

## OVERVIEW AND HISTORY OF DNA TYPING

In the darkness of the early morning hours of 26 August 1999, a young University of Virginia student awoke to find a gun pointed at her head. The assailant forced her and a male friend spending the night to roll over on their stomachs. Terrorized, they obeyed their attacker. After robbing the man of some cash, the intruder put a pillow over the man's head and raped the female student. The female was blindfolded with her own shirt and led around the house while the intruder searched for other items to steal.

Throughout the entire ordeal, the intruder kept his gun to the back of the male student's head, daring him to look at him and telling him if he tried he would blow his head off. The assailant forced the young woman to take a shower in hopes that any evidence of the crime would be washed away. After helping himself to a can of beer, the attacker left before dawn taking with him the cash, the confidence, and the sense of safety of his victims. Even though the assailant had tried to be careful and clean up after the sexual assault, he had left behind enough of his personal body fluids to link him to this violent crime.

The police investigating the crime collected some saliva from the beer can. In addition, evidence technicians found some small traces of semen on the bed sheets that could not be seen with the naked eye. These samples were submitted to the Virginia Department of Forensic Sciences in Richmond along with control samples from other occupants of the residence where the crime occurred. The DNA profiles from the beer can and the bed sheets matched each other, but no suspect had been developed yet. Because of intense darkness and then the blindfold, the only description police had from the victims was that the suspect was black, medium height, and felt heavy set.

A suspect list was developed by the Charlottesville Police Department that contained over 40 individuals, some from the sex offender registry and some with extensive criminal histories who were stopped late at night in the area of the home invasion. Unfortunately, no further leads were available, leaving the victims as well as other University of Virginia students and their parents suspicious and fearful. The police were at the end of their rope and considered asking many of the people on the suspect list to voluntarily donate blood

samples for purposes of a DNA comparison. The top suspects were systematically eliminated by DNA evidence, leaving the police frustrated.

Then, on 5 October, 6 long weeks after the crime had been committed, the lead detective on the case, Lieutenant J.E. 'Chip' Harding of the Charlottesville Police Department, received a call that he describes as being 'one of the most exciting phone calls in my 22 years of law enforcement.' A match had been obtained from the crime scene samples to a convicted offender sample submitted to the Virginia DNA Database several years before. The DNA sample for Montaret D. Davis of Norfolk, Virginia, was among 8000 samples added to the Virginia DNA Database at the beginning of October 1999. (Since 1989, a Virginia state law has required all felons and juveniles 14 and older convicted of serious crimes to provide blood samples for DNA testing.)

A quick check for the whereabouts of Mr Davis found him in the Albemarle-Charlottesville Regional Jail. Ironically, because of a parole violation, he had been court ordered weeks before to report to jail on what turned out to be the same day as the rape. Amazingly enough he had turned himself in at 6 p.m. just 14 hours after committing the sexual assault! Unless he would have bragged about his crime, it is doubtful that Mr Davis would ever have made it on the suspect list without the power of DNA testing and an expanding DNA database. At his jury trial in April 2000, Mr Davis was found guilty of rape, forcible sodomy, and abduction among other charges and sentenced to a 90-year prison term.

DNA typing, since it was introduced in the mid-1980s, has revolutionized forensic science and the ability of law enforcement to match perpetrators with crime scenes. Thousands of cases have been closed and innocent suspects freed with guilty ones punished because of the power of a silent biological witness at the crime scene. This book will explore the science behind DNA typing and the biology, technology, and genetics that make DNA typing the most useful investigative tool to law enforcement since the development of fingerprinting 100 years ago.

## HISTORY OF FORENSIC DNA ANALYSIS

'DNA fingerprinting,' or DNA typing (profiling) as it is now known, was first described in 1985 by an English geneticist named Alec Jeffreys. Dr Jeffreys found that certain regions of DNA contained DNA sequences that were repeated over and over again next to each other. He also discovered that the number of repeated sections present in a sample could differ from individual to individual. By developing a technique to examine the length variation of these DNA repeat sequences, Dr Jeffreys created the ability to perform human identity tests.

