

Spring 2013

12 Angry Men: An Analysis of How Law, Science, and Individual Persistence Saved DNA Profiling

Lindsay Mullineaux

University of Colorado Boulder

Follow this and additional works at: https://scholar.colorado.edu/honr_theses

Recommended Citation

Mullineaux, Lindsay, "12 Angry Men: An Analysis of How Law, Science, and Individual Persistence Saved DNA Profiling" (2013).
Undergraduate Honors Theses. 455.
https://scholar.colorado.edu/honr_theses/455

This Thesis is brought to you for free and open access by Honors Program at CU Scholar. It has been accepted for inclusion in Undergraduate Honors Theses by an authorized administrator of CU Scholar. For more information, please contact cuscholaradmin@colorado.edu.

The Case for DNA:
An Analysis of Early DNA Profiling in the Courtroom

By
Lindsay A. Mullineaux
Department of Anthropology, University of Colorado at Boulder

Defended April 8, 2013

Primary Thesis Advisor:
Dr. Steve Leigh, Dean of the College of Arts and Sciences

Defense Committee:
Dr. Donna M. Goldstein, Department of Anthropology
Dr. Jerome Christopher Mckee, Adjunct Faculty at the University of Colorado Law
School

Abstract

The following analysis stemmed from a broad inquiry regarding the current state of forensic science. After examining some recent problems concerning traditional forensic techniques, including problems with established fingerprint, hair follicle, and bite mark analysis, it became evident that forensic science has shifted toward a new scientific paradigm. The concurrent rise of DNA profiling technology may be partially responsible for this shift. Hence, the purpose of this analysis is to determine the various aspects of forensic DNA profiling that have allowed it to become such a successful tool in the criminal justice system. While in pursuit of this information, I uncovered a number of individual stories and cases that help trace the rise of forensic DNA analysis. Certainly some of these stories are tragic and painful, told by victims or victim's families hurt by violent crime. However, some of these stories, those told by lawyers and scientists, help chronicle the trials, tribulations, and controversies surrounding the introduction of a budding science into the courtroom. In tracing these unique stories I was able analyze the integration of a new scientific technique into the legal system through the eyes of the individuals involved, with the purpose of gaining new insights for forensic science in the future.

Acknowledgements

I would like to thank Steve Leigh for his eternal enthusiasm and kindness in working with me, even when I was at a total loss. His patience allowed me the time I needed to discover my topic and his support guided me through until the end.

I am also so appreciative of Donna Goldstein, who worked with me even with such late notice and who gave me the best advice, especially the time when I didn't even know what I was asking.

I would like to offer my special thanks to Jerome Christopher Mckee for his openness and willingness to help an undergraduate student, especially one with such a nonexistent background in law.

I am particularly grateful to Douglas Bamforth, who was always there for his class and who always knew just the right thing to say.

Last, but not least, I am ever so thankful for my two cats, who somehow always knew just when to sit on my keyboard and force me to take a break.

Table of Contents

Contents	Page
Abstract.....	2
Acknowledgements.....	3
Chapter 1: Constructing the Double Helix.....	5
1.1 Introduction.....	5
1.2 How Does Forensic DNA Analysis Work?.....	10
1.3 DNA Typing: An Early History.....	19
1.4 Evidence Admissibility in the United States.....	22
Chapter 2: Private Science, Public Defenders.....	24
2.1 Early Trials and Tribulations.....	24
2.2 People v. Andrews.....	27
2.3 George Wesley and Cameron Bailey.....	33
2.4 People v. Castro.....	37
2.5 State v. Schwartz.....	44
2.6 State v. Woodall.....	47
Chapter 3: Public Science and the ‘DNA Wars’.....	50
3.1 Change, Change, Change.....	50
3.2 The Jakobetz Case.....	54
3.3 United States v. Yee et al.....	57
3.4 The Population Substructure Debate and the NRC I.....	62
Chapter 4: With Truces Come New Troubles.....	68
4.1 NRC II.....	68
4.2 Broad Implications.....	70
4.3 Specific Events and Cases.....	72
Chapter 5: Conclusion and Implications for Additional Research.....	77

Chapter 1: Constructing the Double Helix

“When law turns to science or science turns to law, we have the opportunity to examine how these two powerful systems work out their differences,” (Lynch et al, 2010).

1.1 Introduction

Forensic science, which can be essentially any scientific technique employed with the purpose of aiding a legal investigation, has long faced the challenge of operating between two very different worlds. The first of these, the scientific, is a place of quantifiable data and testable hypotheses. In the scientific realm, new ideas are tested and retested, debated and protected, questioned and prodded until they are generally accepted or rejected. Even then, debate runs rampant through almost every scientific community. On the other hand, the legal arena is one guided by complex rules and statutes, often governed more by traditions rather than experiments. However, the legal world is often as much guided by individual human perception and interaction, demonstrated by the role of jury members as decision makers. Additionally, from a legal standpoint there are only two possible outcomes. The party must eventually be determined either innocent or guilty. However, for both groups the underlying goal involves a certain quest for the truth (Haack, 2004).

These two starkly contrastable worlds must combine when forensic experts are called to testify in the legal realm, where they present evidence based on scientific techniques, which the jury must then take into consideration. Specific examples of forensic science include fingerprinting, odontology, entomology, botany, and toxicology, among many others. The technological differences between these varied techniques raise

questions between both the scientific and legal communities over which scientific practices should be admissible as evidence, and which ones should not. If a scientific technique is new but tested, can that be presented as evidence? On the other hand, if a scientific method is older and tested, but now under debate, how should that be treated? These questions are important to consider because misuse of a scientific technique in a legal case can lead to misrepresentation of the evidence. Distorted evidence can lead to the incarceration of an innocent person, singlehandedly defeating the purpose of a forensic technique meant to aid the legal system in promoting justice. On the other hand, eliminating all debated forensic methods in a world of constant scientific flux would make most evidence inadmissible, greatly reducing efficiency and accuracy in the criminal justice system. So the forensic sciences as a whole must progress, keeping in mind the importance of practicing reliable techniques capable of functioning within both worlds, in eternal pursuit of a forensic panacea.

Some experts believe that this panacea has been found. Early types of forensic DNA analysis, also referred to generally as DNA profiling or fingerprinting, were discovered in 1984, and at work within the legal system by 1988 (Thompson, 1993). Since then, DNA typing techniques have risen to the frontline of forensic science and crime scene investigation (Thompson, 1993). DNA typing has excelled as an aid for both convicting the guilty as well as exonerating the innocent. By 2012, 305 people in the United States had been released due to DNA based exonerations (Saks and Koehler, 2005). In addition, the true perpetrator was caught in 147 of these cases. While forensic DNA analysis continues to assist in multitudes of convictions and exonerations, other forensic techniques such as fingerprinting and hair follicle analysis have come under

scrutiny regarding scientific reliability and accuracy (Saks and Koehler, 2005). As more and more of these methods come under fire, DNA typing seems to stand alone as the technique capable of consistently providing courtrooms with bold and definitive conclusions (Saks and Koehler, 2005). Experts in the legal field have even started referring to DNA as a “truth machine” as well as “a gold standard for truth telling” (Aronson, 2007). Many forensic and legal experts are beginning to suggest that the methodology and procedural standards surrounding forensic DNA analysis should be revered as a new model for future forensic methods (Aronson, 2007). However, in order to even begin to attempt this, we must first ask, in what ways has forensic DNA analysis developed differently from other techniques?

Certainly the answer is multi-faceted. One explanation is that DNA profiling had early beginnings in academic science, whereas other forensic sciences were developed in forensic labs outside the scholarly world (Aronson, 2007). However, it is unlikely that this is a complete answer. It certainly seems more likely that the correct answer would involve both the scientific and legal worlds. In examining the early history of DNA typing, one can see that the science was not always as reliable as we think of it today (Thompson, 1993). Forensic DNA analysis incorporates a number of different methods and techniques, many of which have evolved dramatically over the last 20 years. By tracking these changes it becomes evidence that DNA typing involves a detailed history, one complicated even further by controversy from both legal and scientific perspectives. Jay D. Aronson, a history professor at Carnegie Mellon University, explains, “developing the technique (DNA typing) for use within the criminal justice system was just as much about social engineering as it was about getting the science right.” (Aronson, 2007). The

arrival of forensic DNA analysis into the legal system was driven as much by businessmen, defense attorneys, prosecutors, and journalists as it was by scientific improvements.

Although forensic DNA analysis was eventually challenged thoroughly in the legal realm, in the beginning the technique was met with almost universal acceptance and admissibility in court. Initially, defense attorneys struggled unsuccessfully to demonstrate the potential flaws and setbacks of the early technology. This legal history, documented by specific case studies, can be divided into three general time frames, each with its own dominating themes. During the first string of cases, law enforcement agencies sent evidence samples to private biotechnology companies for analysis. At the same time, defense attorneys struggled to combat DNA based evidence in court because the technology used by the private companies was strongly backed by the scientific sector. By the arrival of the second stream of cases, the FBI had assumed a role in forensic DNA analysis and attempted to standardize most of the methods and procedures. Controversy from these two periods eventually exploded, igniting a vigorous scientific debate often now referred to as the 'DNA Wars'. With the conclusion of the 'DNA Wars' came a third group of cases, where scientific advancements ignited new concerns regarding laboratory error and regulation. Over many years these scientific and legal concerns for the most part dissipated, permitting forensic DNA analysis to assume the role it fills to this day, a valuable tool in the criminal arena.

Hence, the purpose of this paper is to use individual case studies to demonstrate how persistent defense attorneys and scientists improved and evaluated forensic DNA technology through the legal system. Certainly, the end goal is in no way to discredit

DNA analysis, but to demonstrate how questions, controversy, and eventually compromise from both the legal and scientific worlds improved forensic DNA technology over the course of many years. Hence, the overarching purpose is to demonstrate through these case studies the importance of cooperation and compromise among various disciplines in order to benefit the United States justice system as a whole.

To demonstrate this, I will briefly chronicle the early history of DNA profiling, demonstrating a basic understanding of the technology itself, as well as how it came to exist in the United States in the late 80s. I will then outline the history of admissibility of evidence in the United States, focusing on what standards were relevant to each case. The paper will then unfold by chronicling the struggles faced by those involved in early forensic DNA analysis, from defense attorneys to forensic scientists to population geneticists. To conclude, I will examine modern problems with DNA profiling. By chronicling the complex and multifaceted development of DNA profiling I hope to examine and critique the rise of a successful forensic science, uncovering useful information for developing techniques in the future. However, I also hope to demonstrate the complications involved with the evolution of DNA profiling, stressing that there is no simple recipe to follow guaranteeing the successful integration of a scientific technique into the forensic canon.

As an introductory caveat, it must be clear that the intentions of this paper are in no way to discredit the many prosecutors, scientists, and expert witnesses who researched and testified as proponents of early DNA profiling. Certainly, as with most sophisticated disciplines, there were brilliant and hardworking people on both sides of the debate. However, since these proponents were generally part of the accepted scientific paradigm,

there would be little purpose to discussing them in detail in a paper that attempts to demonstrate how the inclusion of opinions from the non-dominant paradigm in the courtroom setting works to benefit scientific technologies overall.

1.2 How Does Forensic DNA Analysis Work?

In order to understand the technology involved with DNA analysis, it is important to have a basic understanding of what DNA is and how it works. DNA, or deoxyribonucleic acid, is a gene-encoding molecule that exists within every living organism (Kaye, 2010). DNA is unchangeable from cell to cell within an individual, and contains all the genetic information necessary for the “construction of a human being,” (Calladine and Drew, 1997). The DNA molecule consists of subunits made up of four nucleotide bases, adenine, thymine, guanine, and cytosine, also shorted to A, T, G, and C (Kaye, 2010). The basic form of DNA, known as the double helix, consists of two strands that wrap around each other, connected by bonds between these nucleotide bases (Kaye, 2010). Each base will only pair with one complimentary base, as A pairs only with T, and G pairs only with C (Calladine and Drew, 1997). Although each cell contains a complete set of an individual’s DNA, and hence a complete set of genes, the cell only activates the genes specific to the cell’s particular function (Calladine and Drew, 1997).

Most human DNA is tightly wound into 23 pairs of chromosomes, which are found in the nuclei of most cells (Thompson and Krane, 2003). Within these chromosomes there are genes, “which lie at various positions (loci) along the chromosome,” (Kaye, 2010). The term ‘Gene’ is essentially shorthand for molecular biologists and geneticists, as it generally refers to a segment of DNA that codes for a protein (Nelkin and Lindee, 2004). An allele is the alternative form of a gene that is

located at a specific position on a specific chromosome. According to the Human Genome Project, there are approximately 30,000 genes in the human body, which usually range from 1,000 to 10,000 base pairs long (Kent et al., 2002). Genes include coding and non-coding DNA sequences (Kaye, 2010). The coding sequences of the base pairs encode information then used by cells to generate proteins, which are made up of chains of Amino acids (Kaye, 2010). Non-coding sequences regulate gene operation, determining the amount of protein the gene produces (Kaye, 2010). In the mid 1980s, genetic identification was achieved by analyzing the variations in the proteins expressed by genes. This allowed for some forensic testing using serology, a science focused on the serums found in bodily fluids (Kaye, 2010). However, the use of serology in forensics was limited, as test results could only determine that the suspect was either included or excluded from the population that the sample evidence matched. Forensic DNA analysis promised a number of improvements over these previous techniques (Kaye, 2010).

Interestingly, at least 99.5% of the base pairings in human DNA are the same in every individual (Kaye, 2010). Only approximately .05% of the sequence is unique to each individual person (Kaye, 2010). Essentially, DNA typing techniques search for variations in our genetic codes that then allow forensic scientists to, in a best case scenario, identify someone, or at least distinguish one individual sample from another. Although there are various techniques for analyzing forensic DNA, the task is generally accomplished by comparing highly polymorphic regions of DNA, or regions of DNA that are highly variable between individuals (Thompson, 1993). “These polymorphisms result in different forms, or alleles, of genetic markers,” (Weedn, 2007).

The beginnings of forensic DNA analysis involve RFLP (Restriction Fragment Length Polymorphism) analysis. Scientists conduct RFLP analysis by extracting DNA from cells and adding a restriction enzyme (Thompson, 1993). The restriction enzyme then binds to specific areas of the DNA made up of short sequences of base pairs (Kaye, 2010). The restriction enzyme slices the DNA at the specific location within the sequence, creating a fragment thought to be unique to an individual (Kaye, 2010). These restriction fragments are then separated by size using gel electrophoresis. Gel electrophoresis is a technique where strands of DNA are placed in holes in a slab of gel (Kaye, 2010). An electric field is then applied to the gel, and the DNA fragments, which are negatively charged, pull through the gel toward the positive pole (Kaye, 2010). The smaller pieces of DNA are able to move farther through the gel than the larger fragments (Kaye, 2010). Hence, “the length of any particular fragment can be measured by comparing the distance it has traveled with the distances traversed by standard fragments of known size placed in a parallel slot in the gel,” (Kaye, 2010). The fragments are transferred to a sheet of nylon, generating manageable results for analysis (Kaye, 2010). A probe, a short single strand of DNA with a radioactive component, is applied to these DNA strands. When the sequence of bases in the probe meets complementary base sequence in the DNA, they bind together. The nylon is then placed between two sheets of film, and eventually the radioactive probe exposes the film (Kaye, 2010). “The result is an autoradiograph, or autorad- a visual pattern of bands representing DNA fragments that contain the specific base-pair sequence targeted by the probe,” (Kaye, 2010). The pattern on the autorad demonstrates the DNA genotype for the individual at the locus

associated with the probe, for example being heterozygous or homozygous for a specific allele (Kaye, 2010).

