DNA testing is to justice what the telescope is for the stars; not a lesson in biochemistry, not a display of the wonders of magnifying glass, but a way to see things as they really are.

(Barry Scheck and Peter Neufeld, *Actual Innocence*)

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 the hope 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, six 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 investigatory tool to law enforcement since the development of fingerprinting over 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 (see D.N.A. Box 1.1). 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 150 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

The first use of DNA testing in a forensic setting came in 1986. Two young girls, Lynda Mann and Dawn Ashworth, were sexually assaulted and then left brutally murdered in 1983 and 1986. Both murders occurred near the village of Narborough in Leicestershire, England with similar features leading the police to suspect that the same man had committed the crimes. A local man confessed to killing one of the girls and his blood was compared to semen recovered from the crime scenes. The man did not match evidence from either crime! Thus, the first use of DNA was to demonstrate innocence of someone who might otherwise have been convicted.

A mass screen to collect blood for DNA testing from all adult men in three local villages was conducted in a thorough search for the killer. Over 4000 men were tested without a match. About a year later a woman at a bar overheard someone bragging about how he had given a blood sample for a friend named Colin Pitchfork. The police interviewed Mr. Pitchfork, collected a blood sample from him, and found that his DNA profile matched semen from both murder scenes. He was subsequently convicted and sentenced to life in prison.

The story behind the first application of forensic DNA typing or genetic fingerprinting, as it was then called, has been well told in Joseph Wambaugh’s *The Blooding*. The DNA typing methods used were Alec Jeffrey’s multi-locus RFLP probes, which he first described in 1985. Since it was first used almost 20 years ago, DNA testing has progressed to become a sensitive and effective tool to aid in bringing the guilty to justice and in exonerating the innocent.

Source:
http://www.forensic.gov.uk
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, tiny amounts of sample, as little as a single cell in some cases can yield a useful DNA profile.

The gamut of DNA typing technologies used over the past 15 years for human identity testing is 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 may be performed. Note that this diagram does not reflect the usefulness of these markers in terms of forensic cases.

![Figure 1.1](image-url)

**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.
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 groups that are typed – A, B, AB, and O – and 40% of the population is type O. Thus, while the ABO blood 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 though 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 10). In many situations, multiple technologies may be used to help resolve an important case or identify victims of mass disasters, such as those from the World Trade Center collapse (Chapter 24).

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 16). 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 STR markers. STRs are a smaller version of the VNTR sequences first described by Dr. Jeffreys. Samples obtained from crime scenes or paternity investigations are subjected to defined processes involving biology, technology, and genetics (Figure 1.2).

**BIOLOGY**

Following collection of biological material from a crime scene or paternity investigation, 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, or PCR (Chapter 4). PCR produces millions of copies for each DNA segment of interest 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 12). Fluorescence detection methods have greatly aided the sensitivity and ease of measuring PCR-amplified STR alleles (Chapter 13). The primary instrument platform used in the United States for fluorescence detection of STR alleles is currently the ABI Prism 310 Genetic Analyzer (Chapter 14). After detecting the STR alleles, the number of repeats in a DNA sequence is determined, a process known as sample genotyping (Chapter 15).

The specific methods used for DNA typing are validated by individual laboratories to ensure that reliable results are obtained (Chapter 16) and before new technologies (see Chapter 17) are implemented. DNA databases, such as the one described earlier in this chapter to match Montaret Davis to his crime scene, are valuable tools and will continue to play an important role in law enforcement efforts (Chapter 18).

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 suspects that are compared to the crime scene evidence. With paternity investigations, a child’s genotype would be compared to his or her mother’s and the alleged father(s) under investigation (Chapter 23). If there is not a match between the questioned sample and the known sample, then the samples may be considered to have originated from different sources (see D.N.A. Box 1.2). The term used 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 (Chapter 20). 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 (see example in D.N.A. Box 1.3). 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 (Chapter 21).

**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
Can a simple DNA test have the power to impact world events? In 1998, independent counsel Kenneth Starr was investigating allegations that U.S. President William Jefferson Clinton had a sexual relationship with a young White House intern, Monica Lewinsky. President Clinton had publicly denied the allegations quite emphatically and at that time there was no concrete evidence to the contrary.