These DNA repeat regions became known as VNTRs, which stands for variable number of tandem repeats. The technique used by Dr Jeffreys to examine the VNTRs was called restriction fragment length polymorphism (RFLP) because it involved the use of a restriction enzyme to cut the regions of DNA surrounding the VNTRs. This RFLP method was first used to help in an English immigration case and shortly thereafter to solve a double homicide case. Since that time, human identity testing using DNA typing methods has been widespread. The past 15 years have seen tremendous growth in the use of DNA evidence in crime scene investigations as well as paternity testing. Today over 100 public forensic laboratories and several dozen private paternity testing laboratories conduct hundreds of thousands of DNA tests annually in the United States. In addition, most countries in Europe and Asia have forensic DNA programs. The number of laboratories around the world conducting DNA testing will continue to grow as the technique gains in popularity within the law enforcement community.

## COMPARISON OF DNA TYPING METHODS

Technologies used for performing forensic DNA analysis differ in their ability to differentiate two individuals and in the speed with which results can be obtained. The speed of analysis has dramatically improved for forensic DNA analysis. DNA testing that previously took 6 or 8 weeks can now be performed in a few hours.

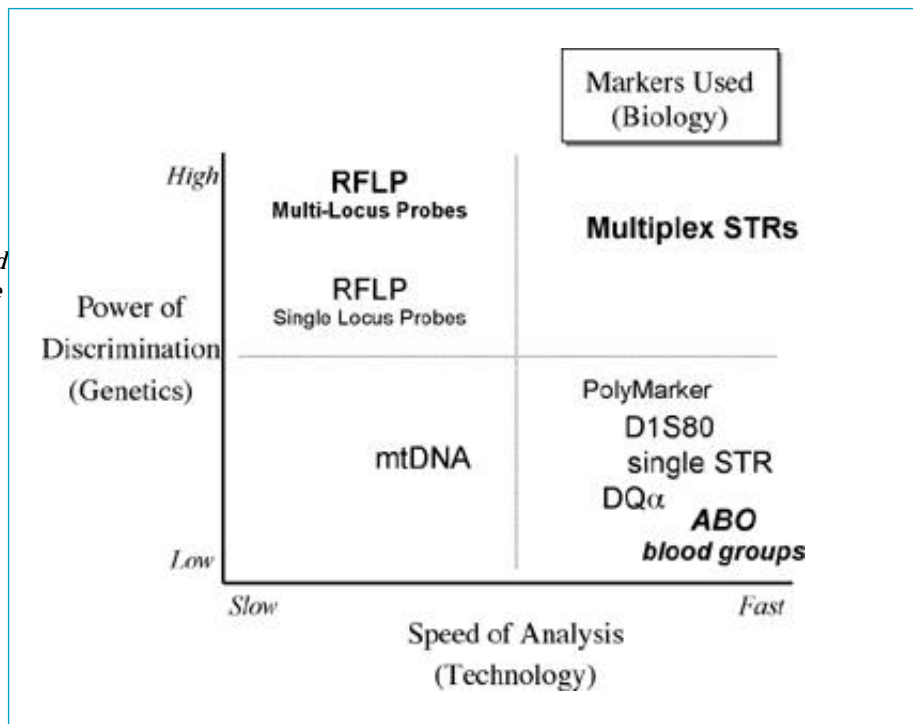
The human identity testing community has used a variety of techniques including single-locus probe and multi-locus probe RFLP methods and more recently PCR (polymerase chain reaction)-based assays. Numerous advances have been made in the last 15 years in terms of sample processing speed and sensitivity. Instead of requiring large blood stains with well-preserved DNA, minute amounts of sample, as little as a single cell in some cases, can yield a useful DNA profile.

The gambit of DNA typing technologies used over the past 15 years for human identity testing are compared in Figure 1.1. The various DNA markers have been divided into four quadrants based on their power of discrimination, i.e. their ability to discern the difference between individuals, and the speed at which they can be analyzed. New and improved methods have developed over the years such that tests with a high degree of discrimination can now be performed in a few hours.