However this technique still fell short in terms of forensic investigation. Alec Jeffreys, the British scientist credited with inventing the first version of DNA profiling, explains “RFLPs...were difficult to find and assay, and did not tell you much about variation between people- you either had the change or you didn’t,” (Jeffreys, 2005). The next significant step forward in DNA analysis involved variable number tandem repeat locus (VNTR). The technique itself was discovered by Jeffrey’s and his team in 1984, and was used to examine VNTR locus, loci consisting of base repetitions in a particularly short sequence that are highly variable from person to person (Aronson, 2007). By 1985, Jeffreys’ team had created a probe that could detect these highly variable loci simultaneously, providing an “individual specific DNA fingerprint,” where multiple bands were produced per probe (Kaye, 2010). The resulting pattern is often described as a barcode (Weedn, 2007). Four months after this, Jeffreys’ team announced more multi-locus probes that could be used for genetic identification (Kaye, 2010). Unfortunately, the multi-locus approach required a substantial amount of high-quality DNA, which is not always present at a crime scene.

The next advancement was a single locus probe (SLP). Single locus probes are constructed so the DNA strands only bind with the alleles at one specific locus, meaning less DNA is required (Kaye, 2010). Because only one locus is involved, the data reveals only one “single, highly polymorphic, restriction fragment length polymorphism,” lessening ambiguity in the data (Jobling and Gill, 2004). The band patterns produced by SLP technology are easier to read than the convoluted multi-locus results (Weedn, 2007).

The results also allow for more precise frequency statistics because data can be collected regarding how often different alleles occur at these specific locations among different populations (Kaye, 2010). However, single locus probe analysis does not produce particularly individualized results. The test was improved by adding other single locus probes and is called multiple single locus probing (Kaye, 2010). For multiple single locus probe tests, the forensic community chose six genetic loci for general use, D1S7, D2S44, D4S139, D10S28, D14S13, and D17S79 (Weedn, 2007). However VNTR testing was still problematic. The technique was not very efficient, gel electrophoresis measurements were messy and far from exact, and large samples of DNA were needed (Kaye, 2010).

The next development in DNA typing technology started when Kary Mullis, a biochemist, had the idea to mimic the natural process of DNA duplication, known as the polymerase chain reaction (PCR) process, in a laboratory setting (Kaye, 2010). The PCR amplification process can be mimicked by heating up a DNA sample, reducing the DNA from its original double helix structure to a single strand structure (Thompson, 1993). When the sample is cooled, these single strands become assessable to primers, which lock onto sections of the DNA where copying will start and stop (Thompson, 1993). An enzyme, DNA polymerase, is added to the mixture, which is then warmed to a temperature at which the polymerase enzyme begins to insert the matching base pairs, rebuilding a second strand to each original template (Kaye, 2010). The resulting product is two identical double-stranded DNA segments (Kaye, 2010). PCR itself is not a typing method, but a way to efficiently create DNA copies, allowing for distinguishing among different alleles. In 1991, Perkin-Elmer (PE), a biotechnology company, introduced a test

kit that could be used to amplify and type the DQ-Alpha gene (HLA DQ alpha/DQA1), a sequence polymorphism found on chromosome 6 (Butler, 2009). The first kit could only test distinguish six different alleles, or define 21 genotypes, making it much less individualized than a VNTR locus analysis (Butler, 2009). PolyMarker (PM+DQA1), an improved kit released by PE in 1993, co-amplified parts of the HLA DQ alpha gene as well as five other DNA segments (Butler, 2009). Starting in the mid 90s, PCR amplification was used with a VNTR named D1S80, which worked because the “D1S80 VNTR is unusually short, which also permits the number of repeats of its sixteen-base-pair core sequence to be determined, thereby avoiding the statistical complications of match windows, binning, and ambiguous single bands,” problems which will all occur in the following case studies (Kaye, 2010). Like RFLP-VNTR analysis, there were also some problems with PCR testing; mainly that it is susceptible to manipulation and contamination (Weedn, 2007). With these developments in PCR testing, VNTR testing was virtually abandoned. In 1988, 43 out of 100 laboratories reported using VNTR for casework, while by 2000; only one laboratory reported using VNTR analysis for casework (Weedn, 2007).

PCR based Short Tandem Repeat analysis (STRs), is a modern technique used for forensic DNA analysis (Weedn, 2007). An STR, also known as a microsatellite, is a DNA locus containing a length polymorphism composed of very short patterns, or core repeats (Weedn, 2007). Because STRs are much shorter than VNTRs, the PCR process is more efficient. By 1997 the FBI had identified 13 standard STR loci to be used for forensic testing (Thompson and Krane, 2003). Since then, a number of commercial

biotechnology companies have developed kits and equipment for analyzing STRs (Thompson and Krane, 2003).

Modern forensics also includes a number of more recent techniques. One interesting new method is the Y-STR test. This test involves examining polymorphic areas from a male Y-chromosome. This technique is valuable in cases of sexual assault, where sample DNA may be a mix of victim and assailant (Thompson and Krane, 2003). Another new technique involves Mitochondrial DNA, or mtDNA, the DNA found in the cell mitochondria rather than the cell nuclei (Thompson and Krane, 2003). mtDNA analysis is useful when the DNA sample is degraded because there are “generally 500 to 2,000 copies of mtDNA per cell, compared to one set of diploid chromosomal, nuclear DNA,” (Weedn, 2007). The high number of mtDNA copies per cell increases the likelihood of obtaining a result, especially when examining older skeletal material (Weedn, 2007). mtDNA is also present in tissues that do not contain nuclear DNA, like fingernails (Weedn, 2007). For identity testing, the mtDNA sequence is obtained and compared to a reference sequence (Weedn, 2007). However, since mtDNA testing involves DNA sequencing it is labor intensive, expensive, and sensitive to contamination (Weedn, 2007).

However, in the forensic setting, using these techniques to report the existence of a match is not enough. The jury needs to know how common or rare the allele frequencies are within a given population. Reporting a random-match probability, or likelihood ratio, is important because “without the probability assessment, the jury does not know whether the matching patterns are as common as pictures with two eyes, or as unique as the Mona Lisa,” (United States v. Yee et al). Although much debate has

occurred over the exact methods for calculating random match probabilities, the underlying structure is demonstrated below.

$$\frac{P(E)|H}{P(E)|\text{not } H} \quad (\text{Lempert, 1997}).$$

Basically, “E is the fact that the DNA in the Evidence and suspect lanes of a gel match and H, the hypothesis in dispute, is that the suspect is guilty (Lempert, 1997). So for DNA evidence,

$$\frac{P(\text{DNA Match})| \text{Defendant's Guilt}}{P(\text{DNA Match})| \text{Defendant's Innocence}} \quad (\text{Lempert, 1997}).$$

However, all this ratio demonstrates is the possibility of a coincidental match, without including possible errors such as contamination or manipulation of the data. The original way to calculate these probabilities is known as the product rule. The product rule is used in statistics and mathematics to calculate the likelihood of a number of independent events occurring together. Scientists and statisticians generate the actual figures used in this ratio by calculating “the frequency with which a given genotype of all the alleles tested will be found in a white, black, or Hispanic population,” (Weedn, 2007). This is calculated by multiplying the allele frequencies of individual alleles in a given ethnic population, using a population database, a technique also referred to as the product rule (Weedn, 2007). The outcome of this calculation is often referred to as *P_m*. However, in order for this calculation to be assumed correct, we must be certain that the events are independent from each other. In the case of DNA profiling, the use of the product rule demonstrates the assumption that both the inheritance of the particular alleles and the

inheritance of the loci are independent events. These things cannot be observed from outward appearance.

For the independence of the alleles themselves we can reference the Hardy Weinberg Principle, an algebraic expression that can be used to describe genetic equilibrium within a population. “According to the Hardy-Weinberg principle, gene frequencies will remain constant from generation to generation within a population unless outside forces act to change it, so long as random mating occurs,” (Harris, 1988). When the gene frequencies of a certain population vary from the expected gene frequencies demonstrated by the Hardy-Weinberg expression, then equilibrium is not attained (Harris, 1988). Assuming a large, randomly mating population, the Hardy Weinberg Equilibrium (HWE) suggests that alleles will associate at random, making it an independent event. It is also necessary to determine if there is any relationship, or “linkage,” between inheriting alleles at one locus to inheriting alleles at another locus, (Aronson, 2007). “Linkage equilibrium” is the idea that alleles are associated randomly with regard to loci. Assuming both Linkage equilibrium and HWE, the product rule would correctly calculate the likelihood of a nonrelated random match.

Arguments against the calculations produced by the product rule mainly involve the concern that the human population may be structured into subgroups, where certain DNA profiles are more predominant in certain ethnic, religious, or geographic zones (Thompson and Krane, 2003). If the reported figure is calculated ignoring this uncertainty regarding population substructure, than the ratio presented to the jury may greatly overstate the value of the evidence (Thompson and Krane, 2003). Hence, the process of calculating probabilities assuming both equilibriums is often referred to as the ‘simple’

product rule (Balding and Nichols, 1994). As of now, in most forensic cases “the evidential impact of a match is very strong and the overstatement of the strength of the evidence by perhaps one or two orders of magnitude may have little practical effect,” (Baling and Nichols, 1994). However, population substructure became a subject of huge debate in the early 90s, and it will be discussed in detail throughout this analysis.

1.3 DNA Typing: An Early History

After Jeffreys’s discovery in 1984, DNA fingerprinting exploded. The technique entered the British legal system by 1985, and moved to the United States shortly after. The United Kingdom was also home to the first instance where the prosecution employed DNA evidence to support a criminal case (Aronson, 2007). The case, brutal by nature, emerged in 1983, when the body of a teenager, eventually identified as Lydia Mann, was discovered in Leicestershire (Aronson, 2007). Although police believed the girl had been viciously raped and murdered, the case remained unsolved. Over time the case grew cold. Years later, at a time when even the family had abandoned all hope for justice, another body was found that demonstrated patterns from the first attack. This time it was the body of fifteen-year-old Dawn Ashworth; stripped naked, raped, and murdered (Aronson, 2007).

A full-scale investigation was launched, and the case became so publicized that a novel chronicling the events entitled “The Blooding,” was eventually published (Aronson, 2007). However, the investigation took a turn for the worse when mentally handicapped seventeen-year-old Richard Buckland was arrested. While under questioning, Buckland even falsely confessed to the first murder (Aronson, 2007). Luckily for Buckland,

Jeffreys agreed to attempt to extract and analyze DNA from sperm collected from the two victims before Buckland was convicted of the crime (Aronson, 2007). An SLP based analysis was conducted (Jobling and Gill, 2004). The results were astounding, demonstrating that Buckland's DNA did not match that the sperm sample take from either of the victims (Aronson, 2007). However, extracted DNA taken from vaginal swabs of both victims was a match, demonstrating that they were killed by the same person (Aronson, 2007). The tale, which includes rounding up all of the men from the town, cannot be discussed in length at the moment. However, it suffices to know that the actual killer, Colin Pitchfork, was eventually arrested and convicted (Jobling and Gill, 2004). This dramatic success story spread the word about DNA profiling through both the European and American forensic communities, showcasing DNA analysis as a powerful new tool capable of convicting the guilty but also protecting the marginalized, such as the Buckland, who could not protect himself.

This dramatic but just ending to a long and drawn out investigation generated free advertising for DNA typing around the globe. The first DNA profiling laboratory in the United States was opened by a United Kingdom based company previously known as Imperial Chemical Industrial PLC but later renamed Cellmark Diagnostics USA (Aronson, 2007). By 1987, Cellmark was offering the "DNA fingerprinting" multi-locus tests. ICI believed they would have difficulty entering the American Market because of more strict admissibility rules in most courts. As a precaution, they hired Daniel Gardener, a man who had worked extensively for the US bureau of alcohol, tobacco, and firearms throughout much of the 80s (Aronson, 2007). However, these concerns seemed unnecessary as DNA typing technology quickly evolved from a tool for investigating

paternity questions and immigration issues to a technique used in criminal cases. By 1988, DNA typing had been admitted as evidence without challenge in almost 200 legal cases (Aronson, 2007).

Prior to 1988, Cellmark found itself at a clear market disadvantage when compared to the biotech company Lifecodes, since Cellmark only offered multi-locus testing (Aronson, 2007). Since both companies were “large publically traded and multi-national conglomerates” it was extremely necessary to dominate forensic DNA typing market as heavily as possible (Aronson, 2007). Facing the competition head on, Cellmark emerged with dual analysis testing featuring both single and multi-locus probes by 1988 (Aronson, 2007). Around this same time, Cellmark launched a gigantic marketing campaign aimed at convincing both legal and scientific professionals as well as the general public that DNA typing could produce “conclusive” and “significant” results in both forensic and paternity cases (Aronson, 2007).

Lifecodes was also working diligently to dominate the market for DNA analysis. Although Lifecodes began as a research company in 1983, the company did not reign in their focus to VNTR testing until around 1986 (Aronson, 2007). Like Cellmark, Lifecodes also launched a huge public relations campaign to convince people of their equally reliable techniques (Aronson, 2007). Demonstrating the effects of the competitive private market, neither company made any attempts to examine possible problems with DNA profiling (Aronson, 2007).

1.4 Evidence Admissibility in the United States

Before jumping into some fairly complex case studies it is also important to have a brief understanding of the admissibility rules that have governed the United States over the last century. The first important ruling regarding admissibility of evidence was Frye versus the United States. The Frye test, also known as the general acceptance test, came out of a lawsuit regarding whether or not the results of polygraph tests should be admissible as evidence (Giannelli, 1980). The result of the Frye vs. United States ruling ensured that scientific evidence in question must be generally accepted by the associated scientific community to be admissible. However, as many scholars will attest to, scientists are rarely in complete agreement, so courts and legal administrators began to work around the Frye rule as much as possible. Some courts even started ignoring the rule entirely (Christensen, 2004).

To deal with these inconsistencies within the courts, in 1975, a new set of rules were enacted, known as the “The Federal Rules of Evidence,” (Christensen, 2004). Among many, Rule 702 specifically stated that,

If scientific, technical or other specialized knowledge will assist the trier of fact to understand the evidence or to determine a fact in issue, a witness qualified as an expert by knowledge, skill, experience, training, or education may testify thereto in the form of an opinion or otherwise, (Christensen, 2004).

Unfortunately this was as problematic as it was successful. Since the Federal Rules lacked any mention the Frye Rule, courts were unsure which rules to follow. Many courts continued to employ the Frye rules, while others exercised the new rule 702. Other courts chose creative models employing what they preferred from each text (Christensen, 2004). Certainly there was no unity or end of confusion brought around the

new rule 702. For example, the Frye standards remained in effect in New York at the time of the trial of Joseph Castro in 1989 (Lynch and Jasanoff, 1998).

