Amazing Advances in Forensic DNA Analysis – past, present and the future

Dr Charan Kishor Shetty, Dr Vidya Bhagat

Abstract— DNA fingerprinting, one of the great discoveries of the late 20th century, has revolutionized forensic investigations. Forensic DNA (deoxyribonucleic acid) analysis or DNA profiling has played a major role in the criminal justice system. New techniques and technologies for DNA profiling continue to evolve every year. This review briefly recapitulates 30 years of progress in forensic DNA analysis which helps to convict criminals, exonerate the wrongly accused, and identify victims of crime, disasters, and war. This paper reviews the literature reported during January 2011 through June 2013 in the field of forensic DNA analysis. Recent advances in almost all aspects of DNA analysis – which include sample collection, storage, and pretreatment, DNA extraction, DNA quantitation, quality assurance of DNA testing, and DNA databases are discussed.

Index Terms— DNA fingerprinting, Forensic DNA profiling, PCR, Short tandem repeat, Lineage markers, Forensic DNA database, CODIS.

1 Introduction

DNA is the chemical code that is found in every cell of an individual's body. Although approximately 99.9 percent of human DNA sequences are the same in every person, forensic scientists are only interested in the 0.1 percent of the DNA that is unique to each individual. As a matter of fact, the likelihood of two unrelated individuals having the exact same DNA profile is ~10-15, or about 1 in 594 trillion individuals [1]. Typically, the following steps are performed during forensic DNA analysis: 1) Sample preparation: crime-scene evidence is collected, stored, and transported to an accredited DNA laboratory; 2) DNA extraction: DNA is isolated from the unknown crime-scene evidence (and/or any bodily fluids from the suspect); 3) DNA amplification: certain regions of DNA are replicated exponentially in order to generate detectable amounts of DNA samples for subsequent analysis; 4) DNA quantitation: DNA fragments of different sizes are separated and detected spectrophotometrically; and 5) DNA profile matching: the profile obtained from the crime-scene evidence is either entered into a DNA database for comparison to locate a possible person of interest, or is compared directly with that from the suspect to determine whether the suspect contributed the DNA at the crime scene.

The increasing public visibility of state-of-the-art forensic methods (e.g., DNA profiling) has been fueled by the growing popularity of several TV shows such as CSI: Crime Scene Investigation, Bones, and Forensic Files etc., and is reflected by a steady increase in enrollment in many forensic science degree programs across the country, since CSI first aired in 2000 [2]. While the general public may regard DNA analysis as a quick and simple process, which can provide infallible evidence for forensic investigations, the reality of forensic science is far less clear and certain than what is portrayed on television [1-3]. Forensic DNA analysis has played an increasingly vital role in the criminal justice system. But as this role expands, many social, legal, and ethical concerns are raised [1]. Among these, how to prevent individual DNA profiles from unauthorized use is, perhaps, of paramount concern. Another important concern is the reliability and robustness of DNA testing itself, as many possible errors exist during the various steps of DNA analysis including evidence collection, sample storage, and DNA extraction etc., which can lead to accusation of the wrong persons.

As DNA analysis becomes more the norm for forensic procedures, other previous techniques have begun to be re-evaluated for effectiveness. In the past 2.5 years (January 2011- June 2014), new technologies for every step of forensic DNA analysis continued to emerge. In this review, recent developments with particular emphasis on new techniques for manipulating and analyzing DNA are summarized.

2 The Start: Blood Groups

To start before we had DNA technology and all of the advanced criminalistics capabilities that we have today. All we had to work with prior to these fantastic advances were the ABO blood group system, which had been discovered by the Austrian scientist Karl Landsteiner in 1901. It wasn't until the late 1970's the serologists, because of the limitations of the ABO system, began to look at biochemical markers. Forensic scientists were examining enzymes found on the red cell membrane. These PGM's (Phosphoglucomutase) or genetic markers were protein enzymes which were found throughout the entire body. The PGM1 was also found in semen, which increased its value in forensic serology because of two alleles. Designated "1" and "2" PGM-1, PGM-2 PGM 2-1. This discovery of three phenotypes, which provided additional genetic information about the blood or sperm recovered from a crime scene, was an exciting forensic advance in the early 1980's.