An ABO blood group determination, which was the first genetic tool used for distinguishing between individuals, can be performed in a few minutes but is not very informative. There are only four possible genotype groups – A, B, AB, and O – and 40% of the population is type O. Thus, while the ABO blood

*Figure 1.1*

*Comparison of DNA typing technologies. Forensic DNA markers are arbitrarily plotted in relationship to four quadrants defined by the power of discrimination for the genetic system used and the speed at which the analysis for that marker may be performed. Note that this diagram does not reflect the usefulness of these markers in terms of forensic cases.*



groups are useful for excluding an individual from being the source of a crime scene sample, the test is not very useful when an inclusion has been made, especially if the sample is type O.

On the other extreme, multi-locus RFLP probes are highly variable between individuals but require a great deal of labor, time, and expertise to produce a DNA profile. Analysis of multi-locus probes (MLP) cannot be easily automated, a fact that makes them undesirable as the demand for processing large numbers of DNA samples has increased. Deciphering sample mixtures, which are common in forensic cases, is also a challenge with MLP RFLP methods, which is the primary reason that laboratories went to single-locus RFLP probes used in serial fashion.

The best solution including a high power of discrimination and a rapid analysis speed has been achieved with short tandem repeat (STR) DNA markers, shown in the upper right quadrant of Figure 1.1. Also because STRs by definition are short, they can be analyzed three or more at a time. Multiple STRs can be examined in the same DNA test, or 'multiplexed.' Multiplex STRs are valuable because they can produce highly discriminating results (Chapter 5) and can successfully measure sample mixtures and biological materials containing degraded DNA molecules (Chapter 7). In addition, the detection of multiplex STRs can be automated, which is an important benefit as demand for DNA testing increases.

It should be noted, however, that Figure 1.1 does not fully reflect the usefulness of these markers in terms of forensic cases. Mitochondrial DNA (mtDNA), which is shown in the quadrant with the lowest power of discrimination and longest sample processing time, can be very helpful in forensic cases involving severely degraded DNA samples or when associating maternally related individuals (Chapter 8). In many situations, multiple technologies may be used to help resolve an important case (Chapter 17).

Over the past 15 years, there has been a gradual evolution in adoption of the various DNA typing technologies shown in Figure 1.1. When early methods for DNA analysis are superseded by new technologies, there is usually some overlap as forensic laboratories implement the new technology. Validation of the new methods is crucial to maintaining high-quality results (Chapter 14). Table 1.1 lists some of the major historical events in forensic DNA typing. The implementation of new methods by the FBI Laboratory has been listed in this historical timeline because the DNA casework protocols used by the FBI create an important trend within the United States and around the world.

## STEPS IN DNA SAMPLE PROCESSING

This book contains a review of the steps involved in processing forensic DNA samples with short tandem repeat markers. STRs are a smaller version of the VNTR sequences first described by Dr Jeffreys. Samples obtained from a crime scene or a paternity investigation are subjected to defined processes involving biology, technology, and genetics (Figure 1.2).