In 1993, following a Supreme Court decision regarding the pill Bendectin, an anti-nausea medication prescribed to pregnant women, a new set of evidentiary standards was enacted. This new set of general suggestions for accepting evidence became known as the Daubert standards, and worked as a compromise between the two previous standards.

The specific Supreme Court decision determined two important things. First, the court decided that Federal Rules of Evidence, having been put in place more recently, superseded the Frye test, opening admissible evidence to minority opinions as long as they are backed by reliable studies and emerging research (Christensen, 2004). Essentially, the court decided that reasonable minority opinion may be admitted into evidence, in the form of well-designed studies. Second, the court included four guidelines to be used, called the Daubert Guidelines, which are outlined below.

“Guidelines for determining whether evidence is scientific and therefore admissible under Federal Rule 702.

- 1) The content of the testimony can be (and has been) tested using the scientific method.
- 2) The technique has been subject to peer review, preferably in the form of publication in peer reviewed literature
- 3) There are consistently and reliably applied professional standards and known potential error rates for the technique
- 4) Consider general acceptance within the relevant scientific community”

(Christensen, 2004).

The cases I have chosen to focus on for this particular paper generally occur before the Daubert Standards were enacted. However, I have chosen to discuss these individual pieces of Daubert because they can be seen unfolding in some of the earlier

case studies. It is also important to keep in mind that even with the arrival of the new standards, state courts could decide which admissibility standards to follow. For instance, at the time of the O. J. Simpson trial in 1995 in California, the court still followed the Frye general acceptance standard (Lynch and Jasanoff, 1998).

Chapter 2: Private Science, Public Defenders

“The legal history of DNA profiling in England and the United States shows a pattern of rise, fall, and then rise again: rapid acceptance in the late 80s, followed by challenges in the courts and science press from 1989 through the mid 1990s, followed by a renewed acceptance at a stronger level.” –Michael Lynch (Lynch et. al, 2010).

2.1 Early Trials and Tribulations

DNA evidence moved into the American law system with an impressively streamlined efficiency. The technique, although very new, was quickly and widely accepted in various courtrooms throughout the late 1980s. This early success was partially due to the extensive campaign efforts set forth by both Cellmark and Lifecodes, as well as positive media attention generated from DNA evidence based success stories reported from England (Aronson, 2007). This positive glow that seemed to be surrounding these new DNA techniques generated a number of problems for practicing defense attorneys. Although there were countless scientists willing to testify for DNA evidence, Cellmark had even released a list of “independent expert witnesses”, defense lawyers had a very difficult time tracking down anyone who was willing to testify as an expert witness against DNA based evidence (Aronson, 2007). Perhaps some scientists were hesitant to testify because they did not want to be discredited by in court by their

peers, opposing expert witnesses who were often respected and well-known scientists. Also, in some states, trial judges denied public defenders' requests to hire expert witnesses while in other states the maximum pay of \$1000 was simply too low to attract scientists (Gerald, 1994-95). Arthur Daemmrich, a professor at Harvard Business School, even argues that companies like Cellmark and Lifecodes set up a kind of vertical integration structure common to large firms where in this case the product was convincing expert testimony (Daemmrich, 1998). This idea of obtaining market control over something so intangible as expert witness testimony demonstrates an alarming level of corporate involvement in the legal system. Equally troubling is the amount at which these two biotech giants influenced the developing technology from a scientific standpoint. This underlying issue of corporate sway over scientific authority is demonstrated by Richard Roberts, who stated in court that, "scientific knowledge is established, assimilated, and transmitted by social trust and authority, rather than by the radical skeptical testing suggested by science's dominant public image," (Aronson, 2007).

Throughout the following case studies it may be helpful to refer to Figure 1, page 26.

Figure 1: A Visual Guide to the Significant DNA Profiling Techniques with Reference to Relevant Cases

	1985	1990	1995	2000
G.E., RFLP specific	1984-1985			
G.E., MLP specific	1986-----1989			
In 1987, Colin Pitchfork, became the first person ever convicted of murder based on DNA profiling evidence.				
G.E., SLP specific	1986-----2000			
*State v. Woodall (1987), **Andrews v. State (1988), ***State of New York v. George Wesley & State of New York v. Cameron Bailey (1988), ****State of New York v. Castro (1989), State v. Schwartz (Minnesota), *****United States v. Jakobetz (1990), *****Yee et. al.(1991), ***** People v. Orenthal James Simpson (1994-1995)				
SLP technology started to gain footing around 1986 and was the predominant method for DNA profiling until the mid 90s. By 2000, most forensic labs had abandoned RFLP technology in favor of PCR				
Early PCR	1988----1989			
The PCR technique was developed in 1983. However PCR in itself is a method for DNA amplification rather than DNA profiling. Some early DNA profiling techniques using PCR amplification appeared briefly in some court cases but only as additional evidence or when the VNTR analysis was unsuccessful.				
DQ-Alpha (PCR)	1991-----1996			
*State v Woodall (1991), People v. Orenthal James Simpson (1994-1995)				
First test kit produced in 1991, second in 1993				
STRs (PCR)	1995-----→			
In 1997 the FBI identified 13 STR loci to be used in forensic testing.				
Modern Technology	2000-----→			
Includes Y-STR, Mini-STR, mtDNA				

Annotations

* In Woodall, a request for DNA analysis from the defense attorney was denied during the pre-trial in 1987. SLP analysis conducted in 1991 demonstrated that Woodall could not have been the perpetrator. Also, PCR analysis of the DQ-Alpha gene demonstrated a matching DQ type for the assailant in both cases but not matching Woodall. The case was dismissed in 1992.

** In Andrews, Lifecodes presented evidence for a DNA match based on SLP analysis. The case involved the first admissibility hearing regarding DNA profiling. The evidence

was found admissible. Interestingly, Lifecodes had just completed a colossal advertising campaign that spanned from 1982-1987.

*** For Wesley-Bailey, a combined admissibility hearing was held in New York; eventually ruling that evidence produced by the same type of Lifecodes test was admissible in court.

****A bloodstain from Castro's watch was collected and sent to Lifecodes for analysis in 1987. Lifecodes reported a random match calculation of 1 in 189,200,000 using SLP testing with three DNA probes. This was the first case where forensic DNA evidence was found inadmissible.

***** In this admissibility hearing involved Cellmark's SLP technology. Eventually the Minnesota Supreme Court deemed the evidence inadmissible.

***** In Schwartz, an FBI laboratory represented results based on SLP analysis, which were admitted.

***** In Yee, SLP based evidence produced by an FBI laboratory was found admissible.

***** In 1995, Simpson withdrew his objections to DNA admissibility, which included SLP and PCR analysis. His lawyers contended that they reserved the right "to challenge the weight and reliability of the prosecution's scientific evidence at trial," (Kaye, 2010).

2.2 People v. Andrews

DNA based evidence faced its first major challenge in Orlando, Florida in 1988 when a neighborhood in Orlando experienced a string of seemingly related rapes (Lewis, 1989). Each occurrence was eerily similar. The perpetrator would sneak into the victim's house, switch off all of the lights, and cover the victim's face with a sleeping bag (Clayborn, 1989-90). Some victims reported the feeling of a razor blade being held to their necks (Kaye, 2010). The perpetrator seemed to leave nothing behind and the police had no leads (Aronson, 2007). To his victims, the rapist existed as an enigma, unknown and unpunished.

With the community on high alert for suspicious behavior, a suspect was eventually found and arrested. The man, Tommie Lee Andrews, was spotted "prowling in a woman's yard in the wee hours of the night," (Aronson, 2007). The police chased Andrews until he crashed his car into a telephone pole (Kaye, 2010). Andrews was

charged with armed robbery and sexual battery against Nancy Hodge, as well as with similar charges for another recent rape that occurred shortly before he was arrested (Aronson, 2007). Although the circumstances certainly seemed to point to his guilt, the evidence was few and far between. There was limited material evidence, consisting only of semen collected from the victims and two fingerprints left on a window (Moss, 1988). Due to the nature of the attacks there was also only one partial eyewitness, Nancy Hodge (Aronson, 2007). Even so, eyewitness accounts are known for being unreliable and prosecutor Tim Berry knew he needed something more in order to incarcerate Andrews (Warden, 2005).

The initial evidence used in the case besides the eyewitness account was blood typing analysis. However the results were inconclusive. Although blood taken from Andrews matched the sample, his blood type is shared by 2/3 of the American male population (Aronson, 2007). The success of the case depended on finding more concrete evidence. To do this, Berry consulted with Jeffrey Ashton, a colleague, who suggested that he look into the new techniques surrounding DNA analysis (Aronson, 2007). Berry sent a sample to Lifecodes for analysis (Aronson, 2007). Dr. Michael Baird a scientist working at Lifecodes conducted the analysis and reported that Andrews DNA matched that of the rapist, with a one in ten billion chance that the match was a coincidence (Thompson, 1989-1990). According to results found by the Lifecodes administered test, “Andrews’ was beyond a doubt the source of the semen in two of the six cases,” (Aronson, 2007). Andrews was eventually convicted in both cases (Moss, 1988).

However before the trial the evidence had to be deemed admissible so a hearing was conducted on the acceptance of single locus VNTR analysis (Kaye, 2010). In 1988,

when the Andrews' case was brought to trial, even courts within the state of Florida were divided on whether to follow the traditional Frye standards or whether to follow the more recent federal rules of evidence, which used a less strict "relevancy approach" from rule 702 (Aronson, 2007).

The Florida Evidence Code at the time stated that:

"if scientific, technical, or other specialized knowledge will assist in trier of fact in understanding the evidence or in determining a fact in issue, a witness qualified as an expert by knowledge, skill, experience, training, or education may testify about it in the form of an opinion; however the opinion is admissible only if it can be applied to evidence at trial," (Aronson, 2007).

For this particular case, it was not that the technique itself needed to be accepted by the general scientific community, but instead that the scientific technique in question needed to be backed by an expert witness, one who could withstand questions of credibility. Berry had little difficulty finding a scientist willing to defend the test results, various expert witnesses quickly agreed to support the prosecution, including David Housman, a biology professor, Alan Giusti, a forensic scientist from Lifecodes, and Michael Baird, a geneticist who worked as the manager of forensic testing at Lifecodes (Kaye, 2010). For his testimony, Housman even personally visited the Lifecodes labs, concluding that the DNA testing occurring there matched the consistency and quality of the DNA tests conducted in academic labs across the globe (Aronson, 2007). Certainly these were brilliant scientists, but their strong affiliations with Lifecodes continue to raise questions of financial motive and corporate gain.

While Berry gathered a team of witnesses to support the prosecution, defense attorney Hal Uhrig struggled on the other side end of the paradigm. He simply could not find any scientist willing to testify against DNA forensic evidence (Aronson, 2007). Uhrig even “recalled that Housman was like ‘an angel’ in the scientific community,” (Aronson, 2007). Nobody wanted to testify against him. The scientists he did ask echoed a consistent theme, that if Housman accepted the quality of the evidence, they had no reason to question his judgment (Kaye, 2010). Uhrig, a man with little background knowledge regarding DNA and no expert witnesses willing to testify for the defense, had to formulate his own strategy (Aronson, 2007).

He formulated a number of responses for the prosecution, one of which included questioning how the chance of a random match occurring was one in ten billion, a number greater than the entire population of the earth (Aronson, 2007). The answer, obvious to the scientific community, is that the figure was a probability based on statistical calculations, making the population of the earth simply not relevant. (Aronson, 2007). However Berry was unprepared for the question and the statistical evidence was deemed inadmissible, resulting in a hung jury (Aronson, 2007). The case was scheduled for retrial.

Upon retrial Housman again testified as an expert witness, along with Michael Baird, a scientist working for Lifecodes. Ashton took over the prosecution, and prepared in advance for questions regarding the admissibility of statistical evidence regarding DNA typing (Aronson, 2007). Again Uhrig could find no one to testify against the DNA based evidence. In the trial, Uhrig brought

up the important issue of human error, and questioned how results from Lifecodes could be trusted when no one had witnessed the actual process (Aronson, 2007). Nonetheless, Andrews was found guilty and sentenced to twenty-two years in prison. When Andrews appeared in court again for the retrial of the Hodge case, Uhrig questioned how DNA evidence could be accepted when only a small fraction of the DNA was examined (Aronson, 2007). Again, he was quickly shot down when Baird explained that only particular regions of the DNA sequence were useful in forensic investigations because they were highly variable from individual to individual (Aronson, 2007). His argument overall can be summed up by his statement, “the tests may be the test of choice today but they may turn out to be fallacy later,” (Benedict, 1988).

Andrews was again found guilty and sentenced to 78 additional years in prison (Moss, 1988). He was the first person in the United States to be convicted of a crime where the prosecution used forensic DNA analysis as evidence (Clayborn, 1989-90).

What is interesting here cannot be observed by looking solely at the outcome of the case. Andrews was found guilty of the crimes that he committed, a result that is difficult to critique. However, what needs to be addressed is the lack of scientific cooperation in questioning DNA typing techniques used in the courtroom. The questions put forth by Uhrig demonstrate the gaps in his understanding of DNA typing, gaps that could have been filled if he had worked an expert witness. One of his main points, regarding the problem of not analyzing the entire DNA molecule, was simply misguided. However his inquiries about

laboratory procedures and standards was spot on. Uhrig, distrustful of the early forensic DNA analysis labs, was one of the first to question the lack of transparency surrounding the budding science. However he could not succeed on his own, and without the assistance of someone from the scientific community he simply could not formulate the right questions.

Uhrig's failure demonstrates the continuation of what was an integral problem for early DNA analysis. Introducing scientific evidence into the American legal system is intended to be an adversarial process, where new techniques are improved and developed through time and debate between experts (Aronson, 2007). However, this process fails when no scientists are willing to question the dominant paradigm, especially one surrounding such a new technology. Without the scientific community to turn to, and with little public information available regarding Lifecodes' and Cellmark's techniques, defense attorneys were left feeling overwhelmed and abandoned.

Things did not improve for defense attorneys anytime soon. Essentially there was very little debate between 1987 and 1988. DNA was quickly and consistently admitted as evidence in courtrooms. However, defense attorneys have never been known to give up easily. There were certainly still questions in need of answers. Mainly, was forensic DNA analysis actually ready for use in the courtroom?

2.3 George Wesley and Cameron Bailey

This absence of scientific debate surrounding DNA fingerprinting in the courtroom certainly meant that defense attorneys around the country were facing an uphill battle (Clarke, 2007). This can be seen in the next case study I have chosen to discuss, which again ended with DNA evidence being determined admissible. However, this case study is once again important because it demonstrates the unwarrantedly high level of public trust placed in private companies like Lifecodes and Cellmark.