- *Author Dr Charan Kishor Shetty is a Medical Lecturer, Department of Forensic Medicine, Faculty of Medicine and Health Sciences, Universiti Sultan Zainal Abidin (UnisZA), Kampus Kota, Jalan Sulta, Mahmud,Kuala Terengganu, 21300, Terengganu, Malaysia. Telephone no. 609-6275546/609-6275740., E-mail: shettykishor.k@rediff.com*

This progress took on speed with the scientific analysis of Deoxyribonucleic Acid (DNA) It was learned that certain areas of DNA vary quite dramatically from one individual to another individual and that these polymorphic regions are so unique to each individual that this technology could be used as a forensic tool.
It wasn't until 1986 that the first forensic application of DNA typing was performed by Dr. Alec Jeffries in England, who was able to match the DNA of a suspect to the biological materials recovered from the bodies of three lust murder victims. The first use of DNA Typing in the U.S. Courts was the Tommy Lee Andrews case involving a series of rapes in Orange County, Florida in November of 1987. This was followed by the Serial Murder case in Virginia of Timothy Wilson Spencer. (1984-1987). Timothy Spencer was the first Appellate Division ruling on DNA and was the first execution based on DNA. However, DNA wasn't always considered such a precise and powerful weapon.

3 DNA Testing Today

Since then, millions of forensic DNA tests have been conducted in the United States and around the world. In a major advance, the analysis of DNA has evolved from a laborious process taking weeks or even months to a procedure that can be completed in a matter of days. The DNA molecule can establish the link between evidential DNA with that of the possible suspect's DNA. It can identify whether the DNA in question is human or non-human and can be used to establish the sex of the specimen.

The RFLP technology of the 1980's, which was the Model T-Ford of DNA analysis, involved the process of identifying the polymorphic regions that are unique to each individual. These Variable Number of Tandem Repeats or VNTR's contained fairly large repeat units with allele sizes being thousands of base pairs long.

The PCR Breakthrough

In 1993 Dr. Kary Mullis received a Nobel Prize for his work during the 1980's that resulted in the invention of with The Polymerase Chain Reaction (PCR), which mimicked the cell's ability to replicate DNA, which enable scientists to take small samples of DNA and essentially copy it a million fold. All the PCR steps have similar basic steps as in RFLP, extraction, amplification and detection. Additional research identified much smaller VNTR's, which were only a few base pairs long, which coupled with PCR was the advent of STR Technology.

PCR Amplification

PCR amplification allows production of many copies of the region of DNA interest. PCR works like a "molecular xerox machine." Millions of copies of a particular sequence of DNA can be made in about 3 hours in a thermal cycler. This is great for Forensic DNA where there is usually very little DNA to start with. Initially, DNA samples that were small or degraded were beyond the reach of DNA-typing techniques. Saliva found on the back of licked postage stamp or an envelope can provide enough genetic material to conduct a sophisticated DNA test. In a Cold Case DNA ruse, the police sent the suspect a letter from a mock law firm with an invitation to join in a bogus class-action suit. The suspect replied to the letter, providing a DNA sample by licking and mailing the enclosed envelope. The DNA found in the saliva on the envelope matched a sample taken from the victim's body. This couldn't have been done without the benefit of STR/PCR technology.

Mitochondrial DNA (mtDNA)

The Mitochondrial DNA genome has been completely sequenced and is 16,569 base pairs in length. Mitochondria contain their own DNA. Every cell in the human body contains hundreds of mitochondria, which are the power plants of the cells. There are many more copies of mtDNA than nuclear DNA present in the cell. The advantage of mtDNA typing over nuclear DNA is the added sensitivity in cases where nuclear mtDNA allows for the examination of bone fragments, hair without root, teeth & other biological evidence that may be limited.

The amount of mtDNA isolated from such specimens may be very small so DNA extraction is followed by PCR amplification. This allows production of many copies of the region of mtDNA of interest. After the amplification is complete the is mtDNA sequenced using conventional sequencing methods.

Short Tandem Repeats (STR)

The STR class of polymorphisms has become the backbone of modern forensic testing. Short Tandem Repeats (STR) loci are polymorphic genetic markers that are well distributed throughout the human genome. The advantage of STR technology is that the small size of STR loci improves the chance of obtaining a result. The interpretation of STR types is simplified through the use of computers, which analyze the sample. DNA profiling uses high throughput instrumentation equipped with detectors.