## BIOLOGY

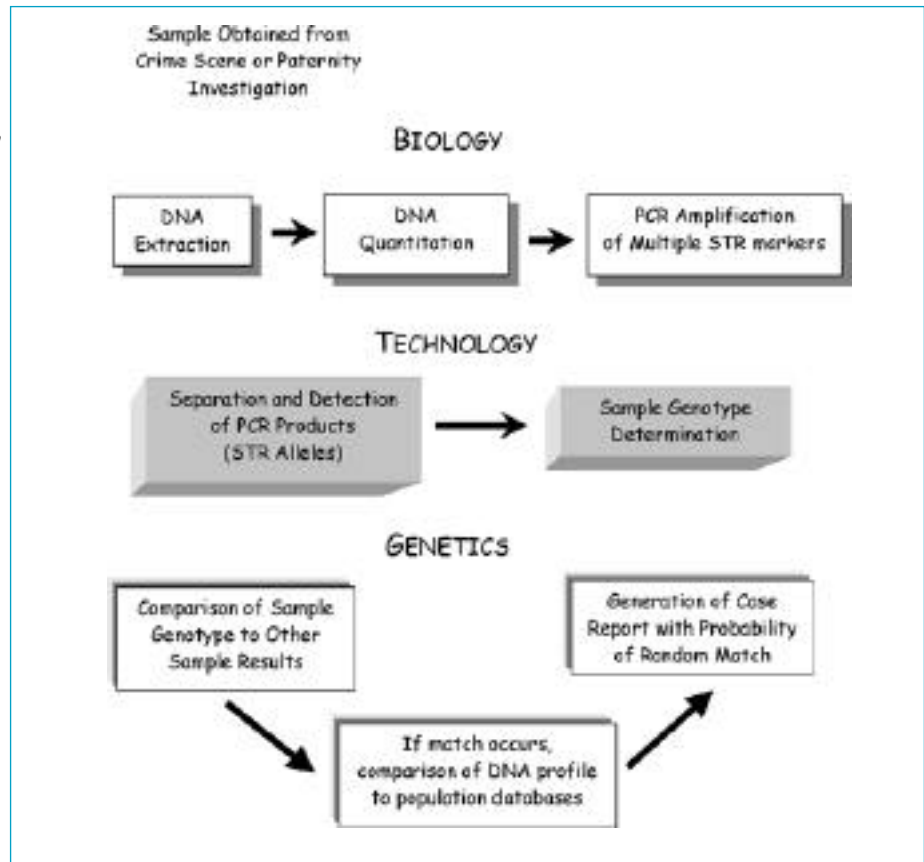
The DNA is first extracted from its biological source material and then measured to evaluate the quantity of DNA recovered (Chapter 3). After isolating the DNA from its cells, specific regions are copied with a technique known as the polymerase chain reaction (Chapter 4). PCR produces millions of copies for each starting DNA molecule and thus permits very minute amounts of DNA to be examined. Multiple STR regions can be examined simultaneously to increase the informativeness of the DNA test (Chapter 5).

## TECHNOLOGY

The resulting PCR products are then separated and detected in order to characterize the STR region being examined. The separation methods used today include slab gel and capillary electrophoresis (CE) (Chapter 9). Fluorescence detection methods have greatly aided the sensitivity and ease of measuring

Figure 1.2

Overview of biology technology, and genetic components of DNA typing using short tandem repeat (STR) markers.



PCR-amplified STR alleles (Chapter 10). The two primary instrument platforms used in the United States for fluorescence detection of STR alleles are the ABI 310 Genetic Analyzer (Chapter 11) and the Hitachi FMBIO II (Chapter 12). After detecting the STR alleles, the number of repeats in a DNA sequence is determined, a process known as sample genotyping (Chapter 13).

### GENETICS

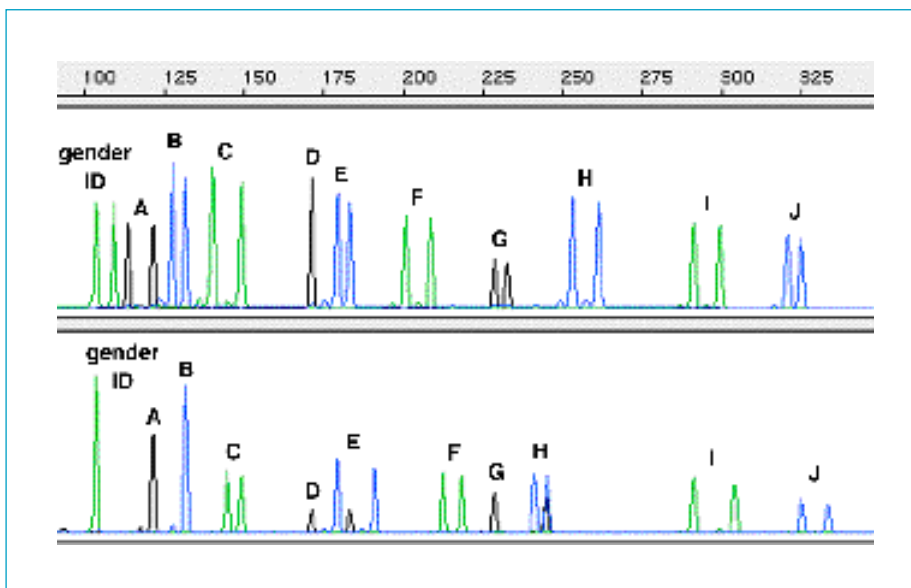
The resulting DNA profile for a sample, which is a combination of individual STR genotypes, is compared to other samples. In the case of a forensic investigation, these other samples would include known reference samples such as the victim or a suspect that are compared to the crime scene evidence. With paternity investigations, the child's genotype would be compared to his or her mother's and the alleged father(s) under investigation. If there is no match between the questioned sample and the known sample, then the samples may be considered to have originated from different sources. The term for failure to match between two DNA profiles is 'exclusion.'