Interestingly, this case study actually involves two different cases, *People of New York v. George Wesley* and *People of New York v. Cameron Bailey* (Aronson, 2007). In the Wesley case, a 79-year-old woman, Helen Kendrick, was found dead in her apartment in Albany (Mahoney, 1992). The defendant, George Wesley was accused of first-degree rape and second-degree murder, along with additional charges of burglary and attempted sodomy (Mahoney, 1992). Although the initial evidence against Wesley was already fairly compelling, the prosecution felt that a DNA print test would provide concrete evidence. After the admissibility hearing, a DNA print test was conducted that compared DNA taken from blood found on the defendant's shirt, sweatpants, and underwear, with DNA taken from the victims hair and the defendants blood (Mahoney, 1992). The test concluded, "that the DNA print pattern on the defendant's T-shirt matched the DNA print pattern from the deceased and that the DNA print pattern from the blood of the defendant was different from that of the decedent," (Smith, 1994).

In the second case, the defendant, Cameron Bailey, was charged with first-degree rape (Mahoney, 1991). The victim had allegedly become pregnant as a result of the attack, and tissue from the aborted fetus provided a possibility for DNA testing (Harris,

1988). Again, an admissibility hearing was necessary. Since the two cases shared many similarities including similar nature of the evidence, similar time frame, and similar location they were combined into one admissibility hearing, which can be simply referred to as the Wesley-Bailey hearing (Aronson, 2007). The hearing began in December of 1987, and continued through the summer of 1988.

The purpose of the hearing was to decide if the DNA-print test offered by Lifecodes met the requirements of the Frye standard, followed by New York courts at the time (Aronson, 2007). If the DNA-print test was deemed admissible as evidence, blood samples could then be taken from the two defendants to be analyzed and used in trial (Harris, 1988). The court operated under the Frye rule, so to determine if the DNA-print test was admissible, the scientific testing in question had to be generally accepted by the relevant scientific community (People v. Middleton, 1981). Michael Baird, the Lifecodes scientist who testified in the Andrews Case, would again be testifying on the side of the prosecution (Aronson, 2007). David Rutnik, the defense attorney working on the case, realized that he would also need to find scientists working outside of Lifecodes in order to stand a chance against Baird, and the prosecution (Aronson, 2007). With the help of a network of defense lawyers, Rutnik was eventually able to track down two scientists, Richard Borowsky and Neville Coleman, who had consulted on earlier scientific admissibility tests, including blood typing analysis (Harris, 1988). Although the two scientists were not directly involved with DNA typing technology, it was the first time that scientists testify as expert witnesses on the side of the defense (Aronson, 2007).

Rutnik's first attempted to discredit the test results by postulating that Baird, as scientist employed by a private company, had something to gain from the outcome of the

trial, and hence was too biased to serve as an expert witness (Aronson, 2007).

Interestingly, the judge saw little problem with the fact that Baird had a financial incentive, as long as the technology could be demonstrated as valid under the Frye standard (Aronson 2007). Rutnik was forced to consider his other options. He eventually attacked DNA fingerprinting by using the two questions as follows:

(1) That even though the theory underlying DNA fingerprinting is valid and generally accepted in the scientific community, Lifecodes' laboratory procedures, methodology, and quality controls are not adequate to assure the reliability and accuracy of its results; and (2) that Lifecodes' population studies are inadequate to establish a claimed power of identity for its results under the laws of population genetics.” (Harris, 1988).

Neville Coleman testified regarding the first issue, arguing both that the science had not been around long enough to generate true peer review, and also that Lifecodes laboratory procedures and methods were not adequate to allow DNA testing as admissible evidence (Harris, 1988). However, the court was quick to note, “Dr. Colman's area of experience is laboratory medicine, and he was held qualified to testify and render an opinion as an expert only with respect to matters of laboratory procedures, laboratory monitoring, and scientific method. He is neither a molecular biologist nor a population geneticist,” (Harris, 1988). Dr. Richard J. Roberts, a leading molecular biologist, assured the court that the technique had been peer reviewed as he had alone had published 98 articles regarding DNA in peer reviewed journals (Harris, 1988). Perhaps most importantly, “Dr. Roberts testified that Lifecodes’ laboratory protocols contained the same quality controls for its procedures that he would have established if he had been setting up Lifecodes.” (Harris, 1988). Dr. Kenneth Kidd, a molecular biologist and geneticist, also testified against Coleman, arguing that he had both reviewed Lifecodes’

published laboratory protocols, and he had also visited the lab and observed the testing process (Harris, 1988). The expert witnesses on the side of the prosecution were found to be much more persuasive than those for the defense.

Borowsky defended the second proposition, arguing that Lifecodes did not attain HWE because the statistical calculations were based on too small of a sample, making it impossible to calculate the probability of a random match (Harris, 1988). Borowsky's argument was not necessarily that Lifecodes' databank demonstrated or did not demonstrate equilibrium but instead that the statistics were based on data far too limited to determine if the populations used by Lifecodes were randomly mating, "at least with respect to the inheritance of particular allele combinations with a specific locus," (Aronson, 2007). Borowsky, like Coleman, was quickly brushed aside. Unsurprisingly, the concerns about the existence of HWE and linkage equilibrium were taken much more seriously in later cases when the expert witness argued strongly against the existence of HWE rather than simply postulating the possibility of insignificant data (Roberts, 1991).

Instead of giving up or moving on to a different plan of attack, Rutnik returned to issues from Coleman's previous testimony. He questioned the expert witnesses for the prosecution about the process and definition of peer review. During a cross examination he thoroughly questioned one expert witness, Richard Roberts. Rutnik's questioning led Roberts to admit on the stand that he had never personally tested the probes used by the Lifecodes laboratory (Aronson, 2007). Roberts even admitted that it was possible that an independent laboratory had never tested the four probes even though the technology was considered peer reviewed (Aronson, 2007). Although these were certainly important points, Roberts responded by explaining he felt he could trust the Lifecodes procedures

because they employed “reputable scientists,” (Aronson, 2007). Unfortunately, the ease of which DNA evidence became admissible in courtrooms may have generated problematic incentives for private companies like Cellmark and Lifecodes. Without little financial competition and even less scientific criticism, these companies had little fiscal incentive to produce quality results.

With incentives in mind, many interesting things can be traced through this case study. First, is the status achieved by Lifecodes. Even though Rutnik succeeded in unearthing accounts of untested laboratory equipment, a finding that should have led to more questions regarding peer review, Lifecodes came through the trial unharmed. Lifecodes’ media generated reputation trumped Rutnik’s concerns, possibly escalating the huge quality control issues that can be seen in later cases. From the outside, it must have seemed as though the metaphoric walls of companies like Lifecodes and Cellmark still held strong. However, on the inside these companies were beginning to cut corners just as defense attorneys were beginning to think about how to attack the foundation itself.

2.4 People v. Castro

On February 5th, 1987, Vilma Ponce was brutally murdered in the Bronx, along with her unborn child and her two-year-old daughter (Patton, 1990). Ponce’s husband reported to the police that he had seen the local handyman, Jose Castro, exiting the area before Ponce entered his house and found his family slain (Aronson, 2007). The police brought Castro in for questioning. Although the interrogation itself revealed nothing, the police noticed a small splatter of blood on Castro’s watch. Castro claimed that the blood was his own but a sample was sent to Lifecodes for analysis nonetheless. A definite

match was reported between the sample taken from the watch and samples taken from the deceased victims. No problems regarding difficulties or ambiguities in the testing were reported.

By late 1989, Joseph Castro was on trial for stabbing Vilma Ponce and her daughter with the DNA evidence against him taking centerfold. However, this actually ended up being the first case with a successful challenge to the admissibility of DNA evidence (Lynch, 2010).

In the pretrial admissibility hearing the defense attorney came prepared, seeking the help of two young attorneys, Barry Scheck and Peter Neufeld (Lynch, 2010). These two attorneys would later become involved with the ever-famous O. J. Simpson case (Lynch, 2010). For now though, they faced a different challenge. Although they were suspicious about the rapid acceptance of DNA typing technology produced by private firms, they still needed to find well-regarded scientists who shared their same concerns. The location of the crime, which occurred in the Bronx, led to a high profile New York admissibility hearing (Lynch, 2010).

Lifecodes once again provided the prosecution with a number of credible expert witnesses, who explained to the court that by using Lifecodes single locus probe method they were able to successfully analyze the sample from Castro's watch and compare it with samples from the two victims (Lynch, 2010). This does not seem too distinct from the cases examined earlier. However the important difference lies in the fact that the defense had finally found a molecular biologist willing to work with them (Aronson, 2007). Eric Lander, a young but brilliant mathematician and molecular biologist provided

the defense with the expertise necessary to analyze Lifecodes results and methodology (Aronson, 2007).

In reviewing the results of the Lifecodes test, Lander found them unreliable at best. The auto radiographic evidence presented showed three lanes comparing the results from the analysis of the blood of the mother and daughter with the analysis of the blood on the watch (Lynch, 2010) Lander believed that the watch lane contained two additional non-matching bands, demonstrating that the test was inaccurate or potentially contaminated (Lynch, 2010). According to Lander, Michael Baird, who was once again testifying for the prosecution, agreed that these non-matching bands were present, however he argued that they were probably produced by some non-human contaminant, meaning that the results were still reliable (Lynch, 2010). His defense of these problematic results may demonstrate a desire to protect Lifecodes credibility and his own finances. However, its also certainly plausible that Baird believed that the results were legitimate. Howard Cooke, a scientist working the Medical Research Council in Edinburgh, Scotland and the man who had actually invented the probe used by Lifecodes to target the DXYS14 locus, eventually testified that without conducting additional experiments, the two extra bands in the watch lane should be seen as evidence that the blood did not match Ponce's (Aronson, 2007).

Spurred on by these inconclusive findings, the defense team continued their search for problematic results. Their next important discovery was that though the Lifecodes final report "stated that both the DNA samples had a 10.25 band at the D2S44 locus, in reality the band in the watch DNA profile was 10.16 kb, while the band in the Ponce's DNA as 10.35 kb. The 10.25 figure turned out to be an average of these two

bands.” (Aronson, 2007). From this, the defense team uncovered a number of methodological errors (Aronson, 2007). Particularly that the 10.16 kb band and the 10.35 kb band differed in size by a standard deviation of 3.06, while the Lifecodes actually employed a rule that two bands must fall between +/- three standard deviations of the average size of the two bands to be considered a forensic match (Aronson, 2007). The defense was then able to demonstrate that if Lifecodes had indeed followed its published methodological standards, they would not have declared the two samples to be a match (Aronson, 2007). Baird was then forced to admit that despite the published methodological standards, Lifecodes relied primarily on visual observation to determine matches (Aronson, 2007).

The defense team continued to uncover additional issues. One in particular involved the sex chromosome test, essentially the test used to demonstrate whether the blood found on Castro’s watch belonged to a male or female. To conduct the sex test, Lifecodes “used a probe that targeted a specific region of the Y-chromosome, called DYZ1” (Aronson, 2007). When Lifecodes analyzed the three samples, they determined that all three were female because none showed a band at this specific locus (Aronson, 2007). The defense argued that this assumption was flawed because the control test also showed no band. Baird argued this, stating that since the control test was derived from the female HeLa (Henrietta Lacks) cell line, a cell line widely duplicated and used in medical research (Aronson, 2007). However, Alan Giusti, the Lifecodes scientist who had also testified in Andrews, contradicted him, stating that the control DNA actually belonged to a male employee, named Arthur Eisenburg, who had a rare genetic condition known as “the short y condition”, where the DYZ1 region is absent, generating a female

result for the specific test (Aronson, 2007). Baird returned to the lab and eventually discovered that the control DNA was from a female Lifecodes technician, Ellie Meade (Aronson, 2007). Although this evidence did not end up detracting as much from Lifecodes' credibility as the earlier discoveries, it still demonstrated that Lifecodes suffered from poor administration, faulty record keeping, and shoddy quality control. This lack of scientific methodology was shocking since Lifecodes had become such a respectable company. However, in hindsight it is not that unreasonable. In the late 80s no one was able to successfully challenge Lifecodes' techniques. Hence the incentive to continue to deliver high quality results was simply nonexistent. As a private company, it made more sense to cut some corners to save time and money.

The questions surrounding statistics and population genetics that had emerged in the Wesley-Bailey case also resurfaced in much more detail during the Castro case (Lynch, 2010). In order to demonstrate the probability of a chance match occurring, it was necessary for the company to specify what reference database it was using. "This is because the alleles marked with DNA probes occur at different rates, on average, in different human groups." (Lynch, 2010). Early databases were sparse and problematic because police offices and blood banks sporadically collected DNA samples (Lynch 2010). In the Castro case, Lander argued that the Hispanic (Castro was Hispanic) database used by Lifecodes was not in Hardy-Weinberg equilibrium (Lynch, 2010). For example, a population demonstrating Hardy-Weinberg equilibrium would exhibit about a 4 percent homozygous rate for any specific allele (Lynch, 2010). The Hispanic database used by Lifecodes exhibited a homozygous rate of "17 percent for the D2S44 locus, and 13 percent for the D17S79 locus," (Aronson, 2007). This possibility of significant non-

random mating within Lifecodes' Hispanic database suggested problems with HWE. However, these numbers may also have been representative of using problematic data for a reference population. Either way, it was clear that additional information was needed to verify the product rule calculations.

As problems regarding the evidence seemed to multiply, the involved expert witnesses actually agreed to hold a meeting sans lawyers, to discuss potential evidentiary problems (Lynch, 2010). Richard Roberts, the scientist questioned by Rutnik in the Wesley-Bailey hearing, also testified for the prosecution in the Castro case. He explained his reasoning behind the meeting as; "we wanted to be able to settle the scientific issues through reasoned argument, to look at the evidence as scientists, not as adversaries," (Lynch, 2010). Although the witnesses for the prosecution, including Roberts, eventually testified that the evidence was sound, the meeting did generate a suggestion to the National Research Council to further investigate these apparent issues (Lynch, 2010).

The outcome of the case was eventually decided by using a three-pronged system, since the court noted difficulties of applying the Frye test to complex scientific evidence (Patton, 1990). The three prongs included:

Prong 1. Is there a theory which is generally accepted in the scientific community, which supports the conclusion that DNA forensic testing can produce reliable results?

Prong 2. Are there techniques or experiments that currently exist that are capable of producing reliable results in DNA identification and which are generally accepted in the scientific community?

Prong 3. Did the testing laboratory perform the accepted scientific techniques in analyzing the forensic samples in this particular case?

(Patton, 1990).

The court found that DNA typing technology fit the two was admissible using the first two prongs (Patton, 1990). However, the court also found that in regard to the third

prong, Lifecodes had not followed the necessary procedures to produce reliable results (Patton, 1990). Overall, the court declared that DNA typing technology in general was admissible under Frye, however in this specific case the evidence provided by Lifecodes for this case was not admissible (Patton, 1990).

