2970185818

MITRE TECHNICAL REPORT

State of the Art Biometrics Excellence Roadmap

Technology Assessment: Volume 3 (of 3) DNA

March 2009

Marc Colosimo, Ph.D.
Russell Graef
Shawne Lampert
Matthew Peterson

Sponsor: B. Scott Swann, Program Manager
Dept. No.: G122
Contract No.: J-FBI-02-093, Task 029 (SABER)
Project No.: 1402F029-BI

The views, opinions and/or findings contained in this report are those of The MITRE Corporation and should not be construed as an official government position, policy, or decision, unless designated by other documentation.

This document reflects the state-of-the-art as of Winter 2009.

MITRE

Executive Summary

This report documents the results of The MITRE Corporation's assessment of the science and technology of Deoxyribonucleic Acid (DNA) biometrics as part of the State of the Art Biometrics Excellence Roadmap (SABER) study for the FBI Criminal Justice Information Services (CJIS) Division Technology Evaluation Standards Test (TEST) Unit.

The great success and wide acceptance of the use of DNA biometrics for human identification has led to tremendous advances in research of new and improved technologies and methodologies, and to adoption of these technologies in commercial industry and the public sector. However, governmental, and specifically law enforcement, use of new DNA biometric technologies, is subject to stringent federal quality assurance measures and controls to protect the quality of collected and stored DNA samples and to ensure the accuracy, privacy, integrity, robustness, and consistency of DNA typing methods and results.

The purpose of this paper is to provide an assessment of the use of DNA biometrics primarily as it pertains to human identification in the U.S. and internationally. Because the intended audience for this document includes management throughout the FBI, other government personnel, and the general public, much of the information presented is high-level, intended to impart broad concepts rather than specific detail best understood by forensic scientists. This assessment is presented from two perspectives:

1. State-of-the-art DNA biometrics across all related communities including research, commercial industry, and the public sector
2. State-of-the-practice DNA biometrics, especially as these biometrics apply to the law enforcement sector.

The paper then identifies high-level challenges facing U.S. national, state, and local law enforcement communities and presents some broad recommendations for CJIS consideration, summarized below:

- Although commercial DNA biometric databases and related services may provide a starting point for limiting the scope of investigations at state and local levels, they have little value at the national level. However, advances in data-mining techniques used in this industry should be monitored for potential governmental use.
- Advances in new biotechnology equipment and methodologies in research should also continue to be monitored for applicability to law enforcement. Additional resources, including facilities, equipment and materials to analyze DNA and validate test results, and most importantly, qualified personnel, should be identified and funded to support this objective.

- Continued active participation by the FBI on various international committees should be supported to ensure harmonization of international standards. From a national perspective, rapid developments in advanced DNA biometric technologies, discussed in the preceding bullet, must be validated for compliance with national standards defined by the Scientific Working Group on DNA Analysis Methods (SWGDAM) and National Institute of Standards and Technology (NIST).
- Continued FBI leadership on international standards committees and in international DNA program implementation efforts is endorsed to ensure 1) international DNA biometric standards are published that will protect the privacy of U.S. DNA information and 2) systems are implemented and tested to ensure compliance with these standards.

Acknowledgements

This report was funded and actively supported by many dedicated experts within the FBI and NIST. The authors wish to acknowledge special thanks to members of the FBI Laboratory and NIST for their comprehensive review and comments.

Page Intentionally Blank

Table of Contents

List of Figures

List of Tables

1 Deoxyribonucleic Acid (DNA) Biometrics

Deoxyribonucleic acid (DNA) is that part of a cell that stores genetic information unique to each living organism. DNA typing, also known as DNA fingerprinting or DNA profiling, is a biometric technology that measures and analyzes DNA to distinguish organisms or individuals with some degree of probability. Today, DNA typing is used to improve livestock and plant crops, identify inheritable diseases and rare medical conditions, discover drugs, determine animal migration patterns, examine human population diversity, and identify humans individually.

1.1 Purpose

The purpose of this report is to provide an assessment of the use of DNA biometrics in the U.S. and internationally, and is one of three volumes addressing a wide range of biometric technologies. This assessment is presented from two perspectives:

1. State-of-the-art DNA biometrics across all related communities including research, commercial industry, and the public sector
2. State-of-the-practice DNA biometrics, especially as these biometrics apply to the law enforcement sector.

The distinction is important because application of state-of-the-art biotechnologies in non-governmental sectors is not required to meet the same quality assurance standards and privacy regulations that are enforced within the federal and state law enforcement communities.

1.2 Scope

The intended audience for this document includes staff and management throughout the FBI, other government personnel, and the general public. Consequently, much of the information contained herein is presented at a high-level to impart broad concepts rather than specific detail understood best by forensic scientists.

To address the general audience, Sections 2 through 4 focus on the science of DNA as a biometric across all sectors that use DNA for human identification. Section 2, Introduction to the Science of DNA and the Human Identifier, presents an overview of the historical use of DNA for human identification, discusses the science of current DNA typing and analysis methods, and highlights examples of specific applications for investigative purposes. Section 3, State of the Industry discusses the DNA life-cycle processes from the initial collection and storage of material from which a DNA sample can be extracted to the creation of the digitized DNA profile to compare the sample profile to other profiles for purposes of identification. Also in Section 3, various methodologies are described and commercial products, identified. Specific limitations for forensic use are noted. Section 4, Database Core Loci, Technologies, and Search Services, describes current DNA profile databases and search services used not only for U.S. and international law enforcement but also for public human identification applications such as ancestry.

The remainder of the paper focuses mainly on DNA within the national and international law enforcement communities. Section 5, Legislation, documents national and international legislation that specifically prescribes the use of DNA within respective law enforcement sectors and judicial systems. Section 6, Standards, presents a discussion of national quality assurance standards and enforcing agencies to ensure the validity, reliability, integrity, and accuracy of DNA typing and searching results such that these results are admissible in a court of law. Section 6 also documents international standardization efforts that can affect U.S. DNA information-sharing efforts with the international community. Section 7, Programs, highlights specific U.S. and international DNA programs.

Section 8, Challenges and Technology Gaps, identifies challenges facing U.S. national, state and local law enforcement communities resulting primarily from an ever-increasing demand for DNA analysis, newly developed typing and processing methodologies that require validation before application to the law enforcement sector, and the additional need for trained and qualified personnel. Section 8 also discusses the issue of maintaining individual privacy.

Section 9, Recommendations, presents high-level recommendations related to database and search services, resource requirements, information privacy and security, and international standards efforts.

1.3 DNA for Human Identification

1.3.1 Non-Forensic

Human DNA is inherited and over time that DNA mutates or changes. By comparing DNA mutation patterns, geneticists can infer the evolutionary history of particular populations and, thereby, determine human migration patterns. Some such studies confirm that Polynesian ancestry traces to Southeast Asia, not to the Americas, and that Native Americans migrated from Siberia, not the Middle East. Another study performed by Dr. Sykes, University of Oxford, determined that the maternal ancestry of most Europeans traces to one of seven women, six of whom where hunter-gathers and the other, an early farmer. These seven women were related and by tracing their maternal genealogy can be found to be descendants of one African woman who lived about 150,000 years ago(Sykes, 2003). In fact, results from the Genographic Project suggest that all humans today are descendants of a group of African ancestors, who began migrating around 60,000 years ago(National Geographic, 2009).

In April, 2008, the National Geographic reported that DNA tests of bone fragments recently exhumed in the Ural mountains are those of two children of Tsar Nicholas II, Alexei and Maria, whose remains had been missing since the family was murdered in 1918. Previous genetic testing had already confirmed the remains of Nicholas, his wife Alexandra, and three of their daughters including those of Anastasia, whom many thought survived. Many claimants worldwide alleged to be Anastasia, most notably, Anna Anderson, whose DNA tests performed after her death determined she was Franziska Schanzkowska, a Polish peasant.

1.3.2 Forensic

Use of DNA for human identification has had a tremendous impact within the law enforcement community and judicial system. Previously unsolved crimes are solved and persons wrongly convicted are exonerated. The use of DNA technology internationally has enabled nations to track, capture, and convict escaped criminals who have crossed international borders and to prove cases of genocide.

The great success and wide acceptance of forensic DNA investigative methods place huge demands on national and state governments, and on local laboratories to evaluate, validate, prioritize, and implement new DNA technologies to accommodate increased demand for DNA testing. In the U.S., stringent federal quality assurance measures and controls are instituted to protect the quality of collected and stored DNA samples and to ensure the accuracy, integrity, robustness, and consistency of DNA typing methods and results. Ongoing audits to ensure compliance to federal quality assurance guidelines are also mandated. Implementation of new DNA typing methods must undergo strict validation procedures before acceptance by the U.S. national law enforcement community.

The Combined DNA Index System (CODIS) program is the U.S. national DNA system that shares, stores, and searches DNA profiles created by federal, state, and local crime laboratories to assist in the identification of suspects in crimes and of missing persons. The FBI Laboratory funds and manages this program. As of November 2008, the National DNA Index (NDIS), which is the national segment of CODIS, contained around 6.5 million offender profiles and 250,000 forensic profiles, and assisted in more than 79,300 investigations(FBI Lab, 2009).

2 Introduction to the Science of DNA and the Human Identifier

In April 1953, researchers Francis Crick and James Watson first described the "double helix" structure of DNA in *Nature*, an international journal of science. This famous work, coupled with new imaging techniques, was the basis for the now accepted understanding that the DNA molecule is a spiral ladder connected by alternating bases, and that the structure relates to the storage of the genetic information necessary to duplicate and maintain the organism. DNA is present in all living organisms; the segments of DNA responsible for carrying genetic information about the organism are called genes. The majority of the DNA is shared by an entire species. With the exception of identical twins, only one tenth of one percent of human DNA differs between any two individuals. For a high-level discussion of the biology of DNA, see Appendix D, Biology of DNA.

The study and subsequent understanding of the unique portions of DNA is the basis for DNA profiling. DNA profiles can also contain information on familial relationships between subjects. Various DNA profiling techniques (also called DNA fingerprinting and DNA typing) are available, depending on the situation and on the quality and nature of the biological samples available for analysis. The following sections first present a general background of the use of DNA for human identification and then discuss current DNA analysis methods for human identification applications that include:

- Autosomal Short Tandem Repeats (STRs)
- Y Chromosome STRs
- X Chromosome STRs
- Single Nucleotide Polymorphisms (SNPs)
- Mitochondrial DNA (mtDNA).

Table 2-1 highlights the state-of-practice application of these methods, discussed in detail in Section 3, State of the Industry.

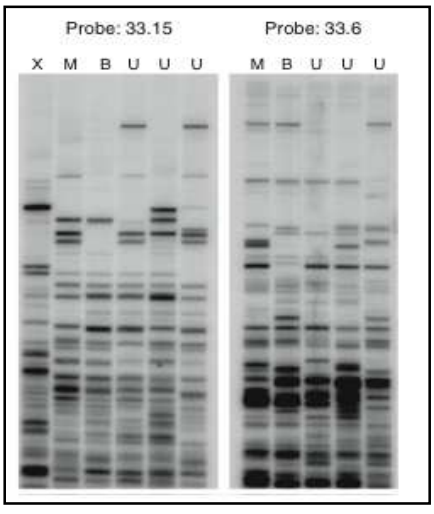
Table 2-1. Current Use of DNA Analysis Methods

Application	STRs	Y-STRs	X-STRs	SNPs	mtDNA
Human Identification	Y	Y	Y	Y	Y
International Law Enforcement	Y	Y	Y	N	Y
U.S. Law Enforcement	Y	Y (Limited Use)	N	Y (Limited Use)	Y

2.1 Background

In the spring of 1985, Sir Alec Jeffreys of Leicester University published the first report on his DNA fingerprinting method that used Restriction Fragment Length Polymorphisms (RFLPs) (Jeffreys, Wilson, & Thein, 1985). Using the RFLP method, DNA samples are treated with a restriction enzyme, which cleaves (or cuts) the DNA in a sequence-dependent manner thereby generating fragments of various sizes. These samples are then separated by size using the technique of slab gel electrophoresis (smaller fragments will travel through the charged gel faster than larger ones). The fragments are then denatured into single strands with heat or an alkaline solution. After transferring the sorted and separated DNA fragments from the gel onto a nylon membrane, the membrane is treated with radioactive (or fluorescent DNA) probes. The probes have specificity for particular sequences within the DNA, such that they bind to specific positions or loci. When the nylon membrane is placed against an X-ray sensitive film, the positions of DNA fragments carrying the radioactive probes appear as a series of dark lines (Jones, 2004) resembling bar codes.

The Jeffreys RFLP technique came to the attention of Sheona York, a lawyer who represented a Ghanaian family of U.K. citizenship entangled in an immigration dispute. The family's youngest son had recently returned from a trip to Ghana, but the Home Office detained him, alleging that he was unrelated to the family and held a forged passport. York asked Jeffreys for his help. Jeffreys compared DNA from the boy, the woman who claimed to be his mother, and her other children.



In this first application of DNA fingerprinting, the technique revealed that the boy was the woman's son and that all of her children shared the same father.

In Figure 2-1 the mother's DNA fingerprint is depicted in column (M); her three undisputed children are indicated in columns (U); the disputed boy, (B); and Alec Jeffreys, (X) because the father was not available (Jeffreys, Brookfield, & Semeonoff, 1985).

Figure 2-1. First Application of DNA Fingerprinting

DNA fingerprinting gained acceptance for forensic application in the U.K. when it was used to solve the 1983 and 1986 murders of 15-year-old schoolgirls Lynda Mann and Dawn Ashworth (Wambaugh, 1989).

In 1987, Lifecodes Corporation performed the first use of RFLP analysis in a U.S. criminal court where Tommy Lee Andrews was convicted of a series of sexual assaults in Orlando, Florida. In 1988, the FBI implemented RFLP analysis, after improving its robustness and sensitivity and collecting extensive data on the frequency of RFLP variants in different populations. This strategy uses single-locus probes, which bind to a single location in the DNA, resulting in a less complex DNA profile.

RFLP analysis gained acceptance in U.S. courts and its usage by forensic laboratories was wide-spread in the U.S. throughout the 1990s. Although powerful in terms of its discrimination power, RFLP analysis has limitations. Not only is the process time-consuming, but original typing methods required microgram quantities of high-molecular weight DNA that was not degraded or compromised, and the analysis of complex images. These constraints precluded its usage for testing several kinds of forensic evidence (i.e., those that may contain limited amounts of biological material, and/or may contain DNA that has been degraded due to aging, environmental exposure, or chemical treatment).

Some initial limitations of DNA analysis were addressed in the late 1980s with the development of the polymerase chain reaction (PCR) for which Kary Mullis received the 1993 Nobel Prize in chemistry, and by the discovery of STR DNA markers (or microsatellites). The PCR process amplifies or multiplies to create a billion copies of a target nucleotide sequence found in DNA. PCR amplification can be performed from minimal amounts (e.g., 500 pg or $5x10^{-8}$ grams) of DNA (Butler, 2007). Thus, PCR analysis is well suited for applications that entail limited amounts of biological sources. In addition, PCR analysis is less labor-intensive and has a faster turn-around, making biometric methods that utilize PCR analysis the current standards (see Section 3.1.5, DNA Amplification).

2.2 Autosomal Short Tandem Repeats

In 1991, Baylor College of Medicine's Thomas Caskey began using microsatellites (short tandem repeats) for his work in clinical genetics. Short tandem repeats are polymorphic DNA loci consisting of a repeated core nucleotide sequence resulting in regions of DNA that vary in length. STRs contain small repeating core unit, typically three (trinucleotide), four (tetranucleotide), or fives (pentanucleotide) bases, with a typical length of less than 300 bases (Martin, Hermann, & Schneider, 2001). Due to their small size, STRs are easily amplified and are ideally suited for typing trace samples, or old or degraded DNA samples. Peter Gill at Britain's Forensic Science Service applied a method that enabled the simultaneous analysis of multiple STRs, called multiplexing, which used four loci for routine profiling. The second generation multiplex used seven autosomal STR loci and another locus for sex determination. These techniques have made STRs the most commonly used method for identification with more than 20 markers being used in a variety of applications.

MiniSTRs are reduced size STR amplicons generated by a different set of PCR primers, which generate smaller length PCR products. One major advantage of using the same loci is that database compatibility can be maintained with samples processed from mass disasters or forensic

evidence using the traditional core STR multiplexes. In addition, smaller length PCR products have a higher success rate with degraded DNA than full length STR multiplexes (Coble & Butler, 2005) and are more favorably analyzed by alternative technologies such as time-of-flight mass spectrometry and rapid microchannel electrophoretic separations (Coble & Butler, 2005).