If a match or 'inclusion' results, then a comparison of the DNA profile is made to a population database, which is a collection of DNA profiles obtained from unrelated individuals of a particular ethnic group. For example, due to genetic variation between the groups, African-Americans and Caucasians have different population databases for comparison purposes. Finally a case report or paternity test result is generated. This report typically includes the random match probability for the match in question. This random match probability is the chance that a randomly selected individual from a population will have an identical STR profile or combination of genotypes at the DNA markers tested.

### STR MULTIPLEX EXAMPLE

An example of DNA profiles obtained from two different individuals using STR markers is shown in Figure 1.3. In a single amplification reaction, unique sites on ten different chromosomes were probed with this DNA test to provide a random match probability of approximately 1 in 3 trillion. Note that every single site tested produces a different result between these two DNA samples. For example, marker A has two peaks in the top panel and only one peak in the bottom panel. Likewise, marker J produces two peaks in both samples but they result in different patterns due to different sizes at the site measured in the two DNA samples. These results can be reliably obtained in as little as a few hours from a very small drop of blood or blood stain.

Each STR allele is distinguished from the others in the amplification reaction by separating it based on its length and color. The color results from a fluores-



*Figure 1.3*  
Comparison of the DNA profiles for two individuals obtained with multiple short tandem repeat markers. STR length variation at unique sites on ten different chromosomes are probed with this DNA test to provide a random match probability of approximately 1 in 3 trillion. A gender identification test also indicates that the top sample is from a male while the bottom sample is from a female individual. These results were obtained from a spot of blood the size of a pin head in less than 5 hours. The DNA size range in base pairs is shown across the top of the plot. Results from each DNA marker are indicated by the letters A–J.

cent dye that is attached during the amplification reaction. In this example, DNA markers B, E, H, and J are labeled with a blue colored dye, markers A, D, and G are labeled with a yellow dye, and markers C, F, I, and the gender identification (ID) are labeled in green. The gender ID results in two peaks for a male sample and a single peak for a female sample.

## COMPARISONS TO COMPUTER TECHNOLOGY

In order to obtain a better feel for how rapidly forensic DNA analysis methods have progressed in the last decade or so, a comparison to computer technology may be helpful. The use of computers at home and in the workplace has increased dramatically since personal computers became available in the mid-1980s. These computers become faster and more powerful every year. It is almost inconceivable that the internet, which has such a large impact on our daily lives, was just an idea a few years ago.

When multi-locus RFLP probes were first reported in 1985, the average computer operating speed was less than 25 MHz. Fifteen years later in the year 2000, computers with computing speeds of 1000 MHz (1 GHz) are now common. Computer processing speeds and capabilities have increased rapidly every year. Likewise, the ability of laboratories to perform DNA typing methods has improved dramatically due to the rapid progress in the areas of biology, technology, and understanding of genetic theories. In addition, the power of discrimination for DNA tests has steadily increased in the late 1990s (see Table 5.3).

Some interesting parallels can be drawn between Microsoft Corporation, the company that has led the computer technology revolution, and the timing for advancements in the field of forensic DNA typing (Table 1.1). In 1985, the year that Alec Jeffreys first published his work with multi-locus RFLP probes, Microsoft shipped its first version of Windows software to serve as a computer operating system. In 1986, as DNA testing began to 'go public' in the United States with Cellmark and Lifecodes performing multi-locus RFLP, Microsoft went public with a successful initial public offering.

In the late 1980s, single-locus RFLP probes began to be used by the FBI Laboratory in DNA casework. Owing to issues over the use of statistics for population genetics and the quality of results obtained in forensic laboratories, RFLP methods were questioned by the legal community in 1989 and the early 1990s. At this same time, Microsoft had quality problems of their own with the Windows 3.0 operating system. However, they 'turned the corner' with their product release of Windows 3.1 in 1991. In the same year, improved methods for DNA typing were introduced, namely fluorescent STR markers and Chelex extraction.