Surprisingly, Castro pled guilty even after the DNA evidence had been thrown out (Lynch, 2010). For defense attorneys it marked a huge step forward in learning to work with DNA evidence. Scheck and Neufeld even started holding meetings to educate other on ways to combat DNA evidence (Lynch, 2010). Unlike the ‘independent expert witness’ lists generated by Cellmark, these meetings did not serve the financial motives of any large corporations, although they probably were financially beneficial to the individual defense attorneys. For Lander, and other interested scientists, it unveiled issues with DNA technology that had not yet been explored, mainly issues regarding population substructure and the databases used for reference populations. Publications by Lander (Nature) and Lewontin and Hartl (Science) demonstrated that the scientific community was no longer in total agreement about DNA typing, as it had seemed in earlier years (Lynch, 2010). As for Lifecodes, the Castro case shattered the attitude of good faith that had protected the company from scrutiny during the late 80s, while also demonstrating from a business perspective the importance of delivering a quality product in forensic science. Since both Lifecodes and Cellmark employed extremely talented scientists and technicians it seems plausible that these blunders could have been avoided had the technology been more successfully challenged early on by both scientists and lawyers.

2.5 State v. Schwartz

While the employees of Lifecodes worked to quell the storm of bad press generated by the outcome of the Castro trial, Cellmark faced their own similar problems. Like Lifecodes, Cellmark had breezed through early admissibility trials, perhaps resulting in the mishaps just down the road. In many ways the Schwartz case is simply a less publicized version of the Castro case, as the courts in both cases found DNA typing to be generally admissible but found the evidence presented in the individual cases problematic (Imwinkelried, 1991).

Everything started in May 1988, when nineteen year old student Carrie Coonrod was raped and stabbed to death in downtown Minneapolis (Kaye, 2010). Eyewitnesses aided the police in tracking down Thomas Schwartz, who was found in the vicinity of the crime and was indicted for stabbing Coonrod to death (Kaye, 2010). Assistant Public Defender Patrick Sullivan was appointed to the case (Aronson, 2007).

The physical evidence against Schwartz was strong even without the use of DNA testing. A pair of bloodstained pants were found in Schwartz' apartment, and a bloodstained shirt allegedly belonging to Schwartz was found near the scene of the crime (Kaye, 2010). A Cellmark run state laboratory reported that “[a]ll bands in the DNA banding pattern obtained from the blood of Carrie Coonrod are contained in the DNA banding pattern obtained from the stain removed from the plaid shirt.” (Kaye, 2010). Cellmark reported that the frequency of this DNA-banding pattern in a Caucasian population was approximately 1 in 33 billion (Kaye, 2010). However this same report also concluded “no definitive conclusion can be reached,” (Kaye, 2010).

To challenge the Cellmark's results, Sullivan started working with multiple molecular biologists and lawyers, including Simon Ford and William C. Thompson (Aronson, 2007). Together they sorted through Jeffrey's original papers in order to better understand Cellmark's single locus probe procedures (Aronson, 2007). Eventually they determined that Cellmark was making claims about its procedures that they did not believe it could support (Aronson, 2007).

The pretrial alone contained testimony from 12 expert witnesses from both the defense and the prosecution (Kaye, 2010). During the pretrial, Sullivan argued that Cellmark was restricting access to their databases, making it impossible for the defense to determine if the databases being used were in Hardy Weinberg equilibrium or not. Sullivan even stated that the company did this to "protect itself from embarrassment and gain a competitive advantage," (Aronson, 2007). Sullivan also demonstrated to the court that Cellmark "had made several errors in a blind proficiency test carried out in late 1987 and early 1988 by the California Association of Crime Laboratory Directors." (Aronson, 2007). Specifically, Cellmark had identified two samples as coming from the same sample, when they actually did not (Kaye, 2010). This demonstrates that even if the population database used by Cellmark to construct the chance of a random match were in Hardy Weinberg equilibrium, the human error present would make the statistic inconclusive. However the prosecution also involved convincing testimony from a number of credible scientists. The trial judge eventually determined that the report was admissible, however the judge also asked the appellate courts to look over this decision before continuing to the trial itself (Kaye, 2010).

During the actual trial, Sullivan argued that although DNA testing should be generally admissible, the results in this particular case should not. The court was troubled by Cellmark's deficient testing procedures and errors found during the CACLD blind trial. The Minnesota Supreme Court determined the particular evidence inadmissible, meaning that until further notice, all DNA evidence provided by Cellmark was inadmissible (Aronson, 2007). The state continued with the prosecution of Schwartz without the DNA evidence (Aronson, 2007).

The outcome of this case is very similar to the outcome of the Castro case. These large private biotech companies were finally starting to be held responsible for poor quality control and management. However, it certainly must be mentioned that standards of quality were simply nonexistent at the time when companies like Lifecodes and Cellmark took on DNA profiling (Aronson, 2007). Hence, the more important message to take from these case studies is that defense attorneys and hardworking scientists were truly making a contribution in the evolution of DNA profiling. The problems highlighted by Lifecodes and Cellmark were important problems that needed to be addressed by the entire scientific community, rather than the two companies alone.

At this point it seems relevant to wonder how Lifecodes and Cellmark so easily dodged the financial concerns raised by many attorneys in earlier cases. In many instances the defense attorneys would attack an individual expert witness, arguing that the evidence should be thrown out due to personal bias or financial gain. However these attempts were never very successful, which is slightly concerning considering the problems that eventually came to light. Lifecodes and Cellmarks' lack of organization and quality assurance is certainly a product of the time in which the companies both

emerged. However, it also seems likely that financial incentives may have played a role in promoting the lax environment.

2.6 State v. Woodall

The Castro case certainly demonstrated that the attitudes surrounding DNA typing technology were starting to shift. Defense attorneys were figuring out ways to question DNA based evidence, while also demonstrating that companies like Lifecodes and Cellmark needed to be held responsible for poor evidentiary standards. Around this same time, State v. Woodall became the first case in West Virginia where the West Virginia Supreme Court ruled on the admissibility of DNA evidence, deeming it admissible (Connor, 1996). However, in this case, the defense, rather than the prosecution employed the evidence.

In 1987 in Huntington, West Virginia, two women were separately abducted at knifepoint and repeatedly raped (Connor, 1996). The victims were forced to close their eyes but both women caught brief glimpses of the perpetrator (Connor, 1996). The first woman noted that the man wore brown pants and was uncircumcised (Connor, 1996). The second woman also recounted that the perpetrator was uncircumcised (Connor, 1996). She also informed the police about the perpetrators hair color as well as the style of boots he was wearing (Connor, 1996).

Glen Woodall was arrested, and eventually found guilty of sexual assault, sexual abuse, kidnapping, and aggravated robbery by a jury in 1987 (Connor, 1996) The evidence used by the prosecution included the two eyewitness accounts, as well as comparison of Woodall's hair with hair found in one of the victim's car, and a state

police chemist's testimony that blood secretions from Woodall matched semen secretions held as evidence (Connor, 1996). One of the victims also identified clothing found in the defendant's house, saying that it matched the clothing the perpetrator had been wearing at the scene of the crime (Connor, 1996). In addition, the Woodall, like the perpetrator, was uncircumcised (Connor, 1996). One rather bizarre piece of the evidence was that both women noted a distinctive odor coming from the perpetrator, which allegedly matched the smell of the defendant's workplace (Connor, 1996). Woodall was sentenced to two life terms without parole, to be served consecutively with 203-335 years in prison (Connor, 1996).

Interestingly, in the pre-trial of the case the judge denied the defense attorney's request to use DNA fingerprinting technology to compare the defendant's blood sample with semen samples collected from the victim's clothing (Connor, 1996). The judge based the denial on the fact that the technology was "experimental", even though DNA typing technology had already been deemed admissible by various courts in the United States, as discussed in the Andrews Case. The main reason for the courts' rejection of the experimental technology lay mainly in the fact that the defense attorney was not able to provide any expert witnesses who could testify about the reliability of the test (Connor, 1996). It is certainly possible that this was an issue of location, as the prosecutors who achieved success with DNA technology were generally in larger cities, such as New York. However, it is interesting that companies like Lifecodes and Cellmark seem to have entirely overlooked the possibility that DNA evidence may also be used by the defense. Lists about expert witnesses willing to testify about Lifecodes and Cellmarks' technology had been distributed and shared among prosecutors but rarely among defense attorneys.

After the trial, a DNA test was conducted but the court found the results inconclusive (Connor, 1996).

On July 6th, 1989 the West Virginia Supreme Court of Appeals affirmed Woodall's conviction (Connor, 1996). However, Woodall continued to file various motions, including several appeal petitions and habeas corpus petitions, requesting DNA testing of the evidence. Eventually the West Virginia Supreme Court released the evidence to the defense and the evidence was sent to the FSA, Forensic Science Associates. The FSA concluded using PCR based testing, that although the samples from the two victims matched, they did not match Woodall's sample (Connor, 1996). The results were even reviewed by Dr. Alec Jeffreys himself (Connor, 1996). The test was also conducted again by the Center for Blood Research (CBR) and the same results were reached (Connor, 1996). On July 15th, 1991, the trial court vacated Woodall's conviction. Woodall was set on electronic home monitoring while further analysis, including RFLP testing, was conducted by CBR. The state of West Virginia also conducted its own DNA test. All results demonstrated that Woodall could not have been the perpetrator. After this additional testing, "West Virginia moved to dismiss Woodall's indictment on May 4th, 1992, and the trial court granted the motion," (Connor, 1996). By this time Woodall had spent four years of his life in prison and one year under home confinement (Connor, 1996). He was awarded \$1 million because of his wrongful conviction and imprisonment (Cannon, 1996).

This case is interesting because it demonstrates a change in overall perceptions of DNA technology following the Castro and Schwartz cases. Had Woodall's defense found a scientist willing to serve as an expert witness from the beginning of the case, it

seems very likely that the evidence would have been accepted. Just as DNA was moving into the mainstream in the early 90s, more lawyers and scientists were voicing serious concerns about DNA profiling technology. In this case, three separate laboratories ran tests on the evidence before declaring Woodall's innocence, a rather stark comparison from earlier cases where testing was done by exclusively by one private company. Certainly, Woodall would have preferred a faster path to freedom. However, the repeated testing and slow acceptance of the evidence demonstrates how the mindset surrounding DNA typing technology was becoming increasingly critical. The days of unwavering faith in DNA profiling were coming to an end.

Chapter 3: Public Science and the 'DNA Wars'

“Science is the search for the truth—it is not a game in which one tries to beat his opponent, to do harm to others,” Linus Pauling (Butler, 2009).

3.1 Change, Change, Change

Together, these cases demonstrate the many difficulties involved with the introduction of any new science into the courtroom setting. And though the focus of this paper is mainly legal and scientific, it is important to note that I in no way condone the horrible crimes committed in the previously discussed cases. However, from a broad and certainly disconnected perspective, one could argue that there was more at stake in these cases than the futures of the criminals themselves.

After Castro and Schwartz, negative press engulfed Cellmark and Lifecodes. These companies had taken a huge public relations hit, and were now being faced with more requests from defense attorneys for all of the raw data associated with specific cases,

to be analyzed by outside scientists (Aronson, 2007). At the same time these companies were finally being inspected by state run forensic science laboratories and communities (Aronson, 2007). The New York State Forensic DNA Analysis Panel, a committee created by the New York governor himself to investigate the admissibility of DNA evidence, determined that the claims of extreme accuracy, such as stating the probability of error as 1 in 1 million, were suspect because outside scientists had not examined them (Aronson, 2007). Issues with these companies' data were turning up all around the country. A prosecutor in Maine, after losing a case due to a mislabeled autorad, wrote to Lifecodes not to send any more bills because, "the state will not suffer further by paying for Lifecodes incompetence." (Kaye, 2010).

As these issues continued to arise it was becoming more and more possible for defense attorneys to contest DNA based evidence. Judges were now examining what could go wrong with DNA typing technology, rather than simply deciding if the technique was generally accepted or not (Aronson, 2007). These questions, generally asked by defense attorneys and scientists, started to create controversy even outside of the courtroom, among the molecular biology community itself. Among the legal community, questions regarding mundane procedures and standards were demonstrating the need for the implementation of some sort of quality assurance (Lynch, 2010).

One of the first major changes following Castro and Schwartz was the introduction of FBI involvement in forensic DNA analysis. To implement this, the FBI worked to standardize procedures across the board. Essentially, the goal of the FBI was to create a web of crime laboratories and scientists who all followed the same set of rules regarding procedures and quality standards (Aronson, 2007). However, this was not a

huge of an improvement over the lack of regulation that occurred with the private companies. The FBI claimed that only people working within the forensic world could understand the complexities involved with DNA sample in a forensic context (Aronson, 2007). Problematically, the forensic scientists chosen by the FBI to spearhead the standardization often had little or no experience with molecular biology (Aronson, 2007). Essentially, the FBI believed that “only other member of the forensic community could determine the best way to conduct proficiency testing, individual credentialing, and laboratory accreditation,” (Aronson, 2007). By the end of 1988, the FBI had created an advisory group to help develop a standard set of probes and procedures that could be put into place in all laboratories that working with DNA based evidence (Kaye, 2010). However, even standardization of the probes and technical aspects of the procedures could not secure quality interpretations and results (Aronson, 2007). By December 1988, the FBI was offering forensic DNA analysis to law enforcement agencies. The FBI began a forensic DNA database by 1990.

FBI scientist Bruce Budowle, who was also a geneticist, visited a number of different laboratories and collected information about the current trends in DNA typing technology (Aronson, 2007). After reporting his findings, the FBI Forensic Science Research and Training Center (FSRTC) attempted to put together a research team that could take this information and use it to conduct experiments and improve the efficiency and reliability of forensic DNA typing (Aronson, 2007). They then planned to simplify their findings and send them out to state and local crime labs, making forensic DNA analysis practical and reliable. To do this, FBI decided to train various forensic technicians, many with no background in molecular science, a process that included a

forty-four hour course on molecular biology (Aronson, 2007). The process was completed with a test where they performed analysis on 122 cases from previously submitted evidence (Aronson, 2007). The FBI determined that the results of this test supported that the FBI laboratory protocol was efficient and reliable, regardless of the fact that these technicians were tested immediately after their intensive training (Aronson, 2007). Interestingly, the FBI did not publish any results in peer-reviewed journals until more than a year after they began testing courtroom evidence in December, 1988 (Aronson, 2007). This implementation of standardization, although certainly good-intentioned, demonstrated problems that were actually very similar to those seen with the private companies, including a general lack of transparency and unwillingness to work with other disciplines.

FBI involvement in DNA testing meant that the FBI was now open to the same scrutiny from defense attorneys that private companies like Cellmark and Lifecodes had faced in earlier years (Kaye, 2010). Not surprisingly, defense attorneys once again found it difficult to argue against DNA evidence, especially now that people believed the process had been improved through FBI involvement. FBI involvement generated kind of an equal but opposite problem. In the beginning of the standardization the FBI was almost as nontransparent as the private businesses before them. In addition, some scientists and defense attorneys believed that not enough research had been conducted to solve the issues with the reference population databanks and the population substructure debate. In some ways defense attorneys were right back where they started.

3.2 The Jakobetz Case

The Jakobetz case, which occurred in 1989, helps analyze the arrival of FBI initiated standardization. (Kaye, 2010). The crime occurred when a young woman from Burlington, Vermont, stopped in at a rest area to make a phone call on her way down Interstate 91 (Kaye, 2010). While at the rest area she was attacked, handcuffed, and her head was covered with a pillowcase (Kaye, 2010). She was then driven away from the area and brutally raped and assaulted. She was eventually abandoned on the side of the road near the Bronx in New York (Kaye, 2010). There she was discovered and taken to a hospital. A semen sample taken from a vaginal swab was sent to an FBI laboratory for analysis.