2.3 Y Chromosome STRs

The sex chromosomes (gonosomes), the X and Y chromosomes, are unique in several aspects from the other nuclear chromosomes (autosomes). Y chromosomal analysis is useful in sexual assault cases where samples may contain mixed male and female cells or when relatively low levels of male DNA are mixed with high levels of female DNA. Because the Y chromosomes are paternally inherited, they are useful for applications that seek to resolve paternal identity. Since parts of the X and Y chromosomes show high sequence similarity, Y amplification products are rarely observed in females (Gusmao, et al., 2005). Thus, extra care must be taken when selecting loci for amplification. In addition, seven STRs (but nine amplified products) are required for database entry (minimal haplotype loci, minHt) and recommended for court use (Pascali, Dobosz, & Brinkmann, 1998).

2.4 X Chromosome STRs

X chromosomes can be used to establish relationships between first or second-degree relatives, such as aunt–niece pairs and cousins. X chromosome typing may be especially useful in deficiency paternity testing where a sample from a putative father is not available and samples from the putative paternal relatives are available.

More than 30 X chromosome STR markers are used in identification applications today with the primary use being in complex kinship analysis; for example, to identify remains resulting from wars or mass disasters (Szibor, 2007; Cerri, Verzeletti, Gasparini, Bandera, & DeFerrari, 2006; Oguzturun, Thacker, & Syndercombe-Court, 2006).

2.5 Single Nucleotide Polymorphisms

A single nucleotide polymorphism (SNP) is a single base sequence variation between individuals at a particular point in the genome. SNPs are highly abundant in the human genome and are actively used to find genes associated with human disease and response to pharmaceuticals. Their low mutation rates, short amplicons, and applicability for high throughput genotyping technologies make SNPs potentially more useful than STRs for identification applications, especially in cases of degraded samples. In addition, SNPs could be used to predict ancestry or phenotypic characteristics, such as hair color or eye color (Butler, Coble, & Vallone, 2007) or used in paternity(Borsting, C., et al., 2008). Future use of SNPs for forensic applications is envisioned (Budowle & van Daal, 2008).

Identification of human remains from the World Trade Center disaster was aided by using a panel of approximately 70 non-phenotypic autosomal SNPs (identification SNPs) whose lengths were less than 100 nucleotides long and suitable for degraded DNA specimens(Biesecker, et al., 2005; Vallone, Decker, & Butler, 2005).

2.6 Mitochondrial DNA

Mitochondrial DNA (mtDNA) analysis is useful for the analysis of hair, teeth, and bones, as well as highly degraded tissues that do not lend themselves to successful nuclear (autosomal and gonosomal) STR DNA analysis. Mitochondria are inherited from the mother so all maternal relatives will have the same mtDNA barring mutations, thus it is not possible to uniquely identify an individual. Forensic profiles typically include, at a minimum, sequences in the hypervariable region I (HVI/HVSI), defined by nucleotide positions 16024-16365, and sequences in hypervariable region II (HVII/HVSII), defined by nucleotide positions 73-340 (SWGDAM, 2003). Profiles are designated as differences from the Cambridge Reference Sequence (CRS) or the revised Cambridge Reference Sequence (rCRS) (Andrews, Kubacka, Chinnery, Lightowlers, Turnbull, & Howell, 1999). Figure 2-2 depicts mtDNA showing the control region and the hypervariable regions (HV1/HV2) (NFSTC, 2008).

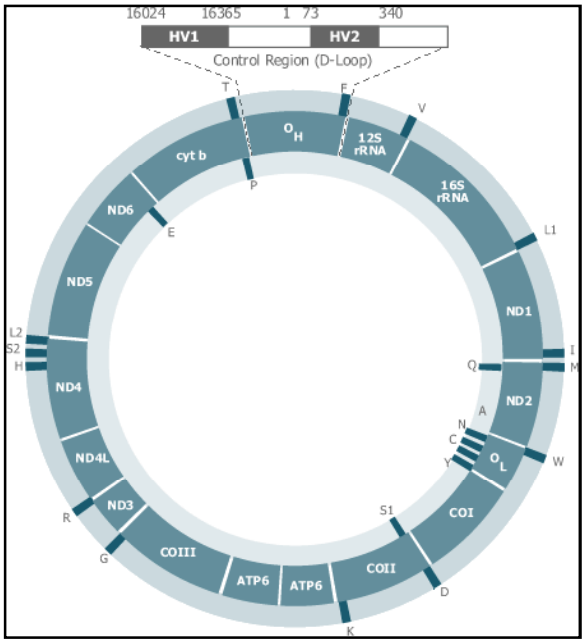


Figure 2-2. mtDNA with Control and Hypervariable Regions [HV1/HV2]

3 State of the Industry

This section presents a high-level discussion of operational requirements, processing methodologies, and commercial products used throughout the DNA life-cycle in the DNA biometrics commercial industries and research sectors, and in law enforcement. Specific limitations for forensic use are noted because the U.S. enforces stringent federal quality assurance measures and controls to protect DNA sources, samples, processing results, and stored profiles when used for law enforcement purposes. Since new DNA typing methods in the law enforcement sector must undergo strict validation procedures, some DNA processing methodologies are not implemented.

3.1 DNA Processes and Related Commercial Biotechnologies

A *DNA source* is the material from which a *DNA sample* can be extracted. Sources of DNA include blood, bones, dental molds, cigarette butts, eating utensils, chewing gum, licked postage stamps, ski masks, licked envelopes, toothbrushes, razor shavings, band aids, or semen found during a medical examination after sexual assault. The DNA sample, or collection of DNA molecules, can then be quantified, amplified, separated, and analyzed. The digital representation of the DNA analysis results is referred to as the *DNA profile*. The distinction is important because only DNA sources (material having DNA samples) and the samples themselves are collected; only DNA samples are extracted, quantified, amplified, separated, and analyzed; only the results of DNA analysis are digitized; and DNA sources, samples, and profiles all can be retained and stored. These terms are often misused or ambiguously used when describing the DNA processes and related biotechnologies during its life-cycle. These processes and related commercial biotechnologies, used by both governmental and private laboratories, are described below with some examples of how they are used by the forensics community, particularly the FBI.

3.1.1 DNA Collection

DNA samples are collected:

- Directly from an individual, typically by drawing blood or swabbing the inside of the cheek using a buccal swab to obtain a sample. For example, a relative of a missing person may voluntarily provide a blood sample to determine possible links between the relative and unidentified human remains. A sample from a known source, such as this case, is referred to as a “reference sample.”
- Indirectly by extracting blood from a bloodstain, often used during crime scene investigations. Other indirect DNA sample sources include hair from a hairbrush, saliva from a toothbrush, or any other object with which the individual had contact. These samples are often” unknown” and “evidentiary.”

Often the method for directly collecting a DNA sample depends upon the type of profile analysis that is required (discussed previously in Section 2, Introduction to the Science of DNA and the

Human Identifier). Sample collection requires use of personal protective equipment to guard against sample contamination and accidental exposure of the collector to infectious materials (FBI, 2008).

3.1.2 DNA Transport and Storage

DNA sample material is normally transported to laboratories that perform DNA analysis. Samples to be shipped are air dried and packaged in paper pouches or boxes (Butler(a), 2005). In most cases sample material is exempt from special carrier instructions, unless suspected or confirmed to be hazardous or to contain an infectious substance such as material recovered from a biological warfare site(IATA, 2008; USDT, 2007). Sample material may also be refrigerated (4°C) or frozen (depending on material), and transported by vehicle or other method.

Purified DNA samples can be stored in water or a buffered solution, such as "TE buffer," and maintained at temperatures of -20°C or at -80°C indefinitely. DNA samples stored on FTA™ bloodstain cards can be kept at ambient temperature without risk of degradation for many years (Kline, Duewer, Redman, Butler, & Boyer, 2002). In all cases, temperature and humidity fluctuations can increase the likelihood of DNA degradation.

Note that the FBI Laboratory requires submitters to be from a duly-authorized law enforcement agency. Proper chain of custody is required to preserve the integrity of recovered material and derived DNA samples (FBI, 2008).

3.1.3 DNA Extraction

A DNA sample is then extracted from the DNA source material and purified. Extraction methods commonly used by forensic laboratories are organic extraction methods such as phenol/chloroform (PC), and non-organic such as Chelex® and commercially available extraction kits, including the FTA® paper processing method (Whatman, PLC), QIAAmp® (Qiagen, Inc.), and DNA IQ (Promega, Inc.). These extraction methods are discussed in more detail in Appendix E.1, DNA Extraction.

3.1.4 DNA Quantitation

The extracted DNA sample is then quantified to determine the amount of DNA required to ensure amplification will yield a full (not partial) profile while avoiding the generation of PCR artifacts resulting from using too much DNA. Quantitation techniques include yield gels, spectrophotometry, fluorometry, slot blot hybridization, AluQuant®, and quantitative PCR (qPCR). The qPCR method is the most efficient and sensitive of the quantitation methods, and is amenable to automation. Appendix E.2, DNA Quantitation discusses these six techniques in more detail.

3.1.5 DNA Amplification

The DNA amplification process is used to create hundreds of millions of copies of DNA segments, such as the STR loci. The most common amplification method used for human identification by forensic laboratories is PCR. Starting with the extracted DNA sample as the

amplification template, the PCR process involves the repetitive denaturing of the double-strands of DNA, annealing short synthetic DNA pieces (called primers) to specific targeted sequences (i.e., STR loci, mtDNA HVRs) in the template, and extending DNA segments from the primers to make a copy of the template. The cyclical repetition of these three steps results in an exponential increase in the amount of amplified target DNA. Appendix E.3, DNA Amplification, describes this process in more detail.

Commercial PCR amplification kits for STR, miniSTR, and Y-STR analyses are available and commonly used by laboratories because of their convenience and quality assurance. Examples of commercial PCR amplification kits include:

- AmpFlSTR® Identifiler™, Profiler Plus ID™, COfiler™, Y-Filer™, and MiniFiler™ offered by Applied Biosystems, Inc.
- PowerPlex® 16, Y, and ES provided by Promega, Inc.

The Profiler Plus ID™ kit in conjunction with the COfiler™ kit and the single-amplification kits Identifiler™ and PowerPlex® 16 amplify the 13 core STR loci that are accepted at NDIS. Eight of the 13 core STR loci are available in the commercial miniSTR kit MiniFiler™.

A complete list of commercial STR kits can be found on the National Institute of Standards and Technology (NIST) Short Tandem Repeat DNA Internet Database (STRBase) web site[1] (Butler, 2008). Only those kits that are approved by NDIS/CODIS (see Section 7.1.2, CODIS, for additional information) and validated at the laboratory performing the amplification are accepted by NDIS/CODIS. Use of Y-STR kits for missing persons' cases at NDIS is planned for 2009.

3.1.6 DNA Separation

The amplified STRs are then separated according to length (a measure of the number of repeats within each STR locus, as well as that of the non-repetitive sequences that flank the repeat region). The most common DNA separation method is capillary electrophoresis, a process by which an electrical current is applied to a narrow, hollow capillary containing a viscous sieving polymer. When the amplified DNA is injected into the capillary, the DNA, which is inherently negatively charged because of its sugar-phosphate backbone, migrates toward the anode (positive pole) of the electric current. Because smaller DNA molecules travel through the polymer more quickly than larger molecules, the molecules are separated by size. Capillary electrophoresis instruments, such

[1] STRBase, Commercially Available STR Multiplex Kits, 23 June 2007, available from: http://www.cstl.nist.gov/biotech/strbase/multiplx.htm; Internet; accessed March 2008.

as the Applied Biosystems, Inc. AB 310 and 3100 series, are the most commonly used instruments for STR allele separation and sizing analysis in forensic applications (Butler, Buel, Crivellente, & McCord, 2004).

Another method for DNA separation is mass spectrometry. Mass spectrometry is an analytical technique that detects compounds by separating ions by their unique mass (mass-to-charge ratios) using a mass spectrometer. The use of mass spectrometry equipment such as the Ibis Biosciences' T5000 Biosensor System to identify SNPs of mtDNA, and to analyze STR and mtDNA is being explored for application to the forensic community (NIST, 2008). These techniques are described in Appendix E.4, DNA Separation.

3.1.7 DNA Profile Analysis

After DNA is separated, the raw image data are analyzed by sophisticated software programs to:

- Identify electrophoretic peaks corresponding to the amplified DNA fragment
- Determine DNA molecule peak values
- Generate the DNA profile based on comparisons to the known "relative size standard."

In law enforcement applications, the profile is then interpreted to determine if the profile 1) is from a single contributor or multiple contributors (mixed sample), and 2) is of sufficient quality for comparison.

U.S. federal quality assurance standards generally require DNA profile interpretation and subsequent technical confirmation by two qualified analysts "to ensure conclusions and supporting data are reasonable and within the constraints of scientific knowledge"(DAB, 1998) before the profiles may be uploaded to NDIS. However, no analyst review is required of single source known offender reference samples that are processed by NDIS-approved expert systems and have good quality data (FBI Laboratory, 2008). Expert systems approved by NDIS include GeneMapper® ID (AB), TrueAllele® (Cybergenetics), and FSS-i3™ Expert Systems Software (Forensic Science Service®) (Roby & Christen, 2007). See Appendix E.5, DNA Profile Analysis, for additional information.

3.1.8 DNA Profile Comparison

It is this common digital representation of DNA sequences (the DNA profile) that allows automated searching and comparison against other stored profiles for commercial and governmental applications to identify paternal relatives, to support disaster victim identification and recovery efforts, or to exchange internationally to identify and track criminals who may have crossed borders, among others.

In law enforcement applications, the DNA profile (also known as the unknown or evidentiary profile) is compared to other DNA profiles (reference or known profiles), resulting in one of three outcomes:

- "If the DNA profiles differ, an exclusion is declared; this means that the known individual is not the source of the DNA obtained from the evidence item.
- "If the DNA profiles match, it is reported that the known individual may be the source of the evidentiary DNA.
- "Insufficient information (called inconclusive) may have been obtained from the DNA analysis, thus precluding reliable comparisons" (FBI Laboratory, 2008).

If an unknown DNA profile matches a reference DNA profile, the significance of the match is then calculated. "This is the probability of selecting an unrelated individual at random having an STR profile matching the DNA of the unknown profile. A random match probability can demonstrate that the DNA profile is so exceedingly rare (i.e., equal to or less than one in six trillion individuals) that it becomes reasonable for a forensic DNA expert to opine that the known individual is the source of the DNA obtained from the evidence item; this is referred to as source attribution" (FBI Laboratory, 2008). This topic is also discussed in Appendix E.6, DNA Profile Comparison.

3.2 Other Related DNA Biotechnologies

Examples of other technologies that enhance the efficiency and effectiveness of DNA processing include robotic platforms, Laboratory Information Management Systems (LIMS), portable DNA typing units, and mass spectrometry equipment. Each of these biotechnologies is briefly described below.

- Robotic platforms are instruments that automate various forensic procedures such as DNA extraction, quantitation, and PCR by mechanically manipulating samples, reagents, and/or other materials thereby reducing, but not eliminating, human intervention. With the growing number of samples (convicted offender samples, in particular) submitted to and processed by forensic laboratories, robotic platforms can improve sample handling precision; decrease the potential for contamination errors; and reduce casework backlog and analysis processing costs, time and labor. By automating these functions, forensic caseworkers can focus on activities requiring human intervention (i.e., evidence processing, case management and administration, decision-making, interpretation, reporting, technical and administrative review, testimony, execution of quality assurance measures, etc.). Several examples of robotic platforms include Star line (Hamilton Company), Freedom Evo (Tecan), Biomek (Beckman), and Wallac DBS Puncher (Perkin Elmer).
- LIMS are designed to: 1) manage and track information for samples, user's activities, instruments (including calibration, quality control, and validation), reagents and quality controls, and standard operating procedures; and 2) automate processes for workflow, audit trailing, invoicing, and other administrative functions. Examples of LIMS include Sample Tracking and Control System (STaCS; Anjura), SampleManager (Thermo Scientific), STARLIMS (Starlims Technologies), and SQL-LIMS (AB).

- Portable DNA typing units are handheld or vehicular laboratory components manufactured to perform onsite DNA sample typing. Although portable units are not used for forensic casework, the concept of portable units and the "mobile laboratory" may offer several advantages for applications such as mass disaster sample processing and biological threat operations (FBI Laboratory, 2008).
- Although mass spectrometry (discussed in Section 3.1.6, DNA Separation) is not currently used in the forensic community, two mass spectrometry techniques are being explored for future forensic applications: 1) matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) to identify SNPs of mtDNA, and 2) electrospray-ionization (ESI) for forensic STR and mtDNA analysis, using the Ibis Biosciences' T5000 Biosensor System[2] (NIST, 2008).

