 in 1 million. Such strong evidence can wield considerable influence in courtroom proceedings, leading defence arguments in criminal cases to focus on the statistical power of a DNA match. The integration of DNA testing into the U.S. legal system has been guided by the Frye criteria, which stipulates that a new scientific technique must gain general acceptance in its respective field before presentation to the jury. Today, most jurisdictions in the United States have established precedents based on the Frye criteria, affirming the admissibility of DNA-based evidence.

However, in jurisdictions where DNA fingerprinting has not been fully embraced, the principal concern revolves around the interpretation of the test's statistical power. The term "fingerprint" aptly describes this phenomenon, as combinations of VNTR markers offer the potential for a pattern that is essentially unique to each individual, making the method potential evidence in jurisdiction operations.

SAFEGUARDING CIVIL LIBERTIES IN THE AGE OF FORENSIC DNA DATABASES

In democratic societies, civil rights and liberties are paramount, and proposals to expand forensic DNA databases to encompass entire populations warrant condemnation. Alec Jeffreys, an early critic, raised concerns about the UK police's approach to DNA profiling, which includes not only convicted individuals but also arrestees without convictions, cleared suspects, and even innocent individuals never charged with an offense. He also highlighted the socioeconomic bias inherent in large national databases like the NDNAD of England and Wales. Critics have pointed out that the majority of database matches relate to minor offenses, with theft accounting for 63% of matches in Germany, while less than 3% relate to serious crimes like rape and murder. In response to legal challenges, the UK amended its database regulations in the Protection of Freedoms bill of 2012, leading to the removal of 1.1 million profiles of innocent individuals from the database by May 2013. Similarly, when Portugal proposed a DNA database containing samples from every citizen in 2005, public outcry led to the limitation of the database to only include samples from criminals. A recent study on public attitudes towards DNA databases revealed that a more critical stance towards widespread databases is associated with the age and education level of respondents. Building greater public awareness regarding the benefits and risks of extensive DNA collections is essential.

DNA FINGERPRINTING BIASNESS: NAVIGATING VARIABILITY AND DISCRIMINATORY POWER

Concern arises from the reliance on the frequency of each VNTR marker within the broader population to determine the statistical power of DNA fingerprinting. Certain subgroups or even familial groups may exhibit a notably higher occurrence of a specific band compared to the general population. This variation complicates the accurate calculation of the statistical power of DNA fingerprinting. However, despite these challenges, VNTR markers continue to be favoured in medical genetics due to their inherent variability, which grants them significant discriminatory power, even within closely related individuals or families.

CHALLENGES IN THE METHOD

1. DNA DEGRADATION: UNVEILING ISSUE

In the realm of crime investigation, tissue samples are often transported to molecular genetic laboratories for analysis long after cell death has occurred. However, the inherent nature of cells dictates that they undergo self-destruction soon after death, leading to the degradation or breakdown of all subcellular components, including DNA. The duration for which cells remain unattended outside the body directly correlates with the extent of DNA degradation. The degradation of nucleic acids poses significant challenges for several reasons. Not only does it diminish the total amount of DNA available for analysis, a critical concern given the typically minimal recovered tissue, but it also hampers the detection of large restriction enzyme fragments. When DNA undergoes breakdown into smaller fragments before reaching the laboratory and is subsequently digested with appropriate restriction enzymes, the resulting cut fragments may be smaller than anticipated. This phenomenon leads to the generation of non-matching patterns, complicating the analysis process further.

2. ELUCIDATION BAND DISPLACEMENT

Band displacement refers to the subtle and gradual variances in the movement of corresponding DNA fragments from different samples. The central question revolves around determining when a minor migration difference suggests a shift in identical fragments versus a genuine mismatch. While it can lead to confusion, it's crucial to understand that band displacement alone doesn't necessarily imply a mismatch. Initially, the threshold for distinguishing bands in DNA fingerprinting was established at three standard deviations, with each deviation representing approximately 0.6 percent of a fragment's molecular weight. Presently, most laboratories adopt a cutoff point of approximately two-and-a-half percent of a fragment's molecular weight, with some opting for even lower deviations.

Additionally, employing positive and negative controls, as well as mixing experiments, aids in differentiating between these scenarios. Band displacement can occur due to various factors. DNA fragments are loaded into wells at one end of the gel, followed by the application of electricity. Despite being loaded sequentially, some diffusion of fragments occurs post-introduction, with early-loaded samples showing more diffusion. However, rapid loading minimizes diffusion and shouldn't result in band displacement.

Another potential cause of band displacement is the degradation of nucleotides at the ends of DNA

fragments, leading to slightly shorter fragments, often found in samples from crime scenes.