Fluorescent detectors identify thirteen different loci that can be analyzed simultaneously. These 13 loci include: D3S1358, vWA,

FGA, D8S1179, D21S11, D18S51, D5S818, D13S317, D7S820, D16S539, Th01, TPOX, and CSF1PO. These 13 loci Form the basis of the National DNA network, CODIS (Combined DNA Index System), which is administrated by the FBI and is installed in forensic laboratories nationwide.

DNA technology is constantly evolving though new applications and innovations. Forensic scientists are combining advances in miniaturization and microchip technologies with well-established techniques of forensic DNA analysis. The fusion of these technologies could revolutionize DNA typing. New methods of DNA technologies have provided for the analysis of previously unsuitable case work samples that are able to be tested.

The Amelogenin Gene

The Amelogenin gene (AAAGTG), is used to identify the sex of an origin of a sample. The standard CODIS set includes amelogenin markers for X and Y chromosomes. If both X and Y are in a sample it is a male. If there is no Y marker, it's a female.

Y-STR DNA

Y-STR-DNA allows for the typing of a portion of the Y chromosome. It detects male DNA only. Small amounts of male DNA can be typed successfully. Mixtures of male DNA (Multiple rapists or rapist with consensual partner) can be resolved and mixtures of male and female DNA may be resolved.

Single Nucleotide Polymorphisms (SNP's)

In order to make new cells, an existing cell divides itself in two. But first it copies its DNA so the new cells will each have a complete set of genetic instructions. Cells sometimes make mistakes during the copying process - kind of like typos or mutations. These typos lead to variations in the DNA sequence at particular locations called Single Nucleotide Polymorphisms or SNP's.

The SNP's are being used to perform DNA profiling of the Y chromosome. In addition to the 13 CODIS loci a number of laboratories have developed multiplexes of SNP's so that male DNA can be individually typed.

Touch DNA

The Touch DNA method was named for the fact that it analyzes skin cells left behind when assailants touch victims, weapons or something else at a crime scene. Humans shed tens of thousands of skin cells each day. These cells are transferred to every surface our skin contacts, i.e. gun grips, eating utensils, steering wheels, etc. If a perpetrator deposits a sufficient number of skin cells on an item at the scene there may be Touch DNA. Touch DNA is not Low Copy Number (LCN) DNA. LCN DNA profiling allows a very small amount of DNA to be analyzed, from as little as 5 to 20 cells. The small amount of starting DNA in LCN samples requires many more cycles of amplification.

Collection of Buccal Cells for DNA Analysis

DNA technology has become so advanced through the extreme sensitivity of techniques like PCR that DNA from epithelial cells that are present in saliva can be swabbed from the surfaces of the oral cavity of suspects. This has become the method of choice in screening a number of suspects in an investigation because the samples can easily and quickly analyzed. During the BTK investigation Kansas detectives collected over 4000 buccal cell samples to compare with their suspect evidence DNA.

SNP Based Ancestry Markers - Biographical Ancestry

Ancestry informative markers are being used by certain DNA firms to help people trace their genographic roots. This technology also has the potential of identifying the four main continental population groups, such as sub-Saharan African, East Asian, Indo-European, and Native American. This technology and could be utilized to allow investigators to concentrate on the specific physical characteristics of the donor of DNA evidence.

CODIS

CODIS stands for Combined DNA Index System. It is the core of the national DNA database, established and funded by the Federal Bureau of Investigation (FBI), and developed specifically to enable public forensic DNA laboratories to create searchable DNA databases of authorized DNA profiles. The CODIS Unit manages the Combined DNA Index System (CODIS) and the National DNA Index System (NDIS) and is responsible for developing, providing, and supporting the CODIS Program to federal, state, and local crime laboratories in the United States and selected international law enforcement crime laboratories to foster the exchange and comparison of forensic DNA evidence from violent crime investigations.

CODIS software enables State, local, and national law enforcement crime laboratories to compare DNA profiles electronically, thereby linking serial crimes to each other and identifying suspects by matching DNA profiles from crime scenes with profiles from convicted offenders.

CODIS uses two indexes to generate investigative leads in crimes for which biological evidence is recovered from a crime scene. The convicted offender index contains DNA profiles of individuals convicted of certain crimes ranging from certain misdemeanors to sexual assault and murder. Each State has different "qualifying offenses" for which persons convicted of them must submit a biological sample for inclusion in the DNA database. The forensic index contains DNA profiles obtained from crime scene evidence, such as semen, saliva, or blood. CODIS uses computer software to automatically search across these indexes for a potential match.