Application of new biotechnologies within the U.S. law enforcement community is subject to quality assurance requirements and operational procedures that govern and protect the integrity and security of the evidence and any DNA profiles submitted to the NDIS, discussed in Section 6.1, U.S. Federal.

[2] IBIS Biosciences, Forensics, 2008, available from: http://www.ibisbiosciences.com/pages.asp?ID=42; Internet; accessed May 2008.

4 Database Core Loci, Technologies, and Search Services

DNA biometric methods and technologies can be used to generate and directly compare two sample profiles. However, the true power of DNA as a biometric is realized when searching against a database of stored sample profiles. A key feature of this approach is the standardization on methods, kits, loci, and systems. Two broad classes of databases and search services are: 1) national and 2) public.

4.1 National

National DNA databases and search systems are primarily used by governments for law-enforcement and forensics purposes. Programs are as follows:

- NDIS/CODIS –The U.S. national DNA database and software, installed in 180 U.S. laboratories and 44 international laboratories in 30 foreign countries(Callaghan, 2009)
- National DNA Database (NDNAD) –The U.K. DNA database, the first established in Europe
- European Standard Set/ Interpol Standard Set of Loci (ESS/ISSOL) –The set of loci defined by Interpol (International Criminal Police Organization), used by 186 member countries to exchange DNA profile information. Note that the ESS is identical to ISSOL (Interpol, 2002)
- European Union (E.U.) Convention –Convention for DNA exchange that requires six of the seven loci in the ESS/ ISSOL and identifies additional loci that may be exchanged. This convention is based on the Prüm Treaty, discussed in detail in Section 5.3, International.

The following sections briefly detail what these national databases exchange and store.

4.1.1 STR Databases

There are several core sets of STR loci currently in use around the world. Table 4-1 documents the autosomal (not the sex chromosomes; see Sections 2.3, Y Chromosome STRs, and 2.4, X Chromosome STRs) STR loci most commonly used in national, commercial, and public DNA databases worldwide.

Table 4-1. Autosomal STR Loci

Locus[3]	Chromo-some	NDIS/ CODIS	NDNAD	ESS/ ISSOL	EU[4]
D3S1358	3	Y		Y	Y
VWA	12	Y	Y	Y	Y
FGA	4	Y	Y	Y	Y
D8S1179	8	Y	Y	Y	Y
D21S11	21	Y	Y	Y	Y
D18S51	18	Y	Y	Y	Y
D5S818	5	Y			Y
D13S317	13	Y			Y
D7S820	7	Y			Y
D16S539	16	Y	Y		Y
THO1	11	Y	Y	Y	Y
TPOX	2	Y			Y
CSF1P0	5	Y			Y
AMEL[5]	X,Y	Y	Y	Y	Y
Penta D	21	Y[6]			Y
Penta E	15	Y[6]			Y
D2S1338	2	Y[6]	Y		Y
D19S433	19	Y[6]	Y		Y
FES	15				Y
F13A1	6				Y

[3] Originally, STRs were named after the genes in which they were found. However, the STRs are now given "D" segment identifiers that are chromosome-specific identifiers in the form of "D[Chromosome]S[Number]" by the Genome Database; for example "D3S1358" is an STR on chromosome "3" (GDB, 2008). Section 9.4E.5.2 describes in more detail the naming conventions used for the STR alleles. For more information about the loci, such as chromosomal position and number of alleles, see (Butler(a), 2005).

[4] Accepts all listed loci, but search is based on ESS/ISSOL.

[5] AMEL (Amelogenin) is not an STR, but is included for completeness here because many STR databases and systems use it for gender identification (sex-typing). There is a six base pair deletion that occurs on the X-chromosome as compared to the Y-chromosomes, which allows Y and X to be distinguished.

[6] Accepted by CODIS at the NDIS but is not one of the 13 core loci that are searched.

Locus[3]	Chromo-some	NDIS/ CODIS	NDNAD	ESS/ ISSOL	EU[4]
SE33 (ACTBP2)	6				Y
CD4	12				Y
GABA[7]	15				Y

4.1.2 Y Chromosome Databases

Table 4-2 documents the Y Chromosome STR (Y-STR) loci types defined in the *Y Chromosome Haplotype Reference Database, Minimal Haplotype (YHRD minHt)* and the *Y Chromosome Haplotype Reference Database, Extended Haplotype (YHRD extHt)*. The YHRD minHt, endorsed by the European Network of Forensic Science Institutes (ENFSI) and accepted as the European standard (Martin P. , 1998), is an international database of polymorphic Y chromosomal sequences used for typing male DNA for forensic and genealogical purposes as well as for kinship testing (Willuweit & Roewer, 2007). The YHRD extHt specifies additional Y-STR loci to allow more discriminatory comparisons (Willuweit & Roewer, 2007). In addition, seven STRs (but nine amplified products) are required for database entry (minimal haplotype loci, minHt) and recommended for court use (Pascali, Dobosz, & Brinkmann, 1998). However, the Scientific Working Group on DNA Analysis Methods (SWGDAM) recommends 11 loci, which includes both minHt and extHt loci, for forensic casework in the U.S. (SWGDAM(a), 2004) Note that neither the Prüm Convention nor the ESS/ISSOL standards specify Y-STR loci and neither the NDNAD nor the NDIS currently collect Y-STR haplotype information. However, the Next Generation CODIS (NGCODIS) will store the SWGDAM-recommended Y-STR loci (FBI, 2008).

Table 4-2. Y-STR Loci

Locus	YHRD minHt	YHRD extHt	SWGDAM Recommended Loci
DYS19/394	Y		Y
DY385 a/b	Y		Y

[7] $GABA_A$ receptor b3 locus is listed in the Prüm implementing agreement treaty (Council of the European Union(a), 2007).

Locus	YHRD minHt	YHRD extHt	SWGDAM Recommended Loci
DYS389 I/II	Y		Y
DYS390	Y		Y
DYS391	Y		Y
DYS392	Y		Y
DYS393	Y		Y
DYS438		Y	Y
DYS439		Y	Y

4.1.3 SNPs Database

SNPs are currently not used for forensics. One issue is that SNPs require a larger number of loci over those required for STRs for the same confidence of matching. In addition, SNPs that are not linked to phenotypic traits and diseases (identification SNPs) are needed. The "SNPs for the Forensic Identification of Persons Consortium" (SNPforID; http://www.snpforid.org) is actively developing a set of loci and kits for forensics use. This is being evaluated by the European DNA Profiling (EDNAP) Group, a working group of the International Society for Forensic Genetics (ISFG) (Sanchez, J.J., Borsting, C., et al, 2008).

4.1.4 mtDNA Databases

The FBI Laboratory's DNA Analysis Unit II (DNAAU-2) maintains the SWGDAM mtDNA Population Database and search tool, available as a downloadable executable online (http://www.fbi.gov/hq/lab/fsc/backissu/april2002/miller1.htm) to non-CODIS participants for forensic analyses and comparison against 4,839 forensic profiles. The CODIS' mtDNA Popstats Population Database contains more mtDNA profiles and is available only to authorized law enforcement laboratories.

4.2 Public

There is a growing for-profit and not-for-profit (mostly academic) sector offering genealogy and ancestry services ranging from simple public databases to detailed ancestry analysis. Additionally, companies offer database resources and search services related to DNA profiling. Internet-based services geared for genealogy and ancestry research are starting to make 'home testing' possible for certain types of analyses. However, recent press reports have questioned the scientific basis for ancestry inference through these commercial tests, which may be primarily of entertainment value (Stahl, 2007). A sampling of some of these resources and services are presented below.

4.2.1 STR Databases

The ENFSI has undertaken the creation of a pan-European (autosomal) STR database (STR-base; http://www.str-base.org/), which is offered to the forensic community via the Internet with the principal demand of the highest conceivable quality of the data (Gill, Foreman, Buckelton, Triggs, & Allen, 2003). STR database uses 10 STR loci (D8S1179, D18S51, D21S11, FGA, TH01, vWA, D2S1338, D3S1358, D16S539, and D19S433) plus the gender marker Amelogenin. Autosomal STR DNA Database (http://www.strdna-db.org) is another example of a public autosomal STR database.

4.2.2 Y Chromosome Databases

The Y chromosome haplotype reference database (http://www.ystr.org/index.html) contains over 54,000 haplotypes from 477 populations (Willuweit & Roewer, 2007). In YHRD (release 23)[8], 97 percent of the population samples (n=53,075) are completely typed for the nine minHt loci and 47 percent (n=25,606) are typed to the eleven SWGDAM recommended loci (see Table 4-2. Y-STR Loci, and Section 2.3, Y Chromosome STRs).

YHRD supports two types of searching: GeoSearch and PopSearch. The former searches samples by their inputted continental affiliation, whereas the latter uses genetic relatedness to a metapopulation (and submetapopulations). These populations are based on prior knowledge (linguistic classification and posterior information based on the genetic distance parameters).

The U.S. Y-STR Database (http://www.usystrdatabase.org/) is another searchable database developed by The Center for Law Enforcement Technology, Training, and Research at the National Center for Forensic Science, University of Central Florida. The U.S. Y-STR Database (Release 1.0) has 13,906 haplotypes from five sources covering five populations (African American, Asian, Caucasian, Hispanic, and Native American). All samples are typed to the eleven SWGDAM recommended loci (SWGDAM(a), 2004) plus partially to six other loci (DYS437, DYS448, DYS456, DYS458, DYS635 and YGATAH4).

4.2.3 SNP Databases

HapMap (http://www.hapmap.org/) is an international effort to generate haplotypes using SNPs. Scientists and funding organizations from six countries are collecting SNPs and are contributing the resultant haplotypes, which are groups that share the same SNPs around a genetic region, to

[8] YHRD 3.0 (Release 24) is available with 59,004 profiles from 499 different populations. However, other statistics are not available on this new release.

the HapMap database. Although currently not used for forensics or biometric DNA identification, SNPS are being actively investigated for use in research and various other applications, such as theranostics where an individual's genetic profile can be used by researchers to tailor therapeutic agents and for personal genomics (see Appendix F). However, if SNPs are to be used in forensics applications, then SNPs that are not linked to phenotypic traits and diseases (identification SNPs) are needed.

4.2.4 mtDNA Databases

The EDNAP MtDNA Population Database (EMPOP; http://www.empop.org) is a Web-based Structured Query Language database(Parson, et al., 2004; Parson & Dür, 2007). The EMPOP project is a scientific collaboration among the DNA laboratory of the Institute of Legal Medicine, the Innsbruck Medical University, and other laboratories performing mtDNA research. The first release (Release 1; 2007) of this "pseudo-public" database held 5,173 mtDNA haplotypes with 4,527 high-quality data (full sequences) and 646 literature data (no appropriate raw sequence data is available) from mainly west Eurasian populations. EMPOP Release 2 added an additional 1,800 high-quality sequences of central and east Asian phylogeny.

4.2.5 Genealogy Databases

There is tremendous growth in commercial personal genetics, biopharmaceutical, and other areas not currently used for law enforcement and general forensic human identification purposes. The technologies used in these rapidly expanding marketplaces are based on those above. However, how they are implemented might not meet the same standards as expected or required for law enforcement. These services also have unknown implications for privacy concerns, since several of them are not U.S.-based and thus not subject to U.S. laws. Two of the largest genetic genealogy DNA database service providers are discussed below, Family Tree DNA and DNA Ancestry. Additional database initiatives are discussed in Appendix F, DNA Databases.

4.2.5.1 Family Tree DNA

Family Tree DNA (FTDNA; http://www.familytreedna.com), founded April 2000 by Mr. Bennett Greenspan, claims to be the first company to develop the commercial application of DNA testing for genealogical purposes. Today, the FTDNA database contains approximately 185,000 individual test records (two-thirds is Y chromosome DNA (Y-DNA) and the remaining is mtDNA) and 75,000 surnames. The FTDNA provides DNA testing services of DNA extractions from samples scraped from the cheek using an FTDNA-supplied in-home sample kit.

The FTDNA is the designated DNA testing company for the Genographic Project, a project sponsored by National Geographic to research the genetic roots of modern man and trace their migratory history, based on these results. To date the FTDNA has processed over 200,000 Genographic Project DNA tests. FTDNA also offers a free service called "Ysearch" (http://www.ysearch.org/) for comparison of Y-DNA from different companies. They currently have 50,242 unique haplotypes from 66,145 records (22 January 2009).

4.2.5.2 Ancestry.com

Ancestry.com (http://www.ancestry.com/), launched in 1997, provides similar DNA testing and database matching services for genealogy purposes. Ancestry.com, a member of The Generations Network, Inc., contains five billion names and 23,000 searchable databases, including the largest online collection of national and international historical records. Ancestry.com recently purchased Relative Genetics (http://www.relativegenetics.com) and now houses their database (http://dna.ancestry.com/).

Sorenson Genomics provides DNA genotyping and sequencing services (Y-DNA and mtDNA) for Ancestry.com using cheek swabs for collection of samples. 23andMe (see Appendix F) and Ancestry.com have partnered to extend access to genetic ancestry expertise.

5 Legislation

The use of DNA typing methodologies and sharing of DNA information within law enforcement in the U.S. and the E.U. is legislated and regulated to ensure stringent quality assurance of processing methods, and privacy and security of processing results. Legislation that prescribes the use of DNA for U.S. federal, state, and international law enforcement purposes is summarized below.

5.1 U. S. Federal

The FBI's CODIS Program was established in 1990 as a pilot project serving 14 state and local laboratories to automate the exchange and comparison of DNA profiles for the purpose of linking crimes with each other and to convicted offenders. In 1994, the *DNA Identification Act* formally authorized the FBI to establish and maintain a national DNA system for law enforcement purposes. Four years later, the FBI implemented the NDIS, the U.S. national system. Federal legislation governing CODIS, which comprises the NDIS as well as state and local systems (see Section 7.1.2, CODIS Program), includes the following:

- *DNA Identification Act of 1994*
- *Crime Information Technology Act of 1998*
- *Paul Coverdell National Forensic Sciences Improvement Act of 2000*
- *21st Century Department of Justice Appropriations Authorization Act*
- *DNA Backlog Elimination Act (2000)*
- *Justice for All Act of 2004*
- *DNA Fingerprint Act of 2005*
- *Reauthorization of the Debbie Smith DNA Backlog Grant Program (2008, part of the Justice for All Act of 2004).*

5.2 U. S. States

Individual state legislation varies regarding collection, analysis, and retention of DNA samples, and entry and expungement of DNA profile information in state and local databases. The FBI requires states to establish procedures to expunge offender and arrestee profiles from NDIS (in compliance with *The DNA Analysis Backlog Elimination Act of 2000* and *The Justice for All Act*) of 1) convicted persons whose qualifying convictions have been overturned or 2) arrested persons whose arrest has been dismissed or who have been acquitted regardless of whether or not the individual state DNA law requires it.

All states collect samples from convicted sex offenders and most states collect samples of all convicted felons. As of early 2009, approximately half of the states had enacted legislation that requires DNA sample retention. Of these states, most impose limitations, such as 1) for specific

timeframes or crime categories, 2) only after the legislation enactment date, or 3) upon petition to retest evidentiary samples to confirm matches to offender profiles(Innocence Project(a), 2009).

5.2.1 DNA Samples

Table 5-1 summarizes by offense type the number of states that have passed legislation requiring DNA sample collection ((DNAResource, 2008) and(NCSL, 2009)).

Table 5-1. State Legislation Requiring Collection of DNA Samples

State Legislation	Number of States
Felony Convictions	
All Convicted Felons	47
Convicted Sex Offenders	50
Juvenile Adjudications	32
Some Juveniles	35
Jail and Community Sentences	49
Retroactive Jail & Prison	40
Retroactive Probation & Parole	26
Not Guilty by Reason of Mental Defect or Guilty But Mentally Ill (GBMI)	9
Misdemeanor Convictions	
Certain Misdemeanors	34
Numerous Misdemeanors	4
Arrests	
Murder Arrestees	14
Sex Crimes Arrests	14
Burglary Arrests	14
All Felony Arrests	7

5.2.2 DNA Profile Databases

According to Privacy International, in 2006 (Privacy International, 2006):

- Three states required expungement of DNA profile information if an arrestee was acquitted or charges were dismissed.
- Thirty-eight states had expungement policies if a convicted individual successfully appealed.
- Three of the 38 states required automatic expungement of DNA profile information.
- Thirty-three of the 38 states required the appellant to initiate the expungement process.
- Only one state required the appellant be informed of his right to have his DNA profile expunged.

5.2.3 Post-Conviction DNA Exonerations

Forty-six states have approved legislation to allow post-conviction DNA testing(Innocence Project(b), 2009). To date, there have been 234 post-conviction DNA exonerations in the U.S(Innocence Project(c), 2009).