The popularity of Microsoft products improved in 1995 with the release of

*Table 1.1 (facing)*

*Major historical events in forensic DNA typing shown by year. The events relating to forensic DNA (first column) are described in context with parallel developments in biotechnology (second column) and key events relating to Microsoft Corporation, which have impacted the computer age (final column).*

Year	Forensic DNA Science and Application	Parallel Developments in Biotechnology	Microsoft Corporation Chronology
1985	Alec Jeffreys develops multi-locus RFLP probes	PCR process first described	First version of Windows shipped
1986	DNA testing goes public with Cellmark and Lifecodes in United States	Automated DNA sequencing with 4-colors first described	Microsoft goes public
1988	FBI begins DNA casework with single-locus RFLP probes		
1989	TWGDAM established; <i>NY v. Castro</i> case raises issues over quality assurance of laboratories	DNA detection by gel silver-staining, slot blot, and reverse dot blots first described	
1990	Population statistics used with RFLP methods are questioned; PCR methods start with DQA1	Human Genome Project begins with goal to map all human genes	Windows 3.0 released (quality problems); exceeds \$1 billion in sales
1991	Fluorescent STR markers first described; Chelex extraction		Windows 3.1 released
1992	NRC I Report; FBI starts casework with PCR-DQA1	Capillary arrays first described	
1993	First STR kit available; sex-typing (amelogenin) developed	First STR results with CE	
1994	Congress authorizes money for upgrading state forensic labs; 'DNA wars' declared over; FBI starts casework with PCR-PM	Hitachi FMBIO and Molecular Dynamics gel scanners; first DNA results on microchip CE	
1995	O.J. Simpson saga makes public more aware of DNA; DNA Advisory Board setup; UK DNA Database established; FBI starts using D1S80/amelogenin	ABI 310 Genetic Analyzer and TaqGold DNA polymerase introduced	Windows 95 released
1996	NRC II Report; FBI starts mtDNA testing; first multiplex STR kits become available	STR results with MALDI-TOF and GeneChip mtDNA results demonstrated	
1997	13 core STR loci defined; Y-chromosome STRs described		Internet Explorer begins to overtake Netscape
1998	FBI launches national Combined DNA Index System; Thomas Jefferson and Bill Clinton implicated with DNA	2000 SNP hybridization chip described	Windows 98 released; anti-trust trial with US Justice Department begins
1999	Multiplex STR kits are validated in numerous labs; FBI stops testing DQA1/PM/D1S80	ABI 3700 96-capillary array for high-throughput DNA analysis; chromosome 22 fully sequenced	
2000	FBI and other labs stop running RFLP cases and convert to multiplex STRs	First copy of human genome completed	Bill Gates steps down as Microsoft CEO

Windows 95. During this same year, forensic DNA methods gained public exposure and popularity due to the O.J. Simpson trial. The United Kingdom also launched a National DNA Database that has revolutionized the use of DNA as an investigative tool. The United States launched their national Combined DNA Index System (CODIS) in 1998, concurrent with the release of Windows 98.

Finally, to aid sample throughput and processing speed, the FBI Laboratory and many other forensic labs have stopped running RFLP cases as of the year 2000. On 13 January 2000, Bill Gates stepped down as the CEO of Microsoft in order to help his company move into new directions.

We recognize that, owing to the rapid advances in the field of forensic DNA typing, some aspects of this book may be out of date by the time it is published, much like a computer is no longer the latest model by the time it is purchased. However, a reader should be able to gain a fundamental understanding of forensic DNA typing from the following pages. While we cannot predict the future with certainty, short tandem repeat DNA markers have had and will continue to have an important role to play in forensic DNA typing due to their use in DNA databases.

The match on Mr Davis described at the beginning of this chapter was made with eight STR markers. These eight STRs are a subset of 13 STR markers described in detail throughout this book that will most likely be used in DNA databases around the world for many years to come. Perhaps with the odds of getting caught becoming greater than ever before, violent criminals like Mr Davis will think twice before carrying out such heinous actions.