Again, evidence seemed to point to Randolph Jakobetz even without DNA testing. A call placed on the same pay phone used by the victim at around the same time was traced back to P. Zanon, who was married to a man named Randolph Jakobetz (Kaye, 2010). Jakobetz's schedule, he worked for a truck driving company, placed him in the general area where the crime was committed (Kaye, 2010). Also, a search of Jakobetz' trailer generated hairs that were similar to the victim's own pubic and head hairs (Kaye, 2010). The victim was also able to positively identify Jakobetz from a photo spread (Kaye, 2010). Jakobetz was charged with kidnapping, a federal crime since he had taken her across state lines (Kaye, 2010).

The DNA evidence further confirmed that the semen sample matched Jakobetz' DNA, with a "one in 300 million chance that the DNA from the semen sample could have come from someone in the Caucasian population other than Jakobetz," (Kaye, 2010). As we now know, this report is slightly flawed because the product rule only generates

the correct probability when assuming no recent shared relations, which is not the same statement as ‘someone in the Caucasian population other than Jakobetz’ (Kaye, 2010).

Although the defense did not bring up this particular issue, they used a number of other tactics to argue that the DNA evidence was inadmissible (Kaye, 2010). Five expert witnesses testified for the prosecution and four testified for the defense over an eight-day hearing. This court had abandoned Frye in favor of the relevancy test so there was no need to determine that the technique was generally accepted by the scientific community (Kaye, 2010).

The DNA testing technique used in the case was gel electrophoresis of VTNR alleles (Kaye, 2010). Using this messy technique, even identical copies of a DNA fragment that are run on the same gel may not produce the exact same results (Kaye, 2010). To deal with this possibility of measurement error, FBI laboratories declared that two fragments matched if they fell within a certain distance from each other, called a match window, the window being five percent (Kaye, 2010).

In the Jakobetz trial, an expert witness for the defense, Joseph Nadeua, argued that the amount for the match window was not consistent with statistical practice (Kaye, 2010). Nadeua believed that the actual size of the window should be about half of what the FBI was following at the time, an assessment supported by other statistician at the time. However, in this particular case Nadeua had chosen to fight the wrong battle. In Jakobetz, all sixteen band matches were within plus or minus 1 percent, making his argument irrelevant (Kaye, 2010). On top of this, the prosecution argued that the FBI’s match window of 5% was reasonable because it kept the chances of both false inclusions and false exclusions as small as possible at the same time (Kaye, 2010).

A more significant issue in the Jakobetz case had to do with the database the FBI used to calculate the random match probability. In Jakobetz, the FBI used a Caucasian database that was generated by taking DNA samples from 225 recruits (Kaye, 2010). The FBI then analyzed these samples “to obtain a list of measurements for the lengths of each allele at each VNTR locus,” (Kaye, 2010). The defense argued that this was not an appropriate sample to use for an entire Caucasian population. Kenneth Kidd, a leading population geneticist, countered this attack by arguing that the sample size could be as small as 100 as long as it was previously determined that the alleles occurred randomly throughout the targeted population (Kaye, 2010). Kidd was backed up by Budowle, the FBI geneticist, and geneticist Thomas Caskey, who both testified that the FBI had used ‘conservative’ adjustments to deal with these issues (Kaye, 2010). Still, the defense remained doubtful of whether it had actually been previously determined that the alleles occurred randomly in the population of recruits.

“The last step in estimating a VNTR frequency is to combine the various allele frequencies into a single number—for example, the one in 300 million frequency for the occurrence of the VNTRs in Jakobetz” (Kaye, 2010). Hence, the next logical argument for the defense was to question the calculation of this single number. In Jakobetz, the expert witnesses testifying for the defense included Richard Lewontin, a geneticist, biologist, and social commentator who found issue with the overarching racial groups used in population genetics. Lewontin argued that because people tend to marry people who share a similar religion, culture, and geographic location in addition to similar race, VNTRs would not occur randomly throughout a broadly defined population, such as Caucasian (Kaye, 2010). However, the court eventually sided with Kidd, who argued

that the FBI had overcompensated for this possibility, and even so his own research showed that subgroups tended to have very small variations in allele frequency (Kaye, 2010). However, at the time neither Kidd nor Budowle had published studies to support these conclusions (Kaye, 2010). Eventually the court concluded that the FBI followed established protocols and standards, and the evidence was deemed admissible.

This case demonstrates that the move to standardization had left some people in the dust. Defense attorneys were once again subject to losing court cases, and many scientists still seriously questioned the methods practiced by these standardized protocol labs. It seemed that the adversarial nature of the legal system was once again at standstill. The FBI demonstrated the same problems of impenetrability and non-transparency in the courtroom that had previously been seen with Cellmark and Lifecodes.

3.3 United States v. Yee et al.

The case of United States v. Yee et al. generated an explosion of arguments surrounding the population substructure debate. In February 1988, three members of the Hell's Angels motorcycle gang located in Cleveland, Ohio were accused of murdering a clerk who they believed to be associated with a rival gang (Kaye, 2010). David Hartlaub, the victim, had been shot fourteen times in his van by a silenced machine gun before his body was thrown from the vehicle (Aronson, 2007). His bloodied and abandoned van was found near the scene of the crime (Aronson, 2007). Apparently, the members of the Hell's Angels had mistaken his van for one driven by a rival gang, the Sandusky Outlaws (Aronson, 2007). Blood was found inside the van, and VNTR testing conducted by the FBI determined that blood found belonged to Hartlaub as well one other individual (Kaye,

2010). The blood matched Hell's Angels member Johnny Ray Bonds, who at the time of arrest had a serious arm wound, believed to have been caused by bullet ricochet (Aronson, 2007). Blood was also found in Bonds car, demonstrating that he was likely the shooter (Aronson, 2007). The reported random-match probability was 1 in 35,000. However the original reported probability was 1 in 270,000, demonstrating that the FBI may have been trying to avoid any major critiques during the trial (Aronson, 2007).

During the pretrial, the Magistrate Judge James Carr conducted an intensive six-week hearing to determine admissibility (Kaye, 2010). Due to the high-prolific nature of the Hell's Angels, along with the recent changes in forensic DNA analysis, the hearing was even referred to as the "Ultimate Showdown," for the reliability of DNA evidence (Aronson, 2007). Once again the defense team included the veterans Scheck and Neufeld (Kaye, 2010). By the end of the brutal two-year process, both the prosecution and defense had accused the other side of having a "win at all costs mindset," (Aronson, 2007).

To attack the evidence, Scheck and Neufeld brought up an interesting point that highlighted problems with both the private DNA testing companies and the new FBI regulated testing. They believed that for DNA evidence to be admissible it had to be peer reviewed from many angles and disciplines, including alternating views from both the forensic, molecular biology, and genetics communities (Aronson, 2007). Essentially, they believed that not all of the relevant scientific communities had been offered a chance to evaluate the new FBI standards, making the approach not generally accepted. One of these 'left out' scientists, Peter D'Eustachio, a professor of biochemistry, testified that the FBI labs did not follow the procedures that they had set out in the implementation

guidelines (Aronson, 2007). He argued that the lack of widespread peer review lead to flawed conclusions and problematic methodology (Aronson, 2007).

The second part of the attack, argued predominately by Lewontin and Hartl, was once again the issue of population substructure in relation to random match probabilities. The technique employed by the FBI assumed that “all major ethnicities with a given race, such as Swedes, Norwegians, Irish, Jewish, and Italian in the Caucasian racial group, tend to intermarry and have gene frequencies that do not diverge significantly from the group average,” (Aronson, 2007). Lewontin and Hartl claimed that technique was incorrect for several reasons. First, they argued that not enough data had been collected to determine if Caucasians subgroups have different or similar allele frequencies. Next, they determined that these various subgroups had only had a few generations to mix together, not enough time allow these gene frequencies to average out (Aronson, 2007). Last, they employed anthropological and sociological evidence to demonstrate that even among one race, people tend to marry people with similar religions, geographic locations, culture, and language, essentially furthering these sub-groups that may have more or less variation in gene frequencies. Essentially, these calculations did not “correct for the possibility that the donors of the two DNA profiles being compared belonged to a subtype more genetically similar than the Caucasian population as a whole,” (Aronson, 2007). Hence, they believed the random-match probabilities were potentially overstated, and misleading to the jury. Lewontin and Hartl also attacked the statistical tests used by the FBI to determine HWE, with Hartl stating that the tests, “are virtually useless as indicators of population substructure because, even for large genetic differences between

subgroups, the resulting deviations from HWE are generally so small as to be undetectable by statistical tests,” (Aronson, 2007).

The prosecution testified back, including Kenneth Kidd, replied that the database used by the FBI was actually in HWE, as American Caucasians tend to marry outside their neighborhoods and outside religious lines (Aronson, 2007). The defense team then used that straightforward argument that even slight deviations from the equilibrium in the database would not make the evidence invalid. The FBI approach to calculations included conservative measures that would more than compensate for population substructure (Kaye, 2010). Kidd, along with other expert witnesses, argued that the phenomenon of population substructure was simply not relevant to forensic DNA analysis, which by nature tended to be much more practical than theoretical (Aronson, 2007).

Interestingly, now that the involvement of private companies had diminished, the defense could still comment on issues of financial motive. Scheck and Neufeld critiqued the expert witnesses testifying for the prosecution, most of whom who were involved with laboratories awaiting pending government grants for the humane genome project (Aronson, 2007). Scheck and Neufeld argued that financial motives and career opportunities kept many of these scientists from closely critiquing the FBI methods. Thomas Caskey, an expert witness for the prosecution, had a huge financial stake in the FBI forensic DNA project because he had adopted the FBI protocol extensively in his own laboratory (Aronson, 2007). Controversy regarding the FBI protocols and methods could lead to changes that could devastate Caskey’s lab. Interestingly, the judge actually believed that Caskey’s adoption of the FBI methods made him more credible as an expert

witness, as he must certainly believe very strongly in a specific protocol in order to adopt it (Aronson, 2007). Despite Scheck and Neufeld's critique, the Judge found Caskey to be a credible witness.

Judge Carr eventually determined that for the particular case, Caskey and the other witnesses for the prosecution represented the views of the relevant scientific population more than Lewontin, Hartl, and other witnesses. The evidence was deemed admissible and the three men involved were found guilty. However, the defense had certainly succeeded in undermining the FBI laboratory procedures (Kaye, 2010). The case may have been lost, but the 'war' was still yet to come.

The interesting aspects of this case lie not only in the scientific questions that arose, but also in the way the United States legal system treats science. The adversarial nature of the courtroom had created bitter rivalries and feuds, rather than fostering the openness and cooperation that so often fosters scientific advancements. Eric Lander, the scientist who worked in the Castro case, was actually so offended by the whole system that he refused to serve as an expert witness again, appearing only as a neutral court witness in Yee (Aronson, 2007). The abuse faced by expert witnesses in the courtroom seemed to echo the frustration felt by defense attorneys, scientists, and even bureau workers who felt unsatisfied with the current state of affairs.

Perhaps most importantly, the outcome of this case also demonstrates a cyclical nature of early DNA profiling. Patterns of reputation trumping science that can be seen in Cellmark and Lidecodes cases were practically replicated with FBI involvement. Concerns raised by defense attorneys were once again brushed aside because of the FBI's trusted status rather than significant scientific data.

3.4 Population Substructure Debate and the NRC I

The early 90s marked the beginning of the so-called ‘DNA Wars’. “Differences of scientific opinion and legal maneuvering intensified to the point where commentators began to speak of “war” and ‘free fire zones,’” (Kaye, 1997). From one perspective, the escalation of the debate could demonstrate more of a legal dispute, with ‘civil libertarians’ on one side, and ‘law and order’ types on the other (Lempert, 1997). However the role of science in the debate, particularly involving issues of population substructure, should not be understated. And despite the media representations of an often snarky conflict, there certainly was a shared end goal to preserve the integrity of DNA analysis rather than to diminish from it. As explained by Richard Lempert, “among the ranks of those who have questioned aspects of the forensic use of DNA evidence, there is no one I know who takes pleasure in seeing the guilty go free (Lempert, 1997).

The Yee case demonstrated to the defense expert witnesses that there was little left to be accomplished within the legal realm. Certainly, the legal realm had served as an arena for disputing new ideas and concepts, but bringing them into practice using testing and analysis would take a more scientific approach. Lewontin and Hartl, disillusioned with their own scientific community as well as the legal system, decided that there was only one thing left to do. They published. In this article they relived their time in the courtroom, referring to it as “their losing courtroom criticism,” (Kaye, 2010).

Overall, the article argued that subgroups demonstrate nonrandom mating among the genotypes, “even though no individual selects a partner based on his or her genes,” (Lewontin and Hartl, 1991). While people are not persistently selecting mates based on

genes alone, “when people chose their mates endogamously, they are unconsciously making a choice among blood groups and other traits correlated with ethnicity,” (Lewontin and Hartl, 1991). They supported this conclusion with the argument that “there is, on average, one-third more genetic variation among Irish, Spanish, Slavs, Swedes, and other subpopulations, than there is on average, between Europeans, Asians, Africans, Americans, and Oceanians,” (Lewontin and Hartl, 1991). The article stated that probability matches could not be determined conclusively using databases made up of general groupings like Caucasian, Black, or Hispanic, at least without further data (Kaye, 2010). Even worse, because the magnitude and direction of the error depend on the specific VNTR locus, the bias cannot be assigned a direction (Lewontin and Hartl, 1991). Until more data was collected, which Lewontin claimed could take 10 to 15 years, they concluded that the probability statements connected to forensic DNA evidence should be deemed inadmissible in court (Roberts, 1991). Lewontin and Hartl also stressed that statistical tests for HWE were essentially useless as indicators of population substructure within subgroups, claiming “the proper approach is the straightforward one of sampling the individual subgroups and examining the differences in the genotype frequencies among them,” (Lewontin and Hartl, 1991).

The publication process itself was a disaster. The editors of *Science* originally refused to publish the article without serious revisions (Kaye, 2010). The full article was eventually published alongside dissenting perspective pieces written by other population geneticists (Kaye, 2010). On the other side of the issue, proponents for forensic DNA analysis, including the FBI as well as accredited scientists like Kenneth Kidd and Thomas Caskey argued once again that their approximations were close enough for forensic issues,

and only required more intense scrutiny when used in scholarly studies of population genetics or molecular biology (Roberts, 1991). Kidd put it persuasively when he explained, “It makes no difference to me if the number is 1 in 800,000 or 1 in 5 million,” later adding that he did not believe it would matter much to a jury either (Roberts, 1991). Ranajit Chakraborty and Kidd published their own rebuttal to the article, which also appeared in *Science* (Roberts, 1991).