The success of CODIS is demonstrated by the thousands of matches that have linked serial cases to each other and cases that have been solved by matching crime scene evidence to known convicted offenders.

4 The future of forensic DNA analysis

The forensic community, as it always has, is facing the question in which direction the DNA Fingerprint technology will be developed. A growing number of colleagues are convinced that DNA sequencing will soon replace methods based on fragment length analysis and there are good arguments for this position. With the emergence of current Next Generation Sequencing (NGS) technologies, the body of forensically useful data can potentially be expanded and analyzed quickly and cost-efficiently. Given the enormous number of potentially informative DNA loci - which of those should be sequenced? In my opinion there are four types of polymorphisms which deserve a place on the analytic device: an array of 20–30 autosomal STRs which complies with the standard sets used in the national and international databases around the world, a highly discriminating set of Y chromosomal markers, individual and signature polymorphisms in the control and coding region of the mitochondrial genome [5], as well as ancestry and phenotype inference SNPs [6]. Indeed, a promising NGS approach with the simultaneous analysis of 10 STRs, 386 autosomal ancestry and phenotype informative SNPs, and the complete mtDNA genome has been presented recently [7]. Currently, the rather high error rates are preventing NGS technologies from being used in forensic routine [8], but it is foreseeable that the technology will be improved in terms of accuracy and reliability. Time is another essential factor in police investigations which will be considerably reduced in future applications of DNA profiling. Commercial instruments capable of producing a database-compatible DNA profile within 2 hours exist [9] and are currently under validation for law enforcement use. The hands-free 'swab in - profile out' process consists of automated extraction, amplification, separation, detection, and allele calling without human intervention. In the US the promise of on-site DNA analysis has already altered the way in which DNA could be collected in future. In a recent decision the Supreme court of the United States held that 'when officers make an arrest supported by probable cause to hold for a serious offense and bring the suspect to the station to be detained in custody, taking and analyzing a cheek swab of the arrestee's DNA is, like fingerprinting and photographing, a legitimate police booking procedure' (Maryland v. Alonzo Jay King, Jr.). In other words, DNA can be taken from any arrestee, rightly or wrongly arrested, as a part of the normal booking procedure. Twenty-eight states and the federal government now take DNA swabs after arrests with the aim of comparing profiles to the CODIS database, creating links to unsolved cases and to identify the person (Associated Press, 3 June 2013). Driven by the rapid technological progress DNA actually becomes another metric of quick identification. It remains to be seen whether rapid DNA technologies will alter the way in which DNA is collected by police in other countries. In Germany for example the DNA collection is still regulated by the code of the criminal procedure and the use of DNA profiling for identification purposes only is excluded. Because national legislations are basically so different, a worldwide system to interrogate DNA profiles from criminal justice databases seems currently a very distant project.

Schematic overview of Haloplex targeting and NGS analysis of a large number of markers simultaneously. Sequence data are shown for samples from two individuals and the D3S1358 STR marker, the rs1335873 SNP marker, and a part of the HVII region of mtDNA.

At present the forensic DNA technology directly affects the lives of millions people worldwide. The general acceptance of this technique is still high, reports on the DNA identification of victims of the 9/11 terrorist attacks [10], of natural disasters as the Hurricane Katrina [11], and of recent wars (for example, in former Yugoslavia [12]) and dictatorship (for example, in Argentina [13]) impress the public in the same way as police investigators in white suits securing DNA evidence at a broken door. CSI watchers know, and even professionals believe, that DNA will inevitably solve the case just following the motto Do Not Ask, it's DNA, stupid! But the affirmative view changes and critical questions are raised. It should not be assumed that the benefits of forensic DNA fingerprinting will necessarily override the social and ethical costs [14].