5.3 International

The Prüm Treaty, (Council of the European Union, 2005) signed in May 2005 by Belgium, Germany, Spain, France, Luxembourg, Netherlands, and Austria, established standards for cooperative exchange of DNA, fingerprint, and vehicle registration information, subject to national legal constraints, to aid in combating terrorism (Balzacq, 2006). The treaty intended that no information other than a "Hit/NoHit" response, and a reference identifier in the case of a hit, be returned to the requesting nation. Further communication between the states would "be conducted at the bilateral level upon the existing national legal and organizational regulations of the respective Parties' sites" (Council of the European Union(a), 2007), which provides protections for DNA profile data.

On June 12, 2007, a revised version of the Prüm Treaty was adopted by the E.U., requiring all E.U. nations to amend their domestic laws within three years to comply with the agreement. The E.U. law differs from the Prüm Treaty in that the E.U. law requires the exchange of requested, *matching and near-matching* DNA profiles, and permits member states to share suspects' personal data (Council of the European Union(b), 2007). The E.U. law also differs from Interpol policies that allow a participating country to impose restrictions to its profile information by other countries.

The information provided in following sections, summarizing the results of an E.U. funded research project, provides a quick overview of E.U. law and policy regarding the collection and removal of DNA profiles from each State's database and retention policies of the actual DNA samples (Van Camp & Dierickx, 2007).

5.3.1 DNA Samples

DNA samples are usually collected from one of three sources: unidentified stains from crime scenes, crime suspects, or convicted offenders. Most E.U. Member States have few restrictions regarding collection of crime scene stains and from suspects, but do not address the issue of consent. Most Member States allow collection without consent of convicted offenders with few restrictions. Likewise, most Member States allow collection from minors and mentally ill (Van Camp & Dierickx, 2007).

5.3.2 DNA Profile Databases

The first operating European national DNA database, the NDNAD, was established in England and Wales in 1995, and today contains over 4 million profiles of individuals and over 300,000 crime-related stains. Austria, Germany, The Netherlands, and France established national forensic databases in 1998; Finland and Belgium, 1999; and Denmark in 2000. Most Member States with a

few exceptions (i.e., Spain and Italy) have operational DNA databases. However, based on the E.U. Council adoption of the Prüm Treaty and recommendation that all E.U. Member States maintain DNA profile databases, these Member States plan to adopt legislation. It is important to note that although a country may not have implemented a database, it may in fact use forensic DNA profiling (Van Camp & Dierickx, 2007).

Most Member States have no restrictions on entry of crime scene stain DNA profiles and retention criteria for crime scene stains vary. Eleven of the 28 Member States studied enter DNA profiles for suspects for any offense and another five for serious offenses. Two Member States have no suspect database. In most cases, suspect DNA samples and profiles are retained unless acquitted or charges are dropped, although suspect DNA profiles are retained indefinitely in the NDNAD. Two Member States said they retained suspect DNA samples indefinitely; one said it retained suspect DNA samples for 75 years (Van Camp & Dierickx, 2007).

All Member States that responded to the study survey collected DNA profiles of convicted offenders. Three Member States retain convicted offender profiles indefinitely, most expunging profiles after an extended period following completion of the offender's sentence or after death. Four Member States retain convicted offender DNA samples indefinitely; one Member State retains these samples for 75 years (Van Camp & Dierickx, 2007).

A similar but less comprehensive survey regarding E.U. Member State policies of the ENFSI DNA Working Group representatives was performed in 2004 and updated in 2005 (Asplen & Lane, 2006).

6 Standards

6.1 U.S. Federal

At the federal level two key organizations, the FBI and NIST, define and publish standards to which NDIS-participating laboratories must comply. While not defined by a standards body, the Daubert standard was established as a result of a U.S. Supreme Court case. In the U. S. Supreme Court case "*Daubert vs. Merrell Dow Pharmaceuticals* (92-102), 509 U.S. 579 (1993)," the Court established criteria for determining if scientific evidence was reliable and hence admissible (Daubert, 1993):

1. Is the evidence based on a testable theory or technique?
2. Has the theory or technique been published and peer reviewed?
3. For a particular technique, does it have a known error rate and standards governing its operational use?
4. Is the underlying science generally accepted within a relevant community?

A robust DNA quality assurance program helps to ensure that 1) DNA evidence is accurate and reliable, and 2) underlying methodologies are standardized, validated, and accepted; hence, evidence presented in court satisfies these criteria.

The following sections present an overview of federal quality assurance standards for NDIS-participating DNA laboratories and federal testing standards established by NIST. Key international standards activities are also documented.

6.1.1 Quality Assurance Standards for Forensic DNA Laboratories

The DNA Identification Act of 1994 requires all public forensic labs receiving federal funds for DNA analysis or using CODIS software, as well as private forensic DNA labs to which DNA samples are outsourced, to comply with FBI quality assurance standards. Participating labs are audited according to the *Quality Assurance Audit for Forensic DNA and Convicted Offender DNA Databasing Laboratories* to ensure compliance.

Additional requirements are levied for those labs participating with the NDIS, documented in their respective Memorandum of Understanding, and agreed to and signed by the lab and the FBI. These requirements relate to submission and exchange of DNA data with the NDIS. Compliance to these standards is measured through audits. Failure to comply may result in suspension of analyst or laboratory participation in CODIS until deficiencies are remedied. Table 6-1 below provides an overview of these standards.

Table 6-1 Quality Assurance Standards for Forensic DNA Laboratories

Date & Authority	Publication Title, Purpose, and Comments
1989 Technical Working Group on DNA Analysis Methods (TWGDAM)	*Guidelines for a Quality Assurance Program for DNA Analysis*. Note: Renamed Scientific Working Group on DNA Analysis Methods (SWGDAM)
Oct1998; Updated Oct 2008, effective July 2009 DNA Advisory Board (DAB)	*Quality Assurance Standards for Forensic DNA Testing Laboratories*: Specifies governance for DNA laboratories that analyze crime scene evidence. 1999 Version: http://www.fbi.gov/hq/lab/html/testinglab.htm and http://www.fbi.gov/hq/lab/fsc/backissu/july2000/codis2a.htm 2008 Revision: http://www.fbi.gov/hq/lab/fsc/current/standards/2008_10_standards01b.htm
Apr1999; Updated Oct 2008, effective July 2009 DAB	*Quality Assurance Standards for Convicted Offender DNA Databasing Laboratories (Offender Standards):* Specifies governance for DNA laboratories that analyze convicted offender samples. 1999 Version: http://www.fbi.gov/hq/lab/html/databasinglab.htm and http://www.fbi.gov/hq/lab/fsc/backissu/july2000/codis1a.htm 2008 Revision: http://www.fbi.gov/hq/lab/fsc/current/standards/2008_10_standards01a.htm
2000	The DAB disbanded, returning responsibility for revisions or additions to the QAS to the SWGDAM
Jan 2001; updated Jul 2004 SWGDAM	*Quality Assurance Audit for Forensic DNA and Convicted Offender DNA Databasing Laboratories*: Sets forth additional guidance for local, state, and federal DNA laboratory auditors to ensure thorough and consistent compliance to the QAS. Note that the American Society of Crime Laboratory Directors/Laboratory Accreditation Board (ASCLD/LAB) and National Forensic Science Technology Center (NFSTC) contributed to the development of this document. 2001 Version: http://www.fbi.gov/hq/lab/fsc/backissu/jan2001/dnaaudit.htm 2004 Version: www.fbi.gov/hq/lab/fsc/backissu/july2004/pdfs/seubert.pdf
Ongoing	Memorandum of Understanding (MOU): Agreed and signed by the FBI and each NDIS participant.

Key tenets of the quality assurance standards that support evidence criteria recommended in the *Daubert* ruling require testing labs to do the following:

- Define, implement, and audit procedures to ensure the integrity, security, and quality of physical evidence including procedures to document chain of custody and to guard against cross-contamination of physical samples and related sample products throughout the process.
- Implement a documented program to maintain, service, and calibrate equipment and to undergo performance checks.
- Implement standard operating procedures for each analytical method employed by the laboratory that include specification of reagents, sample preparation and extraction methods, lab equipment, and controls.

- Use validated DNA analyses methodologies whose scientific principles have been peer reviewed and published.
- Implement guidelines for statistical interpretation of autosomal loci that meet recommendations documented in the National Research Council report entitled "The Evaluation of Forensic DNA Evidence" (1996) and/or a court-directed method.
- Implement procedures to ensure that all DNA profile information satisfies CODIS data integrity requirements.
- Ensure the privacy of all reports, case files, DNA records and databases, and to implement procedures to release such information only in accordance with state and federal law.
- Ensure vendor laboratories to which DNA samples have been outsourced comply with all standards and accreditations required by federal law, and maintain such documentation. The NDIS participating laboratory is required to review DNA test results of the vendor laboratory prior to uploading such data into CODIS.
- Check its DNA procedures annually or whenever substantial changes are made to a procedure against an appropriate and available NIST standard reference material or standard traceable to a NIST standard (see Section 6.1.2, NIST, below).

6.1.1.1 Validation

A key component of a quality assurance program is validation of methods and procedures. Validated methods and procedures ensure the reliability of scientific analytical data and, therefore, admissibility according to criteria established in the *Daubert* ruling. The quality assurance standards define stringent validation requirements for DNA analysis procedures to ensure that they are effective, reliable, robust, and predictable. These standards define two validation stages, development and internal as discussed below:

1. Developmental validation is the acquisition of test data and determination of conditions and limitations of a new or novel DNA methodology for use on database, known, or casework samples. Peer-reviewed publication of the underlying scientific principle(s) of a technology is required.
2. Internal validation is the accumulation of test data within the laboratory to demonstrate that established methods and procedures perform as expected in the laboratory.

6.1.1.2 Accreditation

To participate with NDIS, a laboratory must be both accredited and comply with the NDIS requirements set forth in their Memorandum of Understanding. The NDIS recognizes accreditation programs of The American Society of Crime Laboratory Directors/Laboratory Accreditation Board (ASCLD/LAB) and Forensic Quality Services–International (FQS-I).

The ASCLD/LAB, FQS-I, and the National Forensic Science and Technology Center audit forensic law enforcement DNA testing laboratories for compliance with the *Quality Assurance*

Audit for Forensic DNA and Convicted Offender DNA Databasing Laboratories. The ASCLD/LAB–International and the FQS-I accreditation programs also ensure compliance with International Organization for Standardization/International Electrotechnical Commission (ISO/IEC) 17025:2005 "General Requirements for the Competence of Testing and Calibration Laboratories."

6.1.2 NIST

The NIST publishes Standard Reference Materials (SRM) to support quality assurance and to ensure compatible measurements for calibration of human identity testing assays. The NIST SRMs also enable laboratories receiving accreditation under ISO 17025 requirements to demonstrate measurement traceability to a national reference material. Finally, commercial biotechnology companies use the SRMs to ensure their products comply with regulatory requirements. Table 6-2 documents NIST SRMs that apply to public and forensic DNA quantitation and analysis (NIST, 2009).

Table 6-2 NIST SRM

SRM	Use
2372: Human DNA Quantitation Standard	Human genomic DNA forensic quantitation materials.
2390: DNA Profiling Standard	Forensic and paternity quality assurance procedures and instructional law enforcement or non-clinical research purposes based on RFLP testing.
2391b: PCR-based DNA Profiling Standard	STR information for all genomic DNA samples in the SRM. The STR data includes the CODIS 13 core STR loci, including loci that were commercially available at the time of renewal certification.
2392: Mitochondrial DNA Sequencing	Quality control when performing the PCR and sequencing of human mtDNA for forensic identifications, medical diagnosis, or mutation detection. It may also be used as a control when amplifying PCR and sequencing any DNA.
2392-I: Mitochondrial DNA Sequencing	Compliments and adds another DNA template to SRM 2392 for the amplification and sequencing of human mtDNA. The selection of the HL-60 cell culture line for this additional DNA template was based on a suggestion from the FBI that this DNA would be particularly useful to the forensic community.
2394: Heteroplasmic Mitochondrial DNA Mutation Detection Std	Heteroplasmic mtDNA Mutation Standard.
2395: Human Y-Chromosome DNA Profiling Standard	Forensic and paternity quality assurance procedures for PCR-based genetic testing and for instructional law enforcement or non-clinical research purposes that involve the human Y-chromosome. SRM 2395 may be used to standardize nomenclature for the field of genetic genealogy

6.2 International

The U.S. actively participates with international standards committees focusing on forensic DNA, three of which are presented here.

6.2.1 ISFG/EDNAP

The EDNAP Group of the ISFG has members from 20 European laboratories actively contributing to forensic genetic research as well as performing DNA investigations in crime cases. The ISFG holds meetings every two years. The proceedings of the 2007 ISFG meeting are available at: http://www.fsigeneticssup.com; the main website is: www.isfg.org.

6.2.2 ENFSI

The ENFSI (www.enfsi.eu), which maintains relationships with the ASCLD, the American Academy of Forensic Sciences, Interpol, and other international forensic science organizations, promotes best practices and international standards for quality of ENFSI laboratories. In October 2007, the ENFSI, ASCLD, and other organizations signed a Memorandum of Understanding to build international cooperation focusing on joint research, methodology development and information exchange to enable forensic science facilities within these networks access to the latest crime fighting tools (ENFSI, 2008).

6.2.3 ISO

The ISO/IEC (www.iso.org), with members from 157 countries, develops and publishes international standards and related conformance testing methodologies. The ISO/IEC Joint Technical Committee 1 (JTC 1) focuses on information technology standards. Within the JTC 1, the subcommittee (SC) 37 is responsible for biometric information. This subcommittee is currently drafting the ISO/IEC 19794-14 “Biometric data interchange formats--Part 14: DNA Data” standard to facilitate international interchange of forensic DNA profile information to enable the biometric identification and verification of individuals. This standard includes STR, Y-STR, mtDNA, autosomal SNP, and Y-SNP DNA types; and incorporates data for missing persons and unidentified human remains as well as associated metadata.

The exchange format is based on the Extensible Markup Language (XML) schema. The FBI CODIS Unit plans to implement an import/export exchange capability that conforms to this standard. The ISO/IEC JTC1/SC37 is developing a related standard, 29109-1 “Conformance Testing Methodology for Biometric Data Interchange formats defined in ISO/IEC 19794,” to validate national data exchange format implementations. In addition, the ASCLD/LAB International and FQS-I test for conformance to ISO 17025:2005 “General Requirements for the Competence of Testing and Calibration Laboratories.”

7 Programs

7.1 U. S. Federal

The FBI is responsible for supporting law enforcement efforts ranging from the examination and processing of DNA evidence for criminal investigations and missing person identification to administering the CODIS Program. The FBI Laboratory also provides expert witness testimony at criminal proceedings; trains law enforcement agencies at the federal, state, and local levels; and sponsors the SWGDAM. Other federal programs include those funded for the NIST to develop, validate, and standardize technologies for the forensic DNA and human identity communities, and those funded for the Armed Forces Institute of Pathology DNA Identification Laboratory (AFDIL) to perform forensic DNA analysis research, and to maintain the U.S. military and other authorized personnel DNA reference databank. Key law enforcement programs are presented below.

7.1.1 FBI Laboratory's DNA Analysis Unit Programs

The FBI Laboratory's two DNA Analysis Units (DNAAU) examine evidence and provide expert witness testimony regarding the results of evidence examination for federal, state, and local law enforcement agencies. Generally speaking, the DNAAU-1 performs nuclear STR analysis whereas the DNAAU-2 performs mtDNA analysis. Specific DNAAU-1 programs include[9]:

- The Serology Program to identify types of serological (body fluid) stains.
- The PCR Program to analyze evidence using PCR-based STR typing and to compare the results to those of victims, suspects, and other stains.
- The Federal Convicted Offender Program to collect and analyze samples from federal offenders for comparison against other forensic samples in the NDIS for purposes of confirming a match to provide an investigative lead to the appropriate laboratory.
- The National Missing Persons Program to identify missing and unidentified persons. Training programs to train individuals 1) to analyze evidence specimens or 2) to interpret and testify on analysis results.

[9] Federal Bureau of Investigation, 2008, available from: http://www.fbi.gov/hq/lab/html/dnau1.htm; Internet; accessed May 2008.

The DNAAU-2 performs mtDNA analysis in forensic and missing person cases. Specific DNAAU-2 programs (http://www.fbi.gov/hq/lab/html/mdnau1.htm) include:

- The National Missing Person DNA Database for the identification of missing and unidentified persons in the NDIS.
- The SWGDAM mtDNA Population Database, which is an integrated software and database resource for forensic comparison.