After these publications, the controversy exploded into an all out brawl focused more on delivering personal attacks than on uncovering the truth about a possible phenomenon. Hartl claimed that Wooley, the prosecutor in *Yee*, had threatened him over the phone about the publication, while Kidd personally attacked Lewontin by suggesting he was misrepresenting science to support his own liberal philosophies and political motivations (Kaye, 2010). The debate had certainly become mean-spirited and divisive. Media coverage of the controversy sported diction like battle and attack, leading to the eventual coinage of the controversy as the ‘DNA wars,’ (Aronson, 2007).

Due to the increased controversy and decreased scientific progress, courts struggled with DNA admissibility decisions that would have been easily determined just a year ago. In response to this, the National Academy of Science’s National Research Council (NRC) initiated a study to further examine and solve the problems and limitations associated with forensic DNA technology (Coleman and Swenson, 1995). Victor Mckusick, a dean of American genetics, was picked to head the 14-person council (Coleman and Swenson, 1995). Again, the main issue at hand involved random match probability calculations (Coleman and Swenson, 1995). However, although the board was made up of a multitude of extremely distinguished scientists, only two of the

members of the original board, Eric Lander and Mary Claire-King, were at all experts in molecular biology and population genetics (Kaye, 2010). There were no expert statisticians on the committee (Coleman and Swenson, 1995). The purpose of the report was to come to some sort of compromise between both sides that would also advance the use of forensic DNA analysis back into the courtroom (Kaye, 2010).

The NRC report, entitled *DNA technology in Forensic Science*, came out in April 1992 (Coleman and Swenson, 1995). Lander, who was an early critic of forensic DNA technology, was beginning to doubt the forensic importance of the population substructure due to new statistical studies that supported the original HWE based calculations (Kaye, 2010). However, because the committee could not actually disprove the possibility of the existence of population substructure, they attempted to work around it. The first NRC report argued that this error should burden the state, rather than the defendant. To do this, the committee proposed two methods for calculating random match probabilities, one called the ceiling principle and the other the interim ceiling principle (Kaye, 2010). The ceiling principle involved calculating about a dozen subgroup population samples, and then calculating the ‘ceiling frequency’, the largest frequency of an allele found in any of the samples (Kaye, 2010). The allele ceiling frequencies would then be multiplied to provide an estimate of an upper bound for the genotype frequency (Kaye, 2010). However data for this rule was not available, so a different rule, the interim ceiling principle, was suggested instead. The interim ceiling principle used the highest allele frequencies found in all the major population group databases (white, African American, Hispanic, Asian, native American) to calculate a possible ceiling (Kaye, 2010). If the calculation was less than 10%, an overall minimum

of 10% was to be used instead. The interim ceiling principle received major criticism for being overly conservative and statistically illogical (Balding and Nichols, 1994). “The ceiling principle is unsatisfactory in taking into account populations which have no possible connection with a given crime,” (Balding and Nichols, 1994). Mostly, the report was driven more by policy than by science, and ended up recommending a method for compromise that simply ignored the relevant scientific issues at hand.

The report, while recommending the continued use of forensic DNA analysis in the courts, also called for a number of more generalized changes including increased standardization, mandatory accreditation, and proficiency testing for labs (Coleman and Swenson, 1995). Additionally, the report asked Congress to commission a panel of independent scientists, lawyers, and ethicists to monitor technology developments surrounding DNA analysis (Kaye, 2010). The report also prompted courts to fund expert witnesses for the defense and allow defense attorneys access to all relevant reports and data used in the case (Kaye, 2010). If the report had been more of a success scientifically, it is possible that these well-intentioned measures would have gone into effect, potentially helping to prevent issues of laboratory error that emerged in later cases. However, the NRC I was seen as a failed attempt from almost every perspective community.

Within the legal realm, the NRC report caused additional confusion rather than clarity. Courts across the country were unsure if the report simply demonstrated scientific disagreement, or if it demonstrated that the HWE calculations still being used in many cases were entirely invalid (Kaye, 2010). Following the report, random-match probabilities were deemed inadmissible in two kidnapping cases in California, *People v.*

Barney and *People v. Howard* (Kaye, 2010). In an Arizona State Supreme Court Case, *State v. Bible*, the court determined that although the technical methods for producing the match results were admissible, the reporting of the random match was not (Kaye, 2010). Now prosecutors had to deal with appeals from previous cases, also convincing courts in current cases not to throw out DNA evidence that could place a murderer or serial rapist behind bars (Kaye, 2010). Prosecutors and law enforcement personnel felt that the NRC I did more to exclude DNA evidence than help solidify it (Kaye, 1997).

From the scientific arena, the report generated a number of raised eyebrows. Outspoken individuals quipped about the irrationality of the new rule. Overall, most scientists seemed to think the problem was that the interim rule was attempting to compensate for a bias without knowing anything about the size of the bias itself (Kaye, 2010). However, despite these statistical complaints, even scientists like Lewontin and Hartl, admitted that the rule may often favor the defendant. Claire King, one of the original NRC committee members even agreed, stating that the rule was “a really unbelievable conservative approach, very, very pro-defense approach,” (Kaye, 2010). Budowle and the FBI also rejected the report, claiming the methods were too conservative (Aronson, 2007).

Interestingly, defense attorneys like Neufeld and Scheck also expressed disappointment with the report. They claimed that the board, which had initially included Caskey, was biased toward the prosecution as well as financially motivated (Coleman and Swenson, 1995). This seems unlikely, since most people involved agreed that the ceiling rules were biased toward the defense. It seems more plausible that defense attorneys rejected the report simply because it was easier to win cases when the scientific

community was in controversy, rather than accepting a compromise (Kaye, 2010). Either way, the NRC's attempt to meld science, law, regulation, and technological standardization into one report was scarcely more than a well-intentioned flop.

Chapter 4: With Truces Come New Troubles

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

4.1: NCR II

After the failure of the first NRC report, scientists, lawyers, and the bureau alike were beginning to feel concerned that DNA typing, a gem in the hands of the criminal justice system, would start to lose its power in court. In June of 1993, The NRC created a planning committee, lead by chairman James Crow, to reevaluate forensic DNA analysis, this time with a specific focus on population substructure and statistics (Aronson, 2007). This time the committee was assigned with only one task, “identify the key major issues that must be addressed to resolve quickly the current debate over statistical evaluation of DNA evidence.” (Kaye, 2010). However this did not mean that the board was focused only on population substructure, but instead included all of “the statistical and population genetics issues in the use of DNA evidence,” (Kaye, 2010). This is probably why the second report, which was originally supposed to take only six months, was not released until May 1996. The committee made recommendations in three areas, laboratory procedures, statistical assessments of laboratory results, and social science research (Kaye, 1997). The main finding included in the report was a collection

of empirical data analyzed by the FBI that demonstrated that the population substructure phenomenon was less significant than previously believed.

The report included recommended best practice for determining random match probabilities under various circumstances. The report put forth alternative procedures to the ceiling product rule, which emphasized creating a distinction between populations and subpopulations as follows (Kaye, 2010). The committee reported that when the race of the person is known, an existing database for that race should be used for calculations. However if race is unknown, “calculations for all the racial groups to which possible suspects belong should be made”, employing the basic product rule (Kaye, 2010). If the individual belonged to a specific subgroup, such as a remote village, calculations were to be made using the basic product rule, with the allele frequencies for the involved subgroup. If data was unavailable, as it often was, the committee recommended “adjusting the allele frequencies from the population containing the subgroup according to the population structure equations and then multiplying them according to the product rule,” (Kaye, 2010).

The responses to the NRC II were critical, but less offensively so than responses to the first committee. One of the main reasons for the minimal outcry is the report came out so much later than was originally expected. Many of the statements made in the NRC II were expected. Some scientists and lawyers critiqued small aspects of the report, such as Thompson, who stressed that the report should have contained demands for external regulation of forensic laboratories (Thompson, 1997). Scientists like Koehler and Lempert agreed, arguing that information regarding results of blind proficiency testing for individual labs be reported to juries to help weigh the possibility of errors and false

matches in a case specific fashion (Lempert, 1997). Some scientists critiqued the statistics, including Balding and Morton who suggested some new ideas for calculating random-match probabilities. However, they both admitted that the debate was certainly dwindling down, as it involved mainly small technical issues (Morton, 1997). Even Lewontin claimed that population genetics was no longer at the center of the debate.

When compared to the drama surrounding the initial population substructure controversy, the wind-down of events feels rather uneventful in comparison. However the resolution of the issues at hand cannot be traced solely to the NRC II report, but rather to a number of lesser yet intertwined events that unfolded slightly earlier.

4.3 Broad Implications

First, in 1993, standards of evidence changed with the Supreme Court's decision regarding *Daubert v. Merrell Dow Pharmaceuticals*. However, it is difficult to determine if *Daubert* increased or decreased the admissibility of scientific testimony as a whole because prior to the change, some courts were following the *Frye* standard while others were following the federal rules. Though because DNA admissibility increased dramatically around the same time, it is certainly plausible that the introduction of the *Daubert* standards allowed for increased acceptance of forensic DNA analysis.

However, this paradigm shift away from the problems of population substructure was perhaps influenced most by the rise of a more significant dilemma. Human error and evidence contamination had always existed as an issue in forensic DNA analysis. However scientific developments in PCR testing, including the FBI starting PCR casework in 1994, increased concerns about issues of human error and contamination

(Butler, 2005). When performed correctly, PCR testing generated more reliable and efficient results. However PCR testing was also very susceptible to manipulation or contamination, an issue that had not yet risen to the forefront of the forensic DNA debate. With the rise in PCR testing, this began to change. Even if following the very conservative measures presented in the first NRC report, DNA analysis would still often yield random match probabilities less than 1 in 100,000, or even 1 in 1,000,000, “figures which are almost certainly smaller than any laboratory’s false positive error rate,” (Lempert, 1997). Aided by the extremely prolific nature of the O. J. Simpson case, these problems quickly overshadowed the small calculation errors debated in the ‘DNA Wars’. By the mid 90s, issues of contamination and human error were viewed a more important problem for the forensic community than calculating exact statistical representations of random-match probabilities. Certainly debate still flares up occasionally regarding random-match calculations, but the implications are simply not as drastic as those surrounding problematic laboratory practices and methods. This paradigm shift becomes extremely evident in the Simpson case, where the defense tactics shift from the previous strategy of attacking DNA technology as a whole, to attacking lab specific quality control issues instead.

4.3 Specific Events and Cases

The events discussed in the previous section demonstrate broad-spectrum issues that each probably played a role in the eventual ceasefire. However there were certainly some more explicit cases that occurred.

By the arrival of the ever-controversial Simpson trial, problems from the DNA wars already been resolved due to studies involving increased data and changes in technology. The Simpson Case, which has been the sole forensic focus of multitudes of articles, books, and TV series, is simply too complex for the scope of this paper. The case itself touches on much more than just the reliability of forensic DNA technology, spanning topics from racism to invasive media to the role of affluence in the court system. Hence, I have chosen to discuss the Simpson trial only briefly, examining the main forensic difference between the case and its predecessors to demonstrate changes in the environment surrounding forensic DNA analysis.

A brief description of the crime follows. On June 12th 1994, the second wife of prominent football star and actor Orenthal James Simpson, Nicole Simpson was found stabbed to death in a wealthy suburb of Los Angeles (Kaye, 2010). A man's body lay next to her, also mutilated. The man was Ronald Goldman, a waiter at a fancy restaurant who had come by the house to return a lost pair of eyeglasses. O.J., with his history of violent activity, was immediately the suspected murderer (Kaye, 2010). Although the evidence against Simpson was certainly strong, most people involved with the case hinged on the DNA evidence (Aronson, 2007).

O.J. Simpson, certainly an affluent man himself, hired a capable team of defense lawyers that later became known as the dream team that included Scheck and Neufeld (Aronson, 2007). The prosecution also included very capable lawyers, including San

Diego County Prosecutor, George “Woody” Clarke who has since written his own book chronicling early DNA evidence admissibility in California (Clarke, 2008).

From the prosecutions description, earlier in the evening of the murder, Simpson and his wife had attended their daughter’s dance recital, a meeting at which they were barely civil to each other (Kaye, 2010). After this, Simpson came to his ex-wife’s house, and hid in the bushes. He may have been planning to murder her, scare her, or simply slash her tires, something he had reportedly done before (Kaye, 2010). However, when Goldman arrived at the scene Simpson grew consumed with violent rage, and attacked and murdered both his wife and the waiter. During the attack, Simpson cut his hand, leaving some of his own blood, as well as a trail of bloody footprints. Drops of his blood were found in his car, and a glove was found at the scene of the crime. In the haste of his escape, Simpson left a pair of bloodied socks at the crime scene. “DNA typing linked bloodstains leading from the bodies to O.J., those in the car to Nicole, Goldman, and O.J.; those in the estate grounds and rooms to O.J.; and those on O.J.’s socks to O.J. and Nicole,” (Kaye, 2010). At the time there had never been a case where more DNA evidence was collected and used against one defendant (Thompson, 1996).

Originally the defense team motioned to exclude all DNA evidence. However the samples had been sent to multiple labs, meaning that the defense team would have to demonstrate that no forensic labs in the state could accurately perform VNTR testing (Kaye, 2010). Since the Simpson trial was pre-NRC II, the defense team was able to cite the complaints about the ceiling rule following NRC I in an attempt to demonstrate that the scientific community was still debating population statistics. However, it is important

to realize that this argument fails to take into account the vast amounts of research on the topic being completed and published right around this time.

In response to this, a strange partnership sprang up between Eric Lander and Bruce Budowle, two men who had often faced each other from opposite sides of the battlefield. Lander and Budowle, who had both been working on gathering and analyzing huge amounts of data regarding population genetics, feared that the confusion whipped up by the media frenzy and the star defense team could destroy their recent progress (Aronson, 2007). As a response to Simpson's defense team, Lander and Budowle came together to publish an article claiming that forensic DNA technology was finally reliable and effective in the court setting (Aronson, 2007). In the article, entitled *DNA Fingerprinting Disputes Laid to Rest*, they both admitted that early attempts at DNA typing were extremely problematic but stressed that recent developments had greatly improved DNA typing to a point where it was now reliable and effective. The article cited an enormous FBI conducted study titled *VNTR Population Data: A Worldwide Study*. In this study, the FBI collect data from 25 sub groups spread across more than 50 geographic areas. "The report represented a tremendous amount of work and worldwide cooperation among academic scientists, forensic scientists, research institutions, and crime laboratories around the world." (Aronson, 2007). In conclusion, the report demonstrated that "variations in allele frequency among subgroups were present, but modest and not forensically significant," allowing a return to the original product rule (Aronson, 2007). The article was met with some negativity, mostly from Lewontin and some of his students, who referred to the article as well-placed propaganda. Regardless, for the courts it meant that two extremely influential scientists, one from each side, had

declared a consensus. Admissibility of DNA based evidence increased dramatically around the United States.