During the course of the investigation, a dark blue dress belonging to Monica Lewinsky was brought to the FBI Laboratory for examination. Semen was identified on evidence item Q3243, as the dress was cataloged. The unknown semen stain was quickly examined with seven RFLP single locus probes. Late on the evening of 3 August 1998, a reference blood sample was drawn from President Clinton for comparison purposes (Woodward 1999).

As in the O.J. Simpson case (see D.N.A. Box 3.2), conventional RFLP markers were used to match the sample of President Clinton’s blood to the semen stain on Monica Lewinsky’s dress. At the time these samples were run in the FBI Laboratory (early August 1998), STR typing methods were being validated but were not yet in routine use within the FBI’s DNA Analysis Unit. High molecular weight DNA from the semen stain (FBI specimen Q3243-1) and President Clinton’s blood (FBI specimen K39) was digested with the restriction enzyme HaeIII. A seven-probe match was obtained at all seven RFLP loci examined.

This match was reported in the following manner: ‘Based on the results of these seven genetic loci, specimen K39 (CLINTON) is the source of the DNA obtained from specimen Q3243-1, to a reasonable degree of scientific certainty.’ The random match probability was calculated to be on the order of 1 in 7.8 trillion when compared to a Caucasian population database.

When faced with this indisputable DNA evidence, President Clinton found himself in a tight spot. Earlier statements that he had not had ‘sexual relations’ with Miss Lewinsky were now in doubt. The DNA results along with other evidence and testimony resulted in the impeachment of President Clinton on 19 December 1998 – only the second President in U.S. history to be impeached. This physical evidence played an important role in demonstrating that a sexual relationship had existed between Miss Lewinsky and President William Jefferson Clinton. Although during the Senate impeachment trial, it was determined that his deeds were not serious enough for him to be removed from office, President Clinton’s career will always be tainted by the semen stain on the now famous blue dress.

Sources:
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 bloodstain.

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 fluorescent 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 ID are labeled in green. The gender ID results in two peaks for a male sample (X,Y) and a single peak for a female sample (X,X). Chapter 5 will describe the identity of the DNA markers represented in Figure 1.3.

**COMPARISONS TO COMPUTER TECHNOLOGY**

In order to get a better feel for how rapidly forensic DNA analysis methods have progressed in the last two decades, 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 get 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. Almost 20 years later in the year 2004, computing speeds of 2500 MHz (2.5 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 along a similar timeline due to 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, Table 20.8).

Some interesting parallels can be drawn between the 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

<table>
<thead>
<tr>
<th>Year</th>
<th>Forensic DNA Science &amp; Application</th>
<th>Parallel Developments in Biotechnology</th>
<th>Microsoft Corporation Chronology</th>
</tr>
</thead>
<tbody>
<tr>
<td>1985</td>
<td>Alec Jeffreys develops multi-locus RFLP probes</td>
<td>PCR process first described</td>
<td>First version of Windows shipped</td>
</tr>
<tr>
<td>1986</td>
<td>DNA testing goes public with Cellmark and Lifecodes in United States</td>
<td>Automated DNA sequencing with 4-colors first described</td>
<td>Microsoft goes public</td>
</tr>
<tr>
<td>1988</td>
<td>FBI begins DNA casework with single locus RFLP probes</td>
<td>DNA detection by gel silver-staining, slot blot, and reverse dot blots first described</td>
<td></td>
</tr>
<tr>
<td>1989</td>
<td>TWGDAM established; NY v. Castro case raises issues over quality assurance of laboratories</td>
<td>DNA detection by gel silver-staining, slot blot, and reverse dot blots first described</td>
<td></td>
</tr>
<tr>
<td>1990</td>
<td>Population statistics used with RFLP methods are questioned; PCR methods start with DQA1</td>
<td>Human Genome Project begins with goal to map all human genes</td>
<td>Windows 3.0 released (quality problems); exceeds $1 billion in sales</td>
</tr>
<tr>
<td>1991</td>
<td>Fluorescent STR markers first described; Chelex extraction</td>
<td></td>
<td>Windows 3.1 released</td>
</tr>
<tr>
<td>1992</td>
<td>NRC I Report; FBI starts casework with PCR-DQA1</td>
<td>Capillary arrays first described</td>
<td></td>
</tr>
<tr>
<td>1993</td>
<td>First STR kit available; sex-typing (amelogenin) developed</td>
<td>First STR results with CE</td>
<td></td>
</tr>
<tr>
<td>1994</td>
<td>Congress authorizes money for upgrading state forensic labs; ‘DNA wars’ declared over; FBI starts casework with PCR-PM</td>
<td>Hitachi FMBIO and Molecular Dynamics gel scanners; first DNA results on microchip CE</td>
<td></td>
</tr>
</tbody>
</table>
Table 1.1 (Continued)