In addition, the FBI Laboratory supports the casework of four regional crime laboratories.

7.1.2 CODIS Program

The CODIS Program provides CODIS software and support to authorized federal, state, and local forensic law enforcement laboratories. Today, CODIS software is installed at 180 federal, state, and local laboratories representing all States, Puerto Rico, the U.S. Army Crime Lab, and the FBI. Internationally, CODIS is installed at 44 international laboratories in 30 foreign countries(Callaghan, 2009). The CODIS software is provided to international labs to support search capabilities of their own databases. The FBI Laboratory's CODIS Unit is responsible for the development, installation, training, and support of the CODIS System.

7.1.2.1 NDIS

Operationally, NDIS is a three-tiered (local, state, national) hierarchical system of DNA storage indexes. Communication between labs is over the FBI's Criminal Justice Information Services wide area network using CODIS software.

The DNA profile searches are initially conducted at the level at which the DNA specimen profile data is first entered into the database. If this occurs at the local level, then the local lab searches its Local DNA Index System (LDIS) and saves all matches. Local labs cannot compare their specimen profile data directly with specimen profile data from other LDIS labs; these searches are performed at the state level.

The local lab also uploads its DNA specimen profile data to the state level for comparison at the State DNA Index System (SDIS), which contains specimen profile data from all of that state's LDIS labs as well as its own. The state audits the LDIS uploaded specimen profiles and reports all accepted and rejected specimen profiles to the local lab. For those accepted, the state searches its SDIS and reports matches to the appropriate local lab.

Similarly, state labs upload their DNA specimen profiles to NDIS weekly at predetermined times. The NDIS contains specimen profiles from all SDIS labs, and, therefore, includes specimen profiles of each state's LDIS labs. Searches across states are done at NDIS. Match messages are sent to the appropriate SDIS and LDIS labs. Approximately 85 percent of offender hits are in-state; 15 percent are interstate. Figure 7-1 depicts this process.

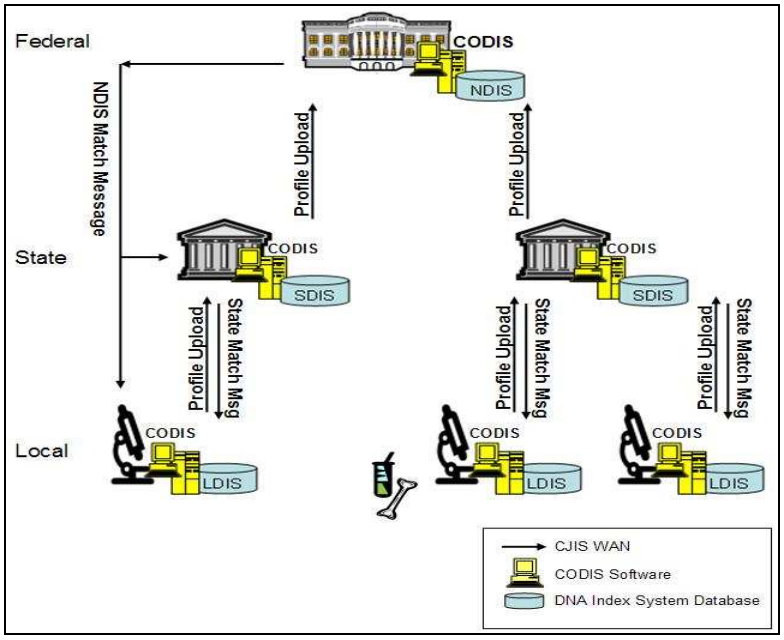


Figure 7-1. Overview of Current NDIS Architecture

Federal law requires that laboratories participating in NDIS are accredited and are externally audited against the FBI Director's Quality Assurance Standards at least every other year. Other stringent procedures are enforced by the NDIS-participating laboratory's CODIS Administrator to ensure 1) the quality, reliability, and security of data stored in NDIS is in accordance with state and federal laws, and NDIS operational procedures; and 2) matches are dispositioned according to NDIS operational procedures. If quality assurance measures or operational procedures are not satisfied, analyst or laboratory participation in NDIS may be suspended until deficiencies are remedied.

International laboratories may request a search of the NDIS through Interpol or an FBI Legal Attaché. Four criteria must be satisfied for request approval(Callaghan, 2007):

- The request must originate from a criminal justice agency.
- The request must be for a search of an authorized specimen category.
- The international laboratory must participate in a quality assurance program.
- The request must contain a sufficient number of core loci for effective searching.

If the international request meets these criteria, the NDIS Custodian searches the NDIS for matches. The results are reviewed and, if approved, forwarded by the NDIS Custodian to Interpol or the Legal Attaché.

7.1.2.2 CODIS Governance

The NDIS is the sole responsibility of the FBI Laboratory Division and is administered by the CODIS Unit. The DNAAU-1 of the FBI Laboratory operates as an SDIS within the system. Casework components of DNAAU-1 and DNAAU-2 reside at the LDIS level.

Three bodies advise the FBI Laboratory on important issues relating to operation of the NDIS (FBI Laboratory, 2008):

1. NDIS Procedures Board – Established in 1997 by the FBI Laboratory, the NDIS Procedures Board defines administrative and routine operational procedures for the NDIS including rules for uploading DNA profile information, searching missing persons' profiles, and maintaining and securing profile information.
2. SWGDAM – The SWGDAM is a body of forensic scientists that includes members from NIST. This body publishes reference material for forensic laboratories to measure the reliability of their equipment and DNA testing processes (*Guidelines for a Quality Assurance Program for DNA Analysis* (TWGDAM, 1989), revised (SWGDAM(b), 2004)). The SWGDAM also formulates and recommends revisions to the FBI Director regarding the federal quality assurance standards for forensic labs, discussed previously in Section 6.1.1, Validation.
3. CODIS State Administrators –The CODIS State Administrators are the administrative authorities for the daily operation of their respective state and local DNA database systems, ensuring that collection and exchange of DNA profile information complies with state laws. The CODIS State Administrators meet semiannually to provide recommendations to the FBI Laboratory on technical and operational issues.

Because members of these groups, the majority of whom are not FBI personnel, are involved in the 1) daily operations of their state and local DNA databases, and 2) quality assurance programs of their laboratories, they are uniquely qualified to advise the FBI.

7.1.3 NIST

The National Institute of Justice (NIJ), through the NIST Office of Law Enforcement Standards, funds NIST to evaluate and develop future technologies for forensic DNA typing. The NIST Human Identity Project team conducts a wide range of research projects each year that enable more effective means for analyzing forensic DNA samples. The team certifies standard reference materials, conducts interlaboratory studies, and produces new assays to enable improved recovery of information from degraded DNA. Standard information and training materials are made available on the NIST STRBase website (http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm).

Recent projects include work on mixture interpretation and low-level DNA, analysis of Y-chromosome variation, examination of rapid PCR amplification conditions, support of DNA quantitation issues, and variant allele sequencing. The NIST also provides resources and training materials for state and local laboratories, and assists in completion of other NIJ-funded projects. In

addition, a number of interlaboratory studies have been conducted by NIST over the past 15 years to study DNA quantitation, mixture interpretation, and instrument sensitivities between and among laboratories. Descriptions of many of these projects can be found on the NIST websites (http://www.cstl.nist.gov/biotech/strbase/NIJprojects.htm and http://www.dna.gov/research/nist). Since 2000, the NIST team has published almost 100 articles, given 250 presentations, and shared 30 training workshops with the forensic DNA community (http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm).

7.1.4 AFDIL

The AFDIL, part of the Armed Forces Institute of Pathology, was chartered in 1991 with the mission to maintain a DNA registry for U.S. soldiers in accordance with Army requirements for the use of DNA analysis in the identification of remains. The organization also provides education services and conducts research related to the forensic use of DNA. The AFDIL comprises three main sections: mtDNA, nuclear DNA, and a research section, and is responsible for generating mtDNA population data for both the EMPOP and SWGDAM mtDNA databases. The lab also supports state and local casework, and mass fatalities incidents. Additional information on AFDIL can be found online (http://www.afip.org/consultation/AFMES/AFDIL/index.html).

7.2 International

Interpol is the largest international police organization, comprising 186 member countries (Interpol, 2008). Forty-two of the more than 50 countries that maintain national DNA databases are members of Interpol (Interpol(a), 2007) including the U.S.

Interpol's I-24/7 global DNA network, known as the DNA Gateway, provides worldwide law enforcement authorities the capability to exchange DNA profile information to identify suspects. The Interpol also maintains a database of DNA profiles that is accessible online to all member countries upon adoption of the charter governing its secure use. The database contains more than "73,000 DNA profiles contributed by 46 member countries ... searches of the database by member countries led to 38 international hits during 2007"(Interpol(b), 2007). This system enforces rules that restrict query access to another participating country's profile information unless specifically authorized. To support international labs that use CODIS software, the CODIS Program has developed an ISSOL-compliant interface to enable export of DNA profiles in the Interpol format.

The "G8" is an international forum comprising Canada, France, Germany, Italy, Japan, Russia, the U.K., and the U.S. The G8 Law Enforcement subcommittee and Interpol recently demonstrated the direct exchange of requests for DNA information among national forensic laboratories in G8 countries using the I-24/7 secure network capability (Interpol(c), 2007).

8 Challenges and Technology Gaps

The great success and wide acceptance of DNA typing for purposes of human identification in government, national defense, commercial, research, and law enforcement sectors have:

- Led to increased demand for DNA processing, analysis and reporting.
- Resulted in rapid advances in research.
- Raised concerns about protecting the privacy of personal DNA information.

Stringent national quality assurance and control measures levy additional challenges in the law enforcement sector including, but not limited to:

- Hiring, training, and retention of fully qualified trained personnel.
- Acquisition of requisite infrastructure and equipment to support current processing demands and to test and validate new research methodologies for forensic use in accordance with national quality assurance standards.
- Definition, implementation, and testing of international standards to protect the privacy of DNA information.

To address these challenges, the DNA biotechnology community in general, and the law enforcement sector in particular, will require additional personnel, infrastructure, and funding. A brief discussion of some of these issues is presented below.

8.1 Resources

8.1.1 Personnel

The demand for qualified, fully trained forensic DNA analysts continues to outpace supply both within the U.S. and internationally for several reasons.

- Loss of qualified DNA analysts across all sectors to support international conflicts, such as those in Afghanistan and Iraq.
- Attrition of DNA analysts from state and local government labs to industry because of higher salaries or increased benefits.
- Existing demands within the law enforcement sector to test the existing backlog of DNA samples.
- Projected demands to process new samples taken from arrested individuals and detainees, required by the *DNA Fingerprinting Act of 2005*.

Once hired, to work within the law enforcement sector, an analyst must initially satisfy training requirements established in the federal quality assurance standards. The analyst then must

continue training to maintain certification and to become qualified in new methods and/or technologies as these methodologies are introduced into the lab.

8.1.2 Infrastructure

Acquisition of additional equipment and facilities may be required at federal, state and local levels to handle the increased demand for DNA testing to alleviate the current backlog of forensic samples and to meet projected increased demands to process new samples taken from arrested individuals and detainees, discussed in the previous section. To help mitigate this challenge, funding grants are made available to state and local labs to acquire additional infrastructure. Other infrastructure may be required to validate and implement newly discovered biotechnologies and methodologies that could be applied to forensic use, discussed below.

8.2 Technical Challenges

Advances in research and development to enhance the accuracy, robustness, and efficiency of DNA processing technologies, especially within the U.S. forensic community, require resources to identify, fund, test, validate for compliance with national quality assurance standards, and implement new capabilities and methodologies. The NIJ is seeking to fund research and development efforts of newly developed biotechnologies and methodologies that could be applied to the law enforcement sector. Areas of interest include (DOJ, NIJ, 2008):

1. Improved DNA analysis methods including rapid screening methods; nondestructive or minimally destructive methods for sample identification, collection, and/or extraction; and genetic screening methods.
2. New or improved methods to separate various components of DNA having mixtures of two or more sources.
3. Methods to identify and characterize various body fluids/cell-types in a DNA sample.
4. Improved tools for DNA preservation and/or storage that retain sample integrity.
5. Improved methods and tools for examining aged, degraded, limited, damaged, inhibited, or otherwise compromised DNA.
6. Identification and characterization of biological markers that can augment DNA source information such as markers that relate to the phenotype of the DNA contributor or that can increase the discriminatory power of DNA.
7. Novel approaches for genetic profiling including those that are portable, high-throughput, more informative, more sensitive, less susceptible to inhibition, non-PCR-based, and/or all-inclusive (e.g., DNA extraction through analysis) methods for analysis of biological evidence.

Examples of specific research that address some of these areas of interest are described below.

8.2.1 Biotechnology Systems

Continued advances in new and improved technologies and methodologies used in the commercial DNA biometrics arena may offer opportunities to increase DNA throughput for law enforcement. Recently, on-site real-time forensic STR analysis using an integrated lab-on-a-chip system was successfully demonstrated at a "at a mock crime scene prepared by the Palm Beach County Sheriff's Office (PBSO). Blood stain sample collection, DNA extraction, and STR analyses on the portable microsystem were conducted in the field, and a successful 'mock' CODIS hit was generated on the suspect's sample within 6 [hours] ... This demonstration of on-site STR analysis establishes the feasibility of real-time DNA typing to identify the contributor of probative biological evidence at a crime scene and for real-time human identification." (Liu, Yeung, Crenshaw, Crouse, Scherer, & Mathies, 2008).

Although rigorous operational procedures and quality assurance requirements[10] in place to protect the integrity and security of DNA profile information submitted to NDIS currently preclude use of on-site equipment within the law enforcement community, advances in this technology should be monitored for future forensic application (see bullets 1 and 7 above).

8.2.2 Mixed Samples

The efforts of scientists at the FBI Laboratory, NIST, and the SWGDAM have served as resources for the forensic DNA community in this area. The NIST has initiated several inter-laboratory studies to assess methodologies related to the interpretation of mixed samples (bullet 2 above). "Laboratories have instruments with different sensitivities. Different levels of experience and training play a part in effective mixture interpretation. Amount of input DNA makes a difference in the ability to detect the minor component (labs that put in 'too much' DNA actually detected minor components more frequently)" (Butler(b), 2005).

Although methods and supporting software that address DNA mixture analysis issues are commercially available, only Expert System software for the interpretation of single source (non-mixed) DNA profiles from offender samples has been authorized and validated for law enforcement use. Other methods and software either not yet validated or in development may demonstrate benefit in the future. For example, future use of electrospray-ionization mass spectrometry for forensic STR and mtDNA, still in research, is a potential prospect for quantitative typing of mixtures.

---

[10] For use in the forensic DNA community, new or modified technologies must be validated according to the *SWGDAM Revised Validation Guidelines*, published July 2004, discussed in Section 6.1.1.1.

As an example of research that is being performed in this area, refer to "A Multiplexed System for Quantification of Human DNA and Human Male DNA and Detection of PCR Inhibitors in Biological Samples" (Barbisina, Fanga, & etal, 2008).

8.2.3 Personal Genome Profiles

The rapid growth of public/pseudo-public genomic profile databases, described in Section 4.1.1, STR Databases, and Appendix F, DNA Databases, presents an unprecedented source of information about populations of individuals, and specific individuals and their families. Potentially, results obtained from searching these databases could assist state and local law enforcement to link individuals to crime scenes. However, search results may be questionable because the quality and integrity of the profile data resulting from unknown DNA sample collection and testing methods are unknown.

From a national perspective, profiles from public/pseudo-public genomic profile databases cannot be searched against profiles in NDIS because of privacy and federal legislation restrictions. For example, the sole purpose of accepting profile information from relatives of missing persons is to facilitate identification of those missing; hence, the search of profiles in the missing persons' index against other DNA profiles in the NDIS is strictly prohibited.

DNA biometric technologies for sequencing the full genome of an individual have advanced in commercial and research applications. Although the amount of material needed is currently several orders of magnitude above that required for current DNA forensic methods, this restriction is expected to abate in the future. With the National Institutes of Health pushing for the $1,000 genome and with companies now offering capabilities to fully sequence personal genomes, the future of DNA biometrics will need to better understand the ramifications of the availability of full genome profiles of individuals and related sequencing technologies.

8.3 Privacy

Like all biometrics used for authentication or identification of a person, the use of DNA as a biometric raises privacy issues regarding the protection of an individual's identity. However, unlike the other biometrics, DNA is unique in that it contains additional information beyond that which authenticates or identifies an individual. Although an in-depth overview of privacy issues related to commercial and research applications is beyond the scope of this assessment, an overview of some of privacy issues facing the U.S. government is presented below.