5 CONCLUSION

In this review, a brief overview of the major developments in the field of forensic DNA analysis during the past 3 years is given. New approaches continued to be explored for more effectiveness. Time continues to be a negative factor in forensic analysis. As DNA tends to degrade under ambient conditions, how to retain the integrity of DNA over an indefinite time of storage becomes a challenge. It is worth noting that although numerous scientific improvements are sure to come, the current methods are reliable and valid. Arguments continue to be made whether SNP markers or even DNA methylation will eventually surpass STR loci as the future main target of forensic DNA analysis. The great potential of automated microfluidic devices coupled with nanotechnology for high throughput DNA analysis is yet to be completely fulfilled. Many more exciting scientific and technological advances are still on the horizon, there is no doubt that the future landscape of forensic DNA analysis will look very different from what we see today. This short article leaves many of such questions unanswered. Alfred Nobel used his fortune to institute a prize for work 'in ideal direction'. What would be the ideal direction in which DNA finger-printing, one of the great discoveries in recent history should be developed?

ACKNOWLEDGMENT

The authors wish to thank the faculty of Medicine and Health Sciences, University Sultan Zainal Abidin (UniSZA) for their support and motivation.

REFERENCES

[1]]. Balding, D. J. 1999. When can a DNA profile be regarded as unique? Science and Justice (in press).
[2] Evett, I. W., and B. S. Weir. 1998. Interpreting DNA Evidence. Statistical Genetics for Forensic Scientists. Sinauer Associates, Sunderland, MA.
[3] NRC1992. DNA Technology in Forensic Science. National Academy Press, Washington, D. C.
[4] NRC1996. The Evaluation of Forensic DNA Evidence. National Academy Press, Washington, D. C.
[5] Parson W, Strobl C, Strobl C, Huber G, Zimmermann B, Gomes SM, Souto L, Fendt L, Delport R, Langit R, Wootton S, Lagacé R, Irwin J. Evaluation of next generation mtGenome sequencing using the Ion Torrent Personal Genome Machine (PGM) Forensic Sci Int Genet. 2013; 4:632–639.
[6] Budowle B, van Daal A. Forensically relevant SNP classes. Biotechniques. 2008; 4:603–608. 610.
[7] Allen M, Nilsson M, Havsjö M, Edwinsson L, Granemo J, Bjerke M. Haloplex and MiSeq NGS for simultaneous analysis of 10 STRs, 386 SNPs and the complete mtDNA genome. Melbourne; 2013. (Presentation at the 25th Congress of the International Society for Forensic Genetics). 2–7 September 2013.
[8] Bandelt HJ, Salas A. Current next generation sequencing technology may not meet forensic standards.Forensic Sci Int Genet. 2012; 4:143–145.
[9] Tan E, Turingan RS, Hogan C, Vasantgadkar S, Palombo L, Schumm JW, Selden RF. Fully integrated, and fully automated generation of short tandem repeat profiles. Investigative Genet. 2013; 4:16. Doi: 10.1186/2041-2223-4-16.
[10] Biesecker LG, Bailey-Wilson JE, Ballantyne J, Baum H, Bieber FR, Brenner C, Budowle B, Butler JM, Carmody G, Conneally PM, Duceman B, Eisenberg A, Forman L, Kidd KK, Leclair B, Niezgoda S, Parsons TJ, Pugh E, Shaler R, Sherry ST, Sozer A, Walsh A. DNA Identifications after the 9/11 World Trade Center Attack. Science. 2005; 4:1122–1123. Doi: 10.1126/science.1116608.
[11] Dolan SM, Saraiya DS, Donkervoort S, Rogel K, Lieber C, Sozer A. The emerging role of genetics professionals in forensic kinship DNA identification after a mass fatality: lessons learned from Hurricane Katrina volunteers. Genet Med. 2009; 4:414–417.
[12] Huffine E, Crews J, Kennedy B, Bomberger K, Zinbo A. Mass identification of persons missing from the break-up of the former Yugoslavia: structure, function, and role of the International Commission on Missing Persons. Croat Med J. 2001; 4:271–275.
[13] Corach D, Sala A, Penacino G, Iannucci N, Bernardi P, Doretti M, Fondebrider L, Ginarte A, Inchaurregui A, Somigliana C, Turner S, Hagelberg E. Additional approaches to DNA typing of skeletal remains: the search for "missing" persons killed during the last dictatorship in Argentina. Electrophoresis.1997; 4:1608–1612. doi: 10.1002/elps.1150180921.
[14] Levitt M. Forensic databases: benefits and ethical and social costs. Br Med Bull. 2007; 4:235–248. doi: 10.1093/bmb/ldm026.