Eventually the defense dropped the motion to exclude evidence (Kaye, 2010). Their new strategy is one that echoes the new concerns regarding DNA technology, concerns of error rates and contamination. The defense story wove a complex web of new possibilities. First, the defense argued that Simpson cut himself while removing his cell phone from his car, leaving the drops of blood on his driveway, his car, and foyer (Thompson, 1996). While in Chicago he heard the news of his ex-wife's murder, and broke a glass in his hotel room, cutting himself again, and resulting in the two cuts present on his finger. The defense then argued that the Rockingham glove was contaminated with Simpson's DNA at a LAPD laboratory, when criminalist Collin Yamauchi spilled some of Simpson's blood from a vial shortly before handling the glove (Thompson, 1996). The defense then argued that the DNA of the true murderer had degraded when the samples were left in a warm, damp environment. The defense was able to demonstrate that the LAPD crime scene investigators were poorly trained and made many serious errors (Thompson, 1996). Eventually, Simpson was acquitted, concluding what is often referred to as the 'trial of the century'.

Most people have probably drawn their own conclusions about the Simpson case, empirically founded or not. More interesting, however, is how the Simpson case demonstrates the ever-changing nature of science within the legal system. At the time of the Simpson case, the controversy that had surrounded the earlier cases was starting to dissipate, making it seem unlikely that the defense would be able to fight massive quantities of DNA evidence. With the publication of the Lander-Budowle article it seem

impossible that the evidence could be deemed inadmissible. Even so, the end of one debate often gives rise to another. As issue regarding DNA profiling shifted to manipulation and contamination of data, the Simpson defense team, having plenty of practice, evolved right along side. And perhaps that is how it should be. As defense council William C. Thompson explained, “There were a number of scientific issues in the case, but they had little to do with the fundamental science underlying DNA technology or the details of laboratory procedures for typing DNA, and everything to do with the potential for cross contamination of samples before they reached the DNA laboratories.” (Thompson, 1996). Opinions of the Simpson trial aside, the laboratory and crime scene team procedures uncovered by the defense team were dismal. Because of the conclusion of the Simpson trial, new regulations emerged.

Following the Simpson trial in 1995, laboratories around the United States had started to improve their procedures for collecting and transporting evidence, as well as preventing contamination within the laboratory itself (Butler, 2005). In 1994, the DNA Identification Act was passed, which commissioned an advisory board with the purpose of developing recommended standards for quality assurance (Federal Bureau of Investigation, 2009). The board was made up of people from both private and public sectors. The board produced two documents that were both recommended to the FBI. By 1999, both “Quality Assurance Standards for Forensic DNA Testing Laboratories” and “Quality Assurance Standards for Convicted Offender DNA Databasing Laboratories” were issued by the Director of the Bureau (Federal Bureau of Investigation, 2009). The 1994 act required that all DNA laboratories that are federally operated, part of National DNA Index System (NDIS), or labs receiving federal funds demonstrate compliance with

the standards issued by the FBI, meaning they now had to comply with these new standards (Federal Bureau of Investigation, 2009). Certainly the issue of human and laboratory error was still present. However, “The issuance of the DNA advisory Board Quality Assurance Standards has helped raise the professional status of forensic DNA testing. It is noteworthy that in a systematic analysis of circumstances normally encountered during casework, no PCR contamination was ever noted according to a recent (1999) study,” (Butler, 2005).

Still, regulation and testing of forensic laboratories is a complex issue. DNA laboratory error rates are extremely hard to estimate accurately. Proficiency tests demonstrate some understanding of error, however blind proficiency testing is quite the misnomer because laboratories often know they are being tested, “and are blind only to what the test results should be,” (Lempert, 1997). This is because true blind tests would be extremely expensive and very difficult to conduct. Because of this, even results produced by these so-called blind proficiency tests probably still understate laboratory error.

Chapter 5: Conclusions and Implications for Additional Research

“It (DNA) is powerful evidence both to convict and exonerate...It’s kind of a truth machine. But any machine when it gets in the hands of human beings can be manipulated or abused,” Peter Neufeld

Forensic DNA testing has come a long way from its start in the mid 80s to eventual acceptance by the late 90s. Today, people often view forensic DNA testing as

the perfect model with which to navigate the future of forensic science. Hence, despite increased admissibility of DNA evidence, future defense attorneys, prosecutors, and scientists will still have roles to play in making sure forensic evidence, including DNA analysis, remains trustworthy and reliable. Like any technology, we must remember that forensic DNA analysis results are only as reliable as the institutions producing them.

This analysis of early DNA history has unearthed a number of interesting issues. First, there was the issue of the private biotech companies, which paraded DNA evidence into courtrooms before any of the necessary tests or standards had been developed. In the future, our government, specifically the FBI, should be more vigilant about the emergence of private companies in forensic technology. However, at least in the case of DNA analysis, proficiency tests and accreditations of forensic laboratories have helped lessen this problem. Further research could be conducted here by asking if there are other similar examples in the United States where private companies have played major roles in producing expert witness testimony and generating technological development. Certainly it is possible that the corporate involvement we see in this analysis of forensic DNA testing may be an isolated issue. However, if some similar patterns can be seen in other disciplines, it may be worth analyzing the role large corporations play in the legal system.

Second, it is evident that problems can often arise even with government supervision and intervention, such as the problems generated by the FBI standardization procedures. The interesting aspect here is that both the private companies as well as the FBI made very similar mistakes, mostly the mistake of focusing on only one aspect of the problem at hand when certainly a multi-dimensional approach would have been

preferable. Like science, regulation is an evolving issue that must be constantly updated in order to remain relevant. The FBI's decision to include people from all disciplines on the DNA Advisory Board demonstrates an important change in structure and belief that has benefitted forensic DNA analysis immensely. The FBI's involvement with consistent updates of the documents suggested by the board demonstrates a desire to keep standards relevant and functional. However, even with these structures in place, problems still arise. For example, as recently as 2002 there was the Houston Crime Lab controversy, where an audit of the crime lab demonstrated lack of employee training, sloppy techniques, and possibly false testimony by lab workers. Possibly "the most compelling evidence of crime lab misconduct was the Sutton case, where a crime lab employee testified before a jury that the DNA from a semen sample strongly suggested that Sutton was a unique match, when in reality the DNA would be a match to approximately 1 in every 16 African American Males," (Eckroth, 2003-2004). This same testimony also omitted evidence from an additional semen sample that would have demonstrated Sutton's innocence (Eckroth, 2003-2004). Misconduct like this suggests that vigilant and consistent audits must remain continuous because evidence manipulation can still occur no matter how advanced the science becomes. However, at least in this case the audit succeeded in unearthing underlying problems of racially motivated corruption. After an in depth investigation, the crime lab was completely revamped. From this perspective, additional research could be conducted by analyzing institutions that are now accredited but have had problems in the past with misconduct and/or negligence.

Third, scientists and legal experts must remember that the forensic techniques exist in a constant state of fluctuation, where life, freedom, and justice are all at stake.

Overall, we must remember that forensic science belongs to two worlds, and the intersection is not always easy or functional. Cooperation, openness, and compromise are often key. Additional research could be conducted here by analyzing something I did discuss in this paper, that being the role of the jury. This paper focused primarily on evolving understandings and compromises between defense lawyers, prosecutors, scientists, and the FBI. However, the jury needs to understand complex evidence in order to ensure the success of a forensic technique. Further research into the best ways for communicating with jury member regarding issues of statistics, random-match probabilities, and error rates could be extremely valuable for both lawyers and scientists.

Despite the early problems with forensic DNA analysis, the evolution of the general technique is really quite extraordinary. In the end, credit must be delivered to the individuals on all sides of the conflict. For example, at times it is difficult to empathize with Neufeld and Scheck, who were often defending criminals guilty of rape or murder. However, as defense attorneys', their responsibility was to their client, and without their dogged determination in the face of the companies like Lifecodes, forensic DNA analysis may have never evolved from its early problematic state. In addition, Neufeld and Scheck now run the Innocence Project at the Benjamin N. Cardozo School of Law, where they use forensic DNA analysis to exonerate innocent people behind bars. Eric Lander is currently a biology professor at MIT and a leader of the Human Genome Project (Broad Institute, 2013). His research consists of researching diseases, including classifying cancer at a molecular level, with the hope of using the humane genome research to advance medicine (Broad Institute, 2013). Kenneth Kidd is a professor of genetics at Yale, where he has created ALFRED, the ALlele FREquency Database, an openly

accessible database of allele frequency data for DNA polymorphism among the human population (Yale School of Medicine, 2013). He has also recently served on DNA advisory panels, using DNA analysis to help identify victims of Hurricane Katrina and the September 11th attacks (Yale School of Medicine, 2013). Daniel Hartl is a professor of biology at Harvard University, where his lab focuses on biological public health hazards such as malaria parasite. In the end, we must remember that despite the bitter rivalries and debates that raged over the decade of forensic DNA development, the involved players are all brilliant people who have worked in their own ways to advance understanding of science and law. In the end, this analysis of the evolution of DNA technology through the legal system is not meant to criticize these hardworking individuals, but instead to take what they have learned and apply it for the benefit of the American public.

Bibliography

- Aronson J. 2007. Genetic Witness: Science, Law, and Controversy in the Making of DNA Profiling: Rutgers University Press.
- Butler JM. 2005. Forensic DNA Typing: Biology, Technology, and Genetics of STR Markers: Academic Press.
- Benedict R. 1988. DNA Profiles Again Match Rape Suspect. The Orlando Sentinel, Florida.
- Butler, J. 2009. Fundamentals of Forensic DNA Typing. Academic Press.
- Calladine CR, Drew H. 1997. Understanding DNA: the Molecule and How It Works: Academic Press.
- Christensen, AM. 2004. The Impact of Daubert: Implications for testimony and research in forensic anthropology (and the use of frontal sinuses in personal identification). J Forensic Sci 49:427.

- Clarke G. 2007. Justice and Science: Trials and Triumphs of DNA Evidence. Rutgers University Press.
- Clayborne, C. 1989. Evidence of DNA Finger Printing Admitted for Identification Purposes in Rape Trial. UALR LJ 12:543.
- Coleman H, Swenson E. 1994. DNA in the Courtroom: A Trial Watcher's Guide. Genelex Corp.
- Connors E. 1996. Convicted by Juries, Exonerated by Science: Case studies in the use of DNA evidence to establish innocence after trial. DIANE Publishing.
- “CODIS—Quality Assurance” 2009. http://www.fbi.gov/about-us/lab/biometric-analysis/codis/codis_quality. Accessed April, 2013.
- Daemmrich A. 1998. The Evidence Does Not Speak for Itself: Expert Witnesses and the Organization of DNA-Typing Companies. Social Studies of Science 28:741-772.
- “Daniel L. Hartl-Faculty”. 2012. <http://www.oeb.harvard.edu/faculty/hartl/hartl-oeb.html>. Accessed March 2013.
- Ebert LB. Frye After Daubert: The Role of Scientists in Admissibility Issues as Seen Through Analysis of the DNA Profiling Cases. University of Chicago Law School Roundtable: 219.
- Eckroth, J. 2003-2004. Tainted DNA Evidence and Post-Conviction Reversals in Houston, Texas: Suggested Solutions to Curb DNA Evidence Abuse. Am. J. Crim. L. 433:
- “Eric. S. Lander: Broad Institute of MIT and Harvard”. 2013. <http://www.broadinstitute.org/history-leadership/scientific-leadership/core-members/eric-s-lander>. Accessed March 2013.
- Gianelli PC. 1980. The admissibility of novel scientific evidence: Frye V. United States, a half-century later. colum l rev 80:1197.
- Greenberg S, Katzer M, Joseph Harris J. The People of The State of New York, Plaintiff v. George Wesley, Defendant. The People of the State of New York, Plaintiff v. Cameron Wesley, Defendant. 140 Misc 2d306.
- Greenburg S, Heran J, Mahoney J. The People of the State of New York, Respondent v. George Wesley. Appellate Division of the Supreme Court of the State of New York, Third Department.
- Haack S. 2004. Truth and Justice, Inquiry and Advocacy, Science and Law. Ratio Juris 17:15-26.

Imwinkelried EJ. 1991. The Debate in the DNA Cases Over the Foundation for the Admission of Scientific Evidence: The Importance of Human Error as a Cause of Forensic Misanalysis. Wash ULQ 69:19.

“The Innocence Project-About Us. Barry Scheck and Peter Neufeld.” http://www.innocenceproject.org/Content/Barry_C_Scheck_Peter_J_Neufeld.php. Accessed March 2013.

Jobling MA, Gill P. 2004. Encoded Evidence: DNA in forensic analysis. Nature Reviews Genetics 5:739-751.

Kaye DH. 2010. The Double Helix and the Law of Evidence: Harvard University Press.

Kaye, DH. 197. DNA, NAS, NRC, DAB, RFLP, PCR, and more: An Introduction to the Symposium on the 1996 NRC Report of Forensic DNA Rvidence. Jurimetrics 37: 395.

“Kenneth Kay Kidd, PhD”. 2010. http://medicine.yale.edu/genetics/people/kenneth_kidd-2.profile. Accessed March 2013.

Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle KH, Zahler AM, Haussler D. 2002. Genome Res 12:996-1006.

Lempert, R. 1997. After the DNA Wars: A Mopping Up Operation. Isr. Rev. 31: 536.

Lewis R. 1989. Genetics Meets Forensics. Bioscience 39:6-9.

Lewontin RC, Hartl DL. 1991. Population genetics in forensic DNA typing. Science 254:1745-1750.

Lynch M, Cole SA, McNally R, Jordan K. 2010. Truth Machine: The Contentious History of DNA Fingerprinting: University of Chicago Press.

Lynch M, Jasanoff S. 1998. Introduction: Contested Identities: Science, Law and Forensic Practice. Soc Stud Sci 28:675-686.

Moss DB. 1988. DNA-The New Fingerprints. ABAJ 74:66.

Morton, NE. 1997. The Forensic DNA Endgame. Jurimetrics 37: 477-494.

Patton SM. 1990. DNA Fingerprinting: The Castro Case. The Harvard Journal of Technology and Law 3.

Roberts L. 1991. Fight erupts over DNA fingerprinting. Science 254:1721.

- Robin GD. 1994. DNA Evidence in Court-The Odds Aren't Even. *Crim Just* 9.
- Saks M, Koehler JJ. 2005. The coming paradigm shift in forensic identification science. *Science* 309:892-895.
- Smith J. *The People & C v. George Wesley*. 83 N Y 2d 417, 633 N E 2d 451, 611 N Y S 2d 97 3.
- Thompson, W. and Krane, D. 2003. Chapter 11: DNA in the courtroom. *Psychological and Scientific Evidence in Criminal Trials*. West Group.
- Thompson W. 1989. DNA typing: Acceptance and weight of the new genetic identification tests. *Va Law Rev*:45.
- Thompson WC. 1996. DNA evidence in the OJ Simpson trial. *U Colo Rev* 67:827.
- Thompson WC. 1993. Evaluating the admissibility of new genetic identification tests : lessons from the D.N.A. war. *The Journal of Criminal Law and Criminology* 84:22-104.
- Warden R. 2005. Illinois Death Penalty Reform: How it Happened, What it Promises. *Journal of Criminal Law and Criminology* 95:381-426.
- Weedn VV, Gaensslen RE. 2007. *Handbook of Forensic DNA*: CRC.
- White . DNA Fingerprinting and the Law. *Modern Law Review* 51:145.