8.3.1 National

U.S. federal legislation specifically precludes the use of information in NDIS that could 1) determine personal physical characteristics[11] (e.g., eye or hair color, health or medical conditions) or 2) identify an individual (e.g., social security number, criminal history, or case-related information). Accordingly, nuclear and mitochondrial DNA information in NDIS has no known association with physical attributes. However, DNA information from non-law enforcement sources such as research or commercial laboratories could contain information that, at present or in the future, has the potential for determining or predicting physical attributes or health status. For this reason, exchange of DNA information between law enforcement labs and non-law enforcement sources through NDIS would raise privacy concerns and, therefore, is not permitted.

8.3.2 International

Privacy issues regarding exchange of DNA profile information within the international community are of special concern because foreign countries may not have the same legislative requirements for collecting, storing, validating, retaining and protecting the privacy of DNA profile information that are enforced within the U.S. Release of U.S. citizen information to the international community requires assurance that this information is protected by all participating states.

[11] Use of this information as metadata in missing persons' cases is allowed.

9 Recommendations

DNA databases and related search capabilities, processing methodologies and equipment, and relevant standards continue to evolve rapidly to meet the growing demands of industry and government. The previous sections 1) provide an understanding of both state-of-the-art and state-of-practice in the commercial, research and forensic DNA communities, and 2) document current technology gaps and challenges. This section presents high-level recommendations for FBI consideration to address these gaps and challenges within the law enforcement domain.

9.1 DNA Database and Search Services

The recent growth in commercial biotechnology coupled with interest in ancestry-related research has created a cottage industry of low cost DNA processing services and large databases. Complete sequencing of the individual may be cost-effective in the near future due to the rapidly declining costs of sequencing and new technologies being developed.

The FBI should consider these databases as having little forensic value at the national level although at state and local levels these databases may provide a starting point for limiting the scope of investigations. In addition, advances in data-mining techniques used in this industry should be monitored for potential governmental use. The FBI should also stay appraised of methods for the storage and searching against full sequences of the whole genome of an individual, mtDNA or STR loci while preserving privacy.

9.2 DNA Biotechnology Equipment and Methods

Driven by factors ranging from new legislative requirements to exonerating the innocent, increased demand for more timely and cost-effective DNA processing and analysis requires the FBI to take the following steps:

- Continue to actively engage in and support the research of new biotechnology equipment and methods.
- Identify opportunities to leverage and validate existing commercial and international DNA biotechnology capabilities.
- Explore unique approaches to combining these rapidly evolving biotechnologies into one integrated overall system to achieve even greater efficiencies.

The FBI should continue to make these objectives high priority and to ensure the availability of critical resources to achieve these objectives. Key resources include the following:

- Qualified personnel, including training to maintain qualifications, and incentives to foster personnel retention
- Additional staff for the CODIS Unit, DNAAU-1, and DNAAU-2 to meet expanding demands and technological advances of DNA analysis
- Laboratory facilities and storage

- Equipment and materials to process DNA and analyze DNA test results
- LIMS to manage samples, instruments, personnel, and other DNA processes
- Automation tools including robotic platforms and expert systems.

Federal grants to state and local governments will also be required to acquire infrastructure to support future demands.

9.3 DNA Standards

Continued active FBI participation on various international standards committees should be supported to ensure harmonization of international standards such as the ISO/IEC 19794 with American national standards such as the ANSI/NIST ITL 2-2008 "American National Standard for Information Systems— Data Format for the Interchange of Fingerprint Facial, & Other Biometric Information." Participation is also vital in those instances wherein international standards are adopted by the U.S. in their entirety, such as the ANSI/ISO/IEC 17025-2005 “General Requirements for the Competence of Testing and Calibration Laboratories.” To ensure CODIS interoperability and effective support of CODIS use internationally, continued FBI involvement in DNA international standards activities is essential.

From a national perspective, rapid developments in advanced DNA biometric technologies should be monitored, tested and validated for compliance with national standards defined by the SWGDAM and NIST (see Section 6, Standards). This may likely require additional resourcing, discussed in the previous section.

9.4 DNA Information Privacy and Security

The U.S. must continue to protect the privacy and rights of all U.S. individuals in the collection, storage, retention, and exchange of DNA information within the U.S. and international communities. The FBI should continue to provide leadership on international standards committees and in international DNA exchange program implementation efforts to ensure 1) international DNA biometric standards are published that will protect the privacy of U.S. DNA information and 2) systems are implemented and tested to ensure compliance with these standards.

Appendix A Acronyms

Acronym	Definition
AB	Applied Biosystems, Inc.
AFDIL	Armed Forces DNA Identification Laboratory
ASCLD/LAB	American Society of Crime Laboratory Directors/Laboratory Accreditation Board
Bp	Base Pair
CJIS	Criminal Justice Information Systems
CNV	Copy Number Variation
CODIS	COmbined DNA Index System
COE	Center of Excellence
CRS	Cambridge Reference Sequence
DAB	DNA Advisory Board
DNA	Deoxyribonucleic Acid
DNAAU	DNA Analysis Unit
DOJ	Department of Justice
E.U.	European Union
EDNAP	European DNA Profiling Group, a working group of the International Society for Forensic Genetics
EMPOP	European DNA Profiling Group (EDNAP) MtDNA Population Database
ENFSI	European Network of Forensic Science Institutes
ESI	Electrospray-Ionization
ESS	European Standard Set
extHt	Extended Haplotype
FBI	Federal Bureau of Investigation
FQS-I	Forensic Quality Services – International
FSS	Forensic Science Service
FTA	Fast Technology for Analysis (of Nucleic Acids)
FTDNA	Family Tree DNA
GBMI	Guilty But Mentally Ill
HV1/HV2	Hypervariable Regions 1 and 2
HVI/HVSI	Hypervariable Region I
HVII/HVSII	Hypervariable Region II
HVR	Hypervariable Region
HWE	Hardy-Weinberg Equilibrium
IEC	International Electrotechnical Commission
Indels	Insertions/deletions
Interpol	International Criminal Police Organization
ISFG	International Society of Forensic Genetics
ISO	International Organization of Standardization
ISSOL	Interpol Standard Set of Loci
JTC	Joint Technical Committee
LDIS	Local DNA Index System

Acronym	Definition
LE	Linkage Equilibrium
LIMS	Laboratory Information Management System
LR	Likelihood Ratio
MAF	Minimum Allele Frequency
MALDI-TOF	Matrix-Assisted Laser Desorption-Ionization Time-of-Flight
minHt	Minimal Haplotype
miniSTR	Mini-Short Tandem Repeat
mtDNA	Mitochondrial DNA
NDIS	National DNA Index System
NDNAD	National DNA Database
NIJ	National Institute of Justice
NIST	National Institute of Standards and Technology
PC	Phenol/Chloroform
PCR	Polymerase Chain Reaction
PE	Probability of Exclusion
qPCR	Quantitative PCR
rCRS	Revised Cambridge Reference Sequence
RFLP	Restriction Fragment Length Polymorphism
RNA	Ribonucleic acid
SABER	State-of-the-Art Biometrics Excellence Roadmap
SC	Subcommittee
SDIS	State DNA Index System
SNP	Single Nucleotide Polymorphism
SRM	Standard Reference Materials
STR	Short Tandem Repeat
STRBase	Short Tandem Repeat DNA Internet Database
SWGDAM	Scientific Working Group on DNA Analysis Methods
TEST	Technology Evaluation Standards Test
U.K.	United Kingdom
U.S.	United States
XML	Extensible Markup Language
Y-DNA	Y Chromosome DNA
YHRD	Y Chromosome Haplotype Reference Database
Y-STR	Y Chromosome Short Tandem Repeat

Appendix B Glossary

Term	Definition
Accredited Laboratory	A DNA laboratory that has received formal recognition that it meets or exceeds a list of standards, including the FBI Director's Quality Assurance Standards, to perform specific tests, by a nonprofit professional association of persons actively involved in forensic science that is nationally recognized within the forensic community in accordance with the provisions of the Federal DNA Identification Act (42 U.S.C. § 14132) or subsequent laws (DAB(a), 2008) and (DAB(b), 2008).
Allele	Any one of the alternative forms of a given gene (e.g., the ABO gene has three major alleles: A, B and O alleles). One of two slightly different versions of a gene that code for different forms of the same trait (www.everythingbio.com).
Allele Frequency	Often called gene frequency. A measure of how common an allele is in a population; the proportion of all alleles at one gene locus that are of one specific type in a population (www.everythingbio.com).
Amplification	The production of many DNA copies from one or a few copies, usually by PCR (www.everythingbio.com).
Assay	A signal amplification technology that detects the presence of specific nucleic acids by measuring the signal generated by many branched, labeled DNA probes (www.everythingbio.com).
Autosome	Any chromosome that is not a sex chromosome (www.everythingbio.com).
Base Pair (bp)	Two nitrogenous (purine or pyrimidine) bases (adenine and thymine or guanine and cytosine) held together by weak hydrogen bonds. Two strands of DNA are held together in the shape of a double helix by the bonds between base pairs. The number of base pairs is often used as a measure of length of a DNA segment (e.g. "500 bp") (www.everythingbio.com).
Capillary Gel Electrophoresis	The adaptation of traditional gel electrophoresis into the capillary using polymers in solution to create a molecular sieve, also known as replaceable physical gel. This allows analytes having similar charge-to-mass ratios to be resolved by size. This technique is commonly employed in SDS-Gel molecular weight analysis of proteins and the sizing of applications of DNA sequencing and genotyping (www.everythingbio.com).
Chromosome	A linear end-to-end arrangement of genes and other DNA, sometimes with associated protein and Ribonucleic acid (RNA). The form of the genetic material in viruses and cells. Humans typically have 46 chromosomes (pair of 23 chromosomes) (www.everythingbio.com).

Term	Definition
Commercial Test Kit	A preassembled kit that allows the user to conduct a specific DNA identification test (DAB, 1999)
Deoxyribonucleic Acid (DNA)	An antiparallel double helix of nucleotides (having deoxyribose as their sugars) linked by phosphodiester (sugar-phosphate) bonds to adjacent nucleotides in the same chain and by hydrogen bonds to complementary nucleotides in the opposite chain. The fundamental substance of which genes are composed (www.everythingbio.com).
Diploid	The state of having each chromosome in two copies per nucleus or cell. A cell or an organism having two chromosome sets or twice the haploid number, or an individual having two chromosome sets in each of its cells. This can be seen by the presence of two of each type of chromosome in a cell nucleus at interphase. In most cells, this is the typical number of chromosomes … In humans this number is 46. It is naturally twice the haploid number of 23 chromosomes contained in human eggs (ova) and sperm (www.everythingbio.com).
DNA Marker	Any unique DNA sequence which can be used in DNA hybridization, PCR or restriction mapping experiments to identify that sequence (www.everythingbio.com).
DNA Profile	The distinctive pattern of DNA restriction fragments or PCR products that can be used to identify, with great certainty, any person, biological sample from a person, or organism from the environment (www.everythingbio.com) (see also DNA Type).
DNA Sequence	The nucleotide sequence of a DNA fragment.
DNA Type	The genetic constitution of an individual at defined locations (also known as loci) in the DNA. A DNA type derived from nuclear DNA typically consists of one or two alleles at several loci (e.g., short tandem repeat loci). The DNA type derived from mitochondrial DNA is described in relation to the revised Cambridge Reference Sequence (Nature Genetics 1999, 23, 147) (SWGDAM(b), 2004).
DNA Typing	The analysis of sections of DNA for purposes of identification.
Double Helix	The normal structural configuration of DNA consisting of two helices winding about the same axis. The structure of DNA first proposed by Watson and Crick with two interlocking helices joined by hydrogen bonds between paired bases (www.everythingbio.com).
Electrophoresis	A technique for separating the components of a mixture of charged molecules (proteins, DNAs, or RNAs) in an electric field within a gel or other support. The movement of electrically charged molecules in an electric field often resulting in their separation (www.everythingbio.com).
E.U. Member State	A country that is a constituent of the European Union.

Term	Definition
Forensic DNA analysis	The process of characterizing DNA obtained from human biological samples (e.g., obtained from evidentiary material from crime scenes, suspects, victims, and convicted offenders) for application to questions of criminal law. The process results in the determination of a DNA type at defined locations in the DNA. (RVG); the process of identification and evaluation of biological evidence in criminal matters using DNA technologies (DAB(a), 2008) and (DAB(b), 2008).
Forensic DNA testing	The identification and evaluation of biological evidence in criminal matters using DNA technologies (DAB, 1999)
Gene	A basic unit of hereditary material; an ordered sequence of nucleotide bases that encodes a product (this product could be just RNA like rRNA or finally coding for a protein) (www.everythingbio.com).
Gene Pair	The two copies of a particular gene present in a diploid cell (one in each chromosome set).
Genetics	The study of inherited traits, genotype/phenotype relationships, and population/species differences in allele and genotype frequencies (DAB(a), 2008) and (DAB(b), 2008).
Genome	The entire complement of genetic material in a chromosome set (www.everythingbio.com).
Gonosome	One of the two chromosomes that specify an organism's genetic sex. Humans have two kinds of sex chromosomes, one called X and the other Y (http://ghr.nlm.nih.gov/glossary).
Haploid	The state of having one copy (one set) of each chromosome per nucleus or cell. A cell having one chromosome set, or an organism composed of such cells (www.everythingbio.com).
Haplotype	A haplotype, a contraction of the phrase “haploid genotype,” is a set of closely linked genetic markers present on one chromosome which tend to be inherited together (not easily separable by recombination). Another way to think about it is that a haplotype is half of a genotype. Some haplotypes may be in linkage disequilibrium … A haplotype can be identified by patterns of SNPs. Maps of SNPs (Haplotype maps) can be used to identify complex genetic variations (www.everythingbio.com).
Heterozygote	An individual having a heterozygous gene pair. A diploid or polyploid with different alleles at a particular locus (www.everythingbio.com).
Heterozygous Gene Pair	A gene pair having different alleles in the two chromosome sets of the diploid individual (e.g., Aa or A1A2) (www.everythingbio.com).
Homozygote	An individual having a homozygous gene pair. A diploid or a polyploid with identical alleles at a locus (www.everythingbio.com).
Homozygous Gene Pair	A diploid gene pair having identical alleles in both copies (e.g., AA or aa) (www.everythingbio.com).
Hybridization	The process of complementary base pairing between two single strands of DNA and/or RNA (SWGDAM(b), 2004).

Term	Definition
Hypervariable Region	Location within 1) nuclear DNA in which base pairs of nucleotides repeat or 2) the D-loop of mitochondrial DNA in which base pairs of nucleotides have. MtDNA has two hypervariable regions: HVR1 (low resolution) and HVR2 (high resolution).
Known sample	Biological material whose identity or type is established (DAB, 1999).
Locus (pl: Loci)	The position of a gene, DNA marker or genetic marker on a chromosome (www.everythingbio.com).
Lysis	The breaking open of a cell by the destruction of its wall or membrane (www.everythingbio.com).
Marker	Any genetic element (locus, allele, DNA sequence or chromosome feature) which can be readily detected by phenotype, cytological or molecular techniques, and used to follow a chromosome or chromosomal segment during genetic analysis (www.everythingbio.com).
Microsatellites	A DNA sequence from two to six nucleotides which are tandemly repeated from about five to 5,000 times (usual range 20–50 repeats). Microsatellites are found at many different loci in the genome, including coding and non-coding regions. A locus containing a microsatellite is often polymorphic because of variation in the repeat number (www.everythingbio.com).
Mitochondrial DNA (mtDNA)	The genetic material found in mitochondria, the organelles that generate energy for the cell. (www.everythingbio.com) mtDNA is inherited from the maternal side.
Multiplex System	A test providing for simultaneous amplification of multiple loci that is either prepared commercially or by a laboratory (DAB(a), 2008) and (DAB(b), 2008).
Multiplexing	An approach that co-amplifies more than one region of DNA using PCR to generate a DNA profile (J.M. Butler, NIST).
Nuclear DNA (nDNA)	Nuclear deoxyribonucleic acid (nDNA) is the DNA contained within a nucleus of cells.
Nucleotide	One of the structural components or building blocks of DNA and RNA. A nucleotide consists of a base (one of four chemicals: adenine, thymine, guanine, and cytosine) plus a molecule of sugar and one of phosphoric acid (www.genome.gov/glossary).
Platform	The type of analytical system utilized to generate DNA profiles such as capillary electrophoresis, real-time gel, and end-point gel instruments or systems (DAB(a), 2008) and (DAB(b), 2008).