3. INADEQUATE DATABASE ASSESSMENT

When comparing DNA fragments from a crime scene to those from a suspect, determining the likelihood of a match occurring by chance alone is essential. Accurately estimating this probability requires knowledge of the frequency at which a specific set of fragments appears in a given human population. Two primary considerations arise here. Firstly, the size of the DNA fingerprint database used to estimate RFLP frequency is crucial. A large sample size is necessary to obtain a representative cross-section of the population and minimize estimation errors. Secondly, the representativeness of the tested group in relation to the suspect's population is paramount. The human population is diverse, comprising various subgroups, often referred to as races, each exhibiting slight genetic variations. Polymorphic DNA fragment differences are documented among ethnic groups, indicating the need for precise comparisons with the suspect's DNA and the population data available. The accuracy of comparing the suspect's DNA with that of a given population hinges on whether the suspect belongs to the group for which frequency data are accessible. If the perpetrator belongs to a different racial or ethnic group, the fragment pattern may be exceptionally rare or common within the screened population. Without a robust scientific database, estimates of finding a specific fragment pattern within a population are likely to be flawed.

Moreover, similar challenges arise in racial or ethnic identification of suspects. Incorrect classification from a genetic perspective can lead to less precise probability calculations. However, such errors won't alter the rarity or commonness of the probability; they simply introduce inaccuracies into the assessment.

THE FUTURE AND PROMISES

In contemplating the trajectory of forensic DNA Fingerprinting, one is confronted with myriad challenges and possibilities. While certain issues can be addressed with existing solutions, others, like uncertainties surrounding band shifting and database inadequacies, may require novel approaches or technologies.

An imperative aspect is the refinement of databases, necessitating deeper exploration into genetic homogeneity within distinct human populations and racial or ethnic clusters. Questions persist regarding the necessity of separate databases for groups like Puerto Ricans and Mexicans, as well as the degree of intermingling within populations, prompting ethical dilemmas regarding genetic data usage. Critical to advancements is probe development, particularly expanding probe capabilities to bind with smaller DNA fragments, which could enhance result accuracy and standardization across laboratories. Moreover, DNA sequencing advancements offer unprecedented potential, enabling comprehensive identification even amidst discrepancies. Innovations such as polymerase chain reaction (PCR) present opportunities for mass DNA production from minimal samples, potentially revolutionizing forensic analysis, especially in crime scene evidence amplification.

Additionally, incorporating complementary RNA probes could heighten sensitivity and decrease DNA sample requirements. Looking ahead, the shift towards DNA sequencing, facilitated by Next Generation Sequencing (NGS) technologies, holds promise. Despite current limitations, improvements in accuracy and efficiency are foreseeable, potentially streamlining police investigations and altering DNA collection procedures.

The integration of DNA analysis into law enforcement practices has already reshaped procedures, as demonstrated by recent court rulings in the US. However, global adoption faces hurdles due to varying national legislations and ethical considerations. Forensic DNA analysis has become an integral part of global justice systems, offering hope in identifying victims of tragedies and aiding investigations. Yet, alongside its benefits, there are growing concerns about societal and ethical implications that must not be overlooked. In the evolution of forensic DNA analysis, while optimism abounds, critical inquiry into its societal impact and ethical dimensions remains essential too.

REFERENCES

1. Herrera, R. J., & Tracey, M. L. (1992). DNA fingerprinting: Basic techniques, problems, and solutions. *Journal of Criminal Justice*, 20(3), 237-248. https://doi.org/10.1016/0047-2352(92)90048-E
2. Debenham, P. G. (1991). DNA fingerprinting. *The Journal of Pathology*, 164(2), 101-106. https://doi.org/10.1002/path.1711640203
3. Aronson, J. D. (2005). DNA fingerprinting on trial: The dramatic early history of a new forensic technique. *Endeavour*, 29(3), 126-131. https://doi.org/10.1016/j.endeavour.2005.04.006

4. Krishnamurthy, V., Manoj, R., & Pagare, S. S. (2011). Understanding the basics of DNA fingerprinting in forensic science. *Journal of Indian Academy of Oral Medicine and Radiology*, 23(4), 613-616. DOI: 10.5005/jp-journals-10011-1233
5. Roewer, L. DNA fingerprinting in forensics: past, present, future. Investig Genet 4, 22 (2013). https://doi.org/10.1186/2041-2223-4-22
6. Billings, P. R. (1995). DNA on Trial. *New England Journal of Medicine*, 333(9), 602. DOI: 10.1056/NEJM199508313330919