<table>
<thead>
<tr>
<th>Year</th>
<th>Forensic DNA Science &amp; Application</th>
<th>Parallel Developments in Biotechnology</th>
<th>Microsoft Corporation Chronology</th>
</tr>
</thead>
<tbody>
<tr>
<td>1995</td>
<td>O.J. Simpson saga makes public more aware of DNA; DNA Advisory Board setup; UK DNA Database established; FBI starts using D1S80/amelogenin</td>
<td>ABI 310 Genetic Analyzer and TaqGold DNA polymerase introduced</td>
<td>Windows 95 released</td>
</tr>
<tr>
<td>1996</td>
<td>NRC II Report; FBI starts mtDNA testing; first multiplex STR kits become available</td>
<td>STR results with MALDI-TOF and GeneChip mtDNA results demonstrated</td>
<td></td>
</tr>
<tr>
<td>1997</td>
<td>13 core STR loci defined; Y-chromosome STRs described</td>
<td></td>
<td>Internet Explorer begins overtaking Netscape</td>
</tr>
<tr>
<td>1998</td>
<td>FBI launches national Combined DNA Index System; Thomas Jefferson and Bill Clinton implicated with DNA</td>
<td>2000 SNP hybridization chip described</td>
<td>Windows 98 released; anti-trust trial with U.S. Justice Department begins</td>
</tr>
<tr>
<td>1999</td>
<td>Multiplex STR kits are validated in numerous labs; FBI stops testing DQA1/PM/D1S80</td>
<td>ABI 3700 96-capillary array for high-throughput DNA analysis; chromosome 22 fully sequenced</td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>FBI and other labs stop running RFLP cases and convert to multiplex STRs; PowerPlex 16 kit enables first single amplification of CODIS STRs</td>
<td>First copy of human genome completed</td>
<td>Bill Gates steps down as Microsoft CEO; Windows 2000 released</td>
</tr>
<tr>
<td>2001</td>
<td>Identifiler STR kit released with 5-dye chemistry; first Y-STR kit becomes available</td>
<td>ABI 3100 Genetic Analyzer introduced</td>
<td>Windows XP released</td>
</tr>
<tr>
<td>2002</td>
<td>FBI mtDNA population database released; Y-STR 20plex published</td>
<td></td>
<td>Windows XP Tablet PC Edition released</td>
</tr>
<tr>
<td>2003</td>
<td>U.S. DNA database (NDIS) exceeds 1 million convicted offender profiles; the U.K. National DNA Database passes the 2 million sample mark</td>
<td>Human Genome Project completed with the ‘final’ sequence coinciding with 50th anniversary of Watson–Crick DNA discovery</td>
<td>Windows Server 2003 released; 64-Bit Operating Systems expand capabilities of software</td>
</tr>
</tbody>
</table>

Butler-01.qxd  09/10/2004  12:20 PM  Page 12

12 FORENSIC DNA TYPING
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 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.

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.

The development and release of Windows 2000 and Windows XP at the beginning of the 21st century continue to improve the capabilities of multi-tasking computer software. In like manner, the development and release of new DNA testing kits capable of single amplification reactions for examining 16 regions of the human genome furthers the capability of multiplexing DNA information (see Chapter 5).

We recognize that due 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 odds of getting caught becoming greater than ever before, violent criminals like Mr. Davis will think twice before carrying out such heinous actions.