Term	Definition
Polymerase Chain Reaction (PCR)	An enzymatic process by which a specific region of DNA is replicated, or amplified, during repetitive cycles to yield many copies of a particular sequence. A PCR cycle consists of the following three steps: 1) Denaturation or conversion of the double-stranded template DNA into its constituent single strands; 2) Annealing of primers to complementary sequences in the DNA template; and 3) Extension of the bound primers by a DNA polymerase (SWGDAM(b), 2004).
Polymorphism	The occurrence in a population (or among populations) of several phenotypic forms associated with alleles of one gene or homologs of one chromosome. Genetic polymorphism is defined as: a difference in DNA sequence among individuals, groups, or populations that gives rise to different forms such as the human blood groups (www.everythingbio.com).
Polymorphism (Genetic)	The occurrence in a population of two or more alleles at a genetic locus, when the frequency of the most common allele is less than 99 percent (SWGDAM(b), 2004).
Quantitative PCR	A method of determining the concentration of DNA in a sample by use of the polymerase chain reaction (DAB(a), 2008) and (DAB(b), 2008).
Reference Material	A material for which values are certified by a technically valid procedure and accompanied by or traceable to a certificate or other documentation that is issued by a certifying body (DAB, 1998).
Restricted Length Fragment Polymorphism (RFLP)	Restriction fragment length polymorphism (RFLP) is generated by cleavage by a specific restriction enzyme, and the variation is due to restriction site polymorphism and/or the number of different repeats contained within the fragments (DAB, 1998).
Restriction Enzyme	Any of a group of enzymes that cleave DNA at specific sites to produce discrete fragments, used especially in gene-splicing. (http://medical-dictionary.thefreedictionary.com/Restriction+Enzyme)
Ribonucleic Acid (RNA)	A single-stranded nucleic acid similar to DNA but having ribose sugar rather than deoxyribose sugar and uracil rather than thymine as one of the pyrimidine bases (www.everythingbio.com).
Short Tandem Repeat (STR)	DNA regions of two or more nucleotides in short repeated sequences (usually 2-6 base pairs in length) directly adjacent to one another.
Technology	The type of forensic DNA analysis performed in the laboratory, such as RFLP, STR, YSTR, or mitochondrial DNA (DAB(a), 2008) and (DAB(b), 2008).
Test Kit	A pre-assembled set of reagents that allows the user to conduct a specific DNA extraction, quantitation or amplification (DAB(a), 2008) and (DAB(b), 2008).

Term	Definition
Validation	A process by which a procedure is evaluated to determine its efficacy and reliability for forensic casework analysis and includes the following: 1) developmental validation is the acquisition of test data and the determination of conditions and limitations of a new or novel DNA methodology for use on forensic samples; 2) internal validation is the accumulation of test data within the laboratory to demonstrate that established methods and procedures perform as expected in the laboratory (DAB(a), 2008) and (DAB(b), 2008).

Appendix C References

23AndMe. (2008, September). *23andMe Democratizes Personal Genetics*. Retrieved October 6, 2008, from 23AndMe: https://www.23andme.com/about/press/20080909b/

AAFS. (2007). Using the Quality Assurance Standards for Forensic DNA Testing Laboratories to Distinguish the Unqualified Forensic DNA Experts From the Qualified Forensic DNA Experts. *Proceedings of the American Academy of Forensic Sciences, XIII*.

Anderson, S., Bankier, A. T., Barrell, B. G., Bruijn, M. H., Coulson, A. R., Drouin, J., et al. (1981). Sequence And Organization Of The Human Mitochondrial Genomes. *Nature , 290*, 457-465.

Andrews, R., Kubacka, I., Chinnery, P., Lightowlers, R., Turnbull, D., & Howell, N. (1999). Reanalysis and Revision of the Cambridge Reference Sequence for Human Mitochondrial DNA. *Nature Genetics , 23*, 147.

Applied Biosystems, I. (2006, December). *Product Bulletin Human Identification - AmpFlSTR® Identifiler® PCR Amplification Kit. Publication 112PB01-07*. Retrieved October 6, 2008, from Applied Biosystems, Inc.: http://www3.appliedbiosystems.com/cms/groups/applied_markets_marketing/documents/generaldocuments/cms_039944.pdf.

Asplen, C. H., & Lane, S. (2006, May). The Non-Forensic Use of Biological Samples Taken for Forensic Purposes: An International Perspective. *American Society of Law Medicine & Ethics Report* .

Balzacq, D. T. (2006). *The Treaty of Prüm and the Principle of Loyalty (ART. 10 TEC).* Centre for European Policy Studies (CEPS) .

Barbisina, M., Fanga, R., & etal. (2008). A Multiplexed System for Quantification of Human DNA and Human Male DNA and Detection. *Forensic Science International: Genetics Supplement Series , 1* (1), 13-15.

Biesecker, L. G., Bailey-Wilson, J. E., Ballantyne, J., Baum, H., Bieber, F. R., Brenner, C., et al. (2005). Enhanced: DNA Identifications After the 9/11 World Trade Center Attack. *Science , 310* (5751), 1122.

Borsting, C., et al. (2008). Performance of the SNPs forID 52 SNP-plex Assay in Paternity Testing. *Forensic Science International: Genetics , 2* (4), 292-300.

Budowle, B., & van Daal, A. (2008). Forensically relevant SNP classes. *Biotechniques , 44* (5), 603-610.

Butler(a), J. (2005). *Forensic DNA Typing: Biology, Technology, and Genetics of STR Markers* (2nd Edition ed.). Academic Press.

Butler(b), J. (2005). Mixture Interpretation: An Open Discussion on Issues and Ways to Increase Casework Throughput. *HITA Workshop.*

Butler, J. (2007). Short Tandem Repeat Typing Technologies Used in Human Identity Testing. *BioTechniques , 43*, Sii-Sv.

Butler, J., Buel, E., Crivellente, F., & McCord, B. (2004). Forensic DNA Typing by Capillary Electrophoresis: Using the ABI Prism 310 and 3100 Genetic Analyzers for STR Analysis. *Electrophoresis* (25), 1397-1412.
Butler, J., Coble, M., & Vallone, P. (2007). STRs vs. SNPs: Thoughts On The Future Of Forensic DNA Testing. *Forensic Sci. Med. Pathol , 3*, 200-205.
Callaghan, T. (2007). Combined DNA Index System Overview. *FBI Laboratory.* VA.
Callaghan, T. (2009). Updated NDIS/CODIS Statistics.
Cerri, N., Verzeletti, A., Gasparini, F., Bandera, B., & DeFerrari, F. (2006). Population data for four X-chromosomal STR loci in a population sample from Brescia (northern Italy). *International Congress Series , 1288*, 286-288.
Coble, M., & Butler, J. (2005). Characterization of New MiniSTR Loci to Aid Analysis of Degraded DNA. *Journal of Forensic Science , January*, 43-53.
Council of the European Union. (2005, July). *Prüm Convention.* Retrieved March 3, 2008, from ICTLEX: www.ictlex.net/wp-content/Prum-ConventionEn.pdf
Council of the European Union(a). (2007, January). *Administrative and Technical Implementing Agreement to the Prüm Convention.* Retrieved March 3, 2008, from gipfelsoli.org: www.gipfelsoli.org/rcms_repos/Antirepression/prum-implementing-agreement-annex.pdf
Council of the European Union(b). (2007, October). *Draft Council Decision on the implementation of Decision 2007/…/JHA on the Stepping Up of Cross-Border Cooperation, Particularly in Combating Terrorism and Cross-Border Crime.* Retrieved March 3, 2008, from Statewatch.org: www.statewatch.org/news/2007/oct/eu-prum-draft-13728-07.pdf
DAB. (1999). *The Quality Assurance Standards for Convicted Offender DNA Databasing Laboratories (Offender Standards).* Washington, D.C.: FBI.
DAB. (1998). *The Quality Assurance Standards for Forensic DNA Testing Laboratories (Forensic Standards).* Washington, DC: FBI.
DAB(a). (2008). *Quality Assurance Standards for DNA Databasing Laboratories Revised.* Retrieved October 22, 2008, from FBI Laboratory: http://www.fbi.gov/hq/lab/html/databasinglab.htm.
DAB(b). (2008). *Quality Assurance Standards DNA Testing Laboratories - Revised.* Retrieved October 22, 2008, from FBI Laboratory: http://www.fbi.gov/hq/lab/html/testinglab.htm.
Daubert v. Merrell Dow Pharmaceuticals, 509 U.S. 579 (1993). No. 92-102 (Certiorari to The United States Court of Appeals for the Ninth Circuit. Supreme Court of the United States June 28, 1993).
deCode. (2008). *Genotyping Services Overview.* Retrieved March 24, 2008, from deCode.com: http://www.decode.com/genotyping/index.php
DNA. (2008). *Advanced and Emerging DNA Techniques and Technologies.* Retrieved March 22, 2008, from President's DNA Initiative: http://technology.dna.gov/

DNAResource. (2008, November). *State DNA Database Laws Qualifying Offenses (As Of November, 2008).* Retrieved March 26, 2009, from DNA Resource: http://www.dnaresource.com/documents/statequalifyingoffenses2008.pdf
DOJ, NIJ. (2008, Nov 13). Solicitation: Forensic DNA Research and Development. D.C.
ENFSI. (2008). *Memorandum signed for International Cooperation in Forensic Science.* Retrieved March 5, 2008, from European Network of Forensic Science Institutes: http://www.enfsi.eu/page.php?uid=1&nom=153
FBI. (2008). Handbook of Forensic Services Revised. *Forensic Science Communications , 10* (e).
FBI Lab. (2009). *CODIS — NDIS Statistics*. Retrieved Jan 17, 2009, from FBI Web Site: http://www.fbi.gov/hq/lab/codis/clickmap.htm
FBI Laboratory. (2008, October). State of the Art Biometrics Excellence Roadmap. Technology Assessment: Volume 3 (of 3). DNA - Draft Comments.
FBI. (2000, July). Quality Assurance Standards for Forensic DNA Testing Laboratories and for Convicted Offender DNA Databasing Laboratories. *Forensic Science Communications* .
GDB. (2008). *Submitting Data to GDB.* Retrieved March 29, 2008, from The GDB Human Genome Database: http://www.gdb.org/gdb/SubmissionDirections.html
Gill, P., Brenner, C., Buckleton, J., Carracedo, A., Krawczak, M., Mayr, W., et al. (2006). DNA Commission of the International Society of Forensic Genetics: Recommendations on the Interpretation of Mixtures. *Forensic Science International , 160* (2-3), 90-101.
Gill, P., Foreman, L., Buckelton, J., Triggs, C., & Allen, H. (2003). A comparison of adjustment methods to test the robustness of an STR DNA database comprised of 24 European populations. *Forensic Science International* (131), 184-196.
Gusmao, L., Butler, J., Carracedo, A., Gill, P., Kayser, M., Mayr, W., et al. (2005). DNA Commission of the International Society of Forensic Genetics (ISFG): An Update of the Recommendations on the Use of Y-STRs in Forensic Analysis. *Forensic Science International* (157), 187-197.
IATA. (2008). Dangerous Goods Regulations. *49th Edition 2008 International Air Transport Association PE 602 and 650* .
Innocence Project(a). (2009, March). *Preservation of Evidence.* Retrieved March 26, 2009, from Innocence Project Web Site: http://innocenceproject.org/Content/253.php.
Innocence Project(b). (2009, March 17). *Mississippi and South Dakota Adopt DNA Access Laws.* Retrieved March 26, 2009, from Innocence Project Web Site: http://www.innocenceproject.org/Content/1912.php
Innocence Project(c). (2009, March). *Facts on Post-Conviction DNA Exonerations.* Retrieved 26 March, 2009, from Innocence Project Web Site: http://www.innocenceproject.org/Content/351.php.
Interpol. (2008, September). *About Interpol.* Retrieved October 6, 2008, from Interpol.int: http://www.interpol.int/public/icpo/default.asp

Interpol. (2002, April). *DNA Profiling–Worldwide Mostly Used Loci for Criminal Investigations/Identification.* Retrieved October 6, 2008, from Interpol.int: http://www.interpol.int/Public/Forensic/dna/loci.asp
Interpol(a). (2007, August). *Interpol's Global Police Communications System: I-24/7.* Retrieved March 4, 2008, from Interpol.int: http://www.interpol.int/Public/NCB/I247/default.asp
Interpol(b). (2007, August 8). *DNA Profiling.* Retrieved March 23, 2009, from http://www.interpol.int/Public/ICPO/FactSheets/FS01.pdf
Interpol(c). (2007, July). *Interpol Media Release.* Retrieved March 4, 2008, from Interpol.int: http://www.interpol.int/Public/ICPO/PressReleases/PR2007/PR200729.asp
Jeffreys, A., Brookfield, J., & Semeonoff, R. (1985). Positive Identification of An Immigration Test-Case Using Human DNA Fingerprints. *Nature* (317), 818-819.
Jeffreys, A., Wilson, V., & Thein, S. (1985). Individual-Specific Fingerprints of Human DNA. *Nature* (316), 76-79.
Jones, P. (2004). DNA Forensics: From RFLP to PCR-STR and Beyond. *Forensic Magazine* (Fall).
Kline, M., Duewer, D., Redman, J., Butler, J., & Boyer, D. (2002). Polymerase Chain Reaction Amplification of DNA from Aged Blood Stains: Quantitative Evaluation of the 'Suitability for Purpose' of Four Filter Papers as Archival Media. *Anal Chem , 15* (74(8)), 1863-9.
Liu, P., Yeung, S. H., Crenshaw, K. A., Crouse, C. A., Scherer, J. R., & Mathies, R. A. (2008). Real-time Forensic DNA Analysis at a Crime Scene Using a Portable Microchip Analyzer. *Forensic Sci Int Genet , 2* (4), 301-9.
Martin, P. (1998, June). *DNA Profiling In Crime Investigations: An European Overview.* Retrieved March 30, 2008, from Promega, Inc : http://www.promega.com/geneticidproc/ussymp10proc/content/17martin.pdf
Martin, P., Hermann, S., & Schneider, P. (2001). A Brief History of the Formation of DNA Databases in Forensic Science Within Europe. *Forensic Science International* (119), 225-231.
McCurdy, L. D., Gioeni, L., L.J., Pennella, T., Fisher, C., Isenberg, A., et al. (2007, August). Multiplex PCR Electrospray-ionization Mass Spectrometry (ESI-MS): Application to Forensic Mitochondrial DNA Examinations. *Forensic Science International: Genetics Supplement Series 1* .
National Geographic. (2009, Jan 26). *The Genographic Project - Atlas of the Human Journey*. Retrieved Jan 26, 2009, from National Geographic: https://genographic.nationalgeographic.com/genographic/lan/en/atlas.html
NCSL. (2008, January). *50 State Laws on DNA Data Banks: State Laws on DNA Data Banks - Qualifying Offenses, Others Who Must Provide Sample.* Retrieved March 16, 2008, from National Conference of State Legislatures: www.ncsl.org/programs/cj/dnadatabanks.htm

NCSL. (2007, December). *December 2007. Post-Conviction DNA Motions.* Retrieved October 6, 2008, from National Conference of State Legislatures: http://ncsl.org/programs/cj/postconviction.htm
NCSL. (2009, February). *NCSL Web Site.* Retrieved March 26, 2009, from State Laws on DNA Data Banks - Qualifying Offenses, Others Who Must Provide Sample: http://www.ncsl.org/programs/cj/dnadatabanks.htm
NFSTC. (2008). *DNA Analyst Training.* Retrieved October 6, 2008, from National Forensic Science Technology Center: http://www.nfstc.org/pdi/Subject09/pdi_s09_m01_01_b.htm
NIH. (2004, March). *A Basic Introduction to the Science Underlying NCBI Resources–What is a Cell.* Retrieved March 29, 2008, from National Institutes of Health - National Center for Biotechnology Information : http://www.ncbi.nlm.nih.gov/About/primer/genetics_cell.html
NIJ. (2000, November). *The Future of Forensic DNA Testing: Predictions of the Research and Development Working Group.* Retrieved March 12, 2008, from National Criminal Justice Reference Service: www.ncjrs.gov/pdffiles1/nij/183697.pdf
NIST. (2009, Jan). *105.8 DNA Profiling (Solid Forms).* Retrieved Jan 2009, from https://www-s.nist.gov/srmors/tables/view_table.cfm?table=105-8.htm
NIST. (2008, October). Review Comments.
NRC. (1996). *The Evaluation of Forensic DNA Evidence.* Retrieved March 4, 2008, from National Academies Press: http://www.nap.edu/catalog.php?record_id=5141
Oguzturun, C., Thacker, C., & Syndercombe-Court, D. (2006). Population Study of Four X-Chromosomal STR Loci in the U.K. and Irish Population. *International Congress Series , 1288*, 283-285.
Parson, W., & Dür, A. (2007, June). EMPOP-A Forensic mtDNA Database. *Forensic Science International: Genetics* , 88-92.
Parson, W., Brandstätter, A., Alonso, A., Brandt, N., Brinkmann, B., Carracedo, A., et al. (2004). The EDNAP Mitochondrial DNA Population Database (EMPOP) Collaborative Exercises: Organisation, Results and Perspectives. *Forensic Science International , 139* (2-3), 215-226.
Parson, W., Niederstätter, H., Lindinger, A., & Gill, P. (2007, December 4). Y-STR Analysis on DNA Mixture Samples—Results of a Collaborative Project of the ENFSI DNA Working Group. *Forensic Science International: Genetics* .
Pascali, V., Dobosz, M., & Brinkmann, B. (1998). Coordinating Y-Chromosomal STR Research for the Courts. *International Journal of Legal Medicine , 112* (1).
PaternityExperts. (2008). *Basic Genetics.* Retrieved October 6, 2008, from PaternityExperts: http://www.paternityexperts.com/basicgenetics.html
Privacy International. (2006, February). *The United States and the Development of DNA Data Banks.* Retrieved March 18, 2008, from Privacy International: http://www.Privacy International.org/article.shtml?cmd%5B347%5D=x-347-528471
Puers, C., Hammond, H., Jin, L., Caskey, C., & Schumm, J. (1993). Identification of Repeat Sequence Heterogeneity at the Polymorphic Short Tandem Repeat Locus

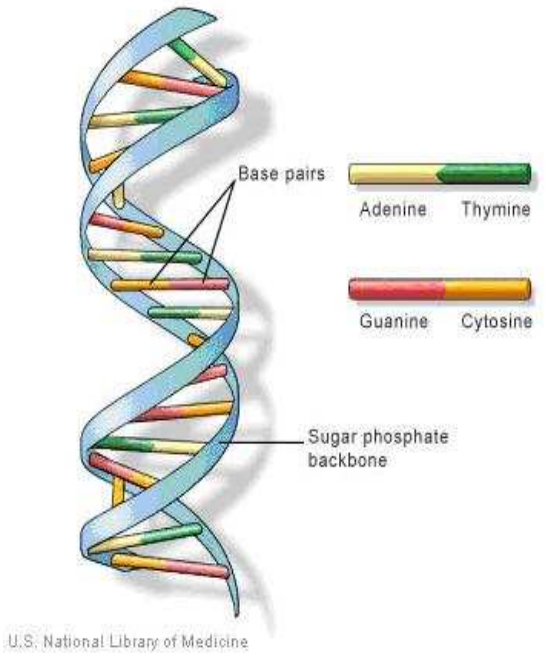
HUMTH01[AATG]. n and Reassignment of Alleles in Population Analysis by Using a Locus-Specific Allelic Ladder. *American Journal of Human Genetics , 53* (4), 953–958.
Roby, R., & Christen, A. (2007, September). *Validating Expert Systems: Examples with the FSS-i3™ Expert Systems Software [Profiles in DNA - Expert Systems]*. Retrieved October 4, 2008, from http://www.promega.com/profiles/1002/ProfilesInDNA_1002_13.pdf
Sanchez, J.J., Borsting, C., et al. (2008). Forensic Typing of Autosomal SNPs with a 29 SNP-Multiplex--Results of a Collaborative EDNAP Exercise. *Forensic Science International: Genetics , 2* (3), 176-183.
Stahl, L. (2007, October 7). Reconstructing the Family Tree. CBS 60 Minutes.
SWGDAM. (2003). Guidelines for Mitochondrial DNA (mtDNA) Nucleotide Sequence Interpretation. *Forensic Science Communications , 5* (2).
SWGDAM. (2000). Short Tandem Repeat (STR) Interpretation Guidelines. *Forensic Science Communications , 2* (3).
SWGDAM(a). (2004). Report on the Current Activities of the Scientific Working Group on DNA Analysis Methods Y-STR Subcommittee. *Forensic Science Communications , 6* (3).
SWGDAM(b). (2004). Revised Validation Guidelines. *Forensic Science Communications , 6* (3).
Sykes, B. (2003, Oct 9). *Mitochondrial DNA and human history.* Retrieved Jan 26, 2003, from Wellcome: http://genome.wellcome.ac.uk/doc_WTD020876.html
Szibor, R. (2007). X-Chromosomal Markers: Past, Present and Future. *Forensic Science International: Genetics , 1* (2), 93-99.
TWGDAM. (1989). *Guidelines for a Quality Assurance Program for DNA Analysis.* Washington DC: FBI.
USDT. (2007). *Code of Federal Regulation 49 Parts 173.196 and199.* U.S. Department of Transportation.
Vallone, P., Decker, A., & Butler, J. (2005). Allele Frequencies for 70 Autosomal SNP Loci with U.S. Caucasian, African-American, and Hispanic Samples. *Forensic Science International* (149), 279-286.
Van Camp, N., & Dierickx, K. (2007). National Forensic DNA Databases in the EU. *European Ethical-Legal Papers N°9* .
Walsh, P., Metzger, D., & Higuchi, R. (1991, October). Chelex 100 as a Medium for Simple Extraction of DNA for PCR Based Typing from Forensic Material. *Biotechniques* , 506-513.
Wambaugh, J. (1989). *The Blooding: True Story of the Narborough Village Murders.* UK: Bantam Press.
Willuweit, S., & Roewer, L. (2007). Y Chromosome Haplotype Reference Database (YHRD): Update. *Forensic Science International: Genetics , 1* (2), 83-87.
Wilson, M. R., Allard, M., Monson, K., Miller, K., & Budowle, B. (2002, October). Further Discussion of the Consistent Treatment of Length Variants in the Human Mitochondrial DNA Control Region. *Forensic Science Communications* .

.

Page Intentionally Blank

Appendix D Biology of DNA

DNA is packaged in chromosomes, threadlike microscopic bodies, found in the inner part of cells known as the **nucleus**. A typical human cell has 46 chromosomes, 23 chromosomes from the mother and 23 from the father. Two of the 46 are the sex chromosomes, X and Y. The remaining chromosomes are numbered "1" (largest) to "22" (smallest). Small amounts of DNA are also found outside the cell nucleus in cell **mitochondria**, which generate energy to ensure cells function properly. Mitochondria are inherited from the mother. DNA found in the cell's nucleus is called "**nuclear DNA**." DNA found in mitochondria is termed "**mitochondrial DNA (mtDNA)**." An organism's complete set of nuclear DNA is called its **genome**. ((NIH, 2004), (PaternityExperts, 2008))



A DNA molecule has two strands, which run in opposite directions and are coiled into a **double helix.** Each strand is composed of **nucleotides**. A **nucleotide** contains: 1) a nitrogenous base, 2) a phosphate molecule, and 3) a sugar molecule. A nitrogenous base is one of four chemicals: Adenine (A), Guanine (G), Thymine (T), or Cytosine (C).

The phosphate and sugar molecules within the nucleotides form the backbone of each strand. The nitrogenous base of a nucleotide on one strand is paired by a hydrogen bond with a base on the other strand to form a **base pair**, somewhat like the rungs of a ladder. Note that "A" always pairs with "T" and "G" with "C."

The order, or sequence, of base pairs within each chromosome determines genetic individuality. Each human has around three billion base pairs of which around 99 percent are identical in all individuals. The number of nucleotide bases that repeat also may vary. Regions of DNA that have highly repeated sequences are called **hypervariable regions** or **mini-satellite DNA**. About 97 percent of DNA has no known function and is referred to as "**junk DNA**." Repeated sequences in this region vary more than in gene regions and are most often used determine an individual's **DNA profile**, used in forensic analysis, thereby protecting individual privacy.

More detailed information regarding the biology of DNA can be found at the website for the National Center for Biotechnology Information, National Institutes of Health: (http://www.ncbi.nlm.nih.gov/About/primer/genetics_cell.html).

Page Intentionally Blank

Appendix E DNA Processing Biotechnologies

This appendix describes in more detail the various DNA processes and related biotechnologies presented in Section 3.1, DNA Processes and Related Commercial Biotechnologies.

E.1 DNA Extraction

There are two main types of extraction methods: 1) traditional organic methods, such as phenol/chloroform (PC) and 2) non-organic such as Chelex® and extraction kits, including Whatman, PLC (General Electric Healthcare) FTA® paper processing, Qiagen QIAAmp, and the Promega, Inc. DNA IQ.

Organic extraction is performed by sequentially combining several reagents (substances that chemically react to detect or prepare a substance) with the DNA sample material. The PC method separates the aqueous layer, containing DNA, from the organic layer, containing cell debris, thereby extracting the DNA sample. Organic extraction is useful for sources yielding small amounts of DNA, such as hair, teeth, or bones. However, this method has several disadvantages: it is manually intensive, requires more time than other methods, increases the possibility of contamination, and uses hazardous chemicals.

Chelex® extraction of DNA is another method by which a simple resin is added directly to sample material (Walsh, Metzger, & Higuchi, 1991), which simplifies the process thereby reducing processing time, decreasing the likelihood of contamination, and limiting safety risks. However, Chelex® extraction is a relatively harsh treatment that can harm some forensic samples. Differential lysis is a modified organic extraction method that splits female and male DNA, allowing separate amplification and profile identification. This method is extremely useful in sexual assault cases where male DNA can be recovered.

The FTA® paper extraction method isolates and preserves DNA. The sample source, such as whole blood, buccal/vaginal swabs, or other material, is applied directly to the FTA® paper. When the material having the DNA contacts the FTA® paper, the DNA sample cells lyse, disintegrating cell debris from the material and directly binding the DNA sample to the paper. The advantages of this method are 1) the sample remains stable at room temperature when stored long-term and 2) ease of use.

The QIAAmp® method uses a spin column containing a mixture of alcohol and lysate that is inserted into a tube containing both the DNA source and lysis buffer. Through a process of multiple washes and centrifuge, sample debris is removed and the purified DNA, eluted.

Two solid phase extraction methods use silica and/or magnetic bead technologies, and it is the combination that is used by forensic laboratories. In both cases DNA binds to the beads as the particular technology process washes away debris and elutes purified DNA. Silica requires high-salt concentrations whereas magnetic beads require a magnetized source such as a rack or tube holder. However, silica may give lower yields than magnetic beads due to filtering.

The DNA IQ System from Promega uses magnetic particles to prepare clean samples. This system can extract either a constant 100ng of DNA from samples from a variety of sources such as bloodstain cards, buccal swabs, or liquid blood, or purified DNA from casework sample types.

E.2 DNA Quantitation

Various DNA quantitation techniques include yield gels, spectrophotometry, fluorometry, slot blot hybridization, AluQuant®, and quantitative PCR (qPCR). These are presented in the following table.

Table E-1. DNA Quantitation Techniques

Quantitation Technique	Analysis Type	DNA Concentration Method (Determined By:)
Yield Gels	Agarose Gel Electrophoresis Analysis	Mixing the DNA sample with agarose gel, separating the DNA from the mixture by applying an electric current, and comparing the fluorescence intensity of the separated DNA to known DNA standards.
Spectrophotometry	Optical Density Analysis	Measuring the absorbance of the DNA at an optical density of 260 nanometers.
Fluorometry	Fluorescence Analysis	Binding a fluorescent dye to the DNA, exciting the fluorescent dye, and measuring the output fluorescence on a fluorometer. PicoGreen® is an example of one of the more common fluorescent dyes.
Slot Blot Hybridization	Hybridization Analysis	Binding the DNA to a nylon membrane, applying a probe that contains a chemical activator to the DNA, and detecting the hybridized complex via several different methods. This was the most common technique for quantitation until recently; QuantiBlot® was the preferred kit.
AluQuant®	Alu Analysis	Denaturing the DNA sample, incubating the DNA with AluQuant Enzyme and Probe solutions, and measuring the concentration after chemical reaction using a luminometer.

Quantitation Technique	Analysis Type	DNA Concentration Method (Determined By:)
Quantitative PCR	qPCR - Real Time PCR Analysis	Measuring the accumulation of fluorescent dyes as the DNA is amplified through a normal PCR process. The qPCR method is the most efficient and sensitive of the quantitation methods. Kits used in qPCR analysis include SYBR® Green and Fluorescent Resonance Energy Transfer, and can be applied on a number of different instruments such as AB 7500 and AB 7900.

E.3 DNA Amplification

The PCR amplification process is the most common method used today to amplify DNA. This process comprises three steps:

1. denaturing: heating DNA to a temperature at which the double-stranded DNA helix is split into two single strands
2. annealing: cooling the DNA and then adding primers (short exogenous DNA sequences that target specific areas for amplification) that anneal (bind) to the DNA
3. extension: heating the DNA to a temperature just below that which would cause denaturing (see first step above), causing the primer binding sites to extend or lengthen.

This process is repeated 20 to 40 times (cycles) depending on the amplification protocol. By repeating this process the DNA is amplified in an exponential fashion (two template helixes become four, four become eight, and so on). The primers used in the amplification of STR loci are tagged with fluorescent labels enabling their subsequent detection following DNA separation.

Examples of commercial PCR amplification kits include:

- AmpFlSTR® Identifiler™, Profiler Plus ID™, COfiler™, Y-Filer™, and MiniFiler™ offered by Applied Biosystems, Inc.
- PowerPlex® 16, Y, and ES provided by Promega, Inc.

Quality assurance measures conducted during manufacture of these kits, in addition to laboratory quality assurance, reduces potential for contamination. The FBI's Quality Assurance Standards for Forensic DNA Testing Laboratories (DAB, 1998) require reference samples to be processed separately (both in time and space) from crime scene samples to ensure against cross-contamination.

E.4 DNA Separation

Separation of alleles by capillary electrophoresis is carried out in three steps, injection, separation, and detection:

1. Injection of amplified material requires placing the sample, diluted in either water or formamide, at the end of a capillary (narrow tube), which is filled with a sieving polymer, and applying an electric current. The current forces the migrating DNA molecules to "stack up" as the molecules are injected, thus concentrating the DNA.
2. Separation occurs as the continued application of electric current causes the sieving polymer in the capillary to filter DNA molecules by size from smallest to largest.
3. Detection of amplified material is accomplished by exciting the fluorescent tags introduced during the PCR amplification process (with the specific loci primers from STR kits) using a laser and then capturing the image with an image sensor camera (a device that converts optical images to electrical output). These images are automatically camera-computerized for later analysis.

Two other techniques include slab gel and mass spectrometry. The slab gel method uses three processes: 1) PCR amplification, 2) polyacrylamide gel electrophoresis, and 3) visualization on a slab gel-based instrument. Although the slab gel method may still be used by a few laboratories, in general this method has been replaced by newer technologies. The Hitachi FM BIO series is an example of slab gel equipment.

Mass Spectrometry "is a technique that allows the detection of compounds by separating ions by their unique mass (mass-to-charge ratios) using a mass spectrometer. The method relies on the fact that every compound has a unique fragmentation pattern (mass spectrum). The sample is ionized; the sample ions are separated based on their differing masses and relative abundance." (DNA, 2008)

Two ionization techniques used in mass spectrometry are the MALDI-TOF and ESI. These techniques are first introduced in Section 3.1.6, DNA Separation.

Research efforts into the use of MALDI-TOF mass spectrometry to identify SNPs of mtDNA have been published widely.[12] A similar technique that results in more multiply-charged ions is

[12] Examples of research include: "A Rapid and Accurate Approach to Identify Single Nucleotide Polymorphisms of Mitochondrial DNA Using MALDI-TOF Mass Spectrometry," (http://www.ncbi.nlm.nih.gov/pubmed/18254712), and "SNPs and MALDI-TOF MS: Tools for DNA Typing in

ESI. Research using the ESI technique focuses on STR and mtDNA analysis. Use of ESI offers the advantages of quantitative typing of mixtures, quicker analysis throughput, automation, and increased sensitivity, while retaining a high degree of resolution (McCurdy, et al., 2007). The Ibis Biosciences' T5000 Biosensor System is an example of equipment used for research in this area.

E.5 DNA Profile Analysis

After completion of the separation process, the DNA is analyzed.

E.5.1 Electropherogram and Plot

The following paragraphs describe the loci and related standards that use these loci that are amplified by the AB AmpFlSTRIdentifiler PCR Amplification Kit, and present the resulting software electropherogram and plot. This material is quoted directly from the Applied Biosystem, Inc. "Product Bulletin Human Identification," which can be viewed at: http://www3.appliedbiosystems.com/cms/groups/applied_markets_marketing/documents/generaldocuments/cms_039944.pdf.

> "The AmpFlSTR® Identifiler® PCR Amplification Kit simultaneously amplifies 15 STR loci plus the Amelogenin gender-determining marker in a single, robust PCR assay. Well characterized tetranucleotide loci co-amplified in the Identifiler® Kit include the thirteen core STR loci as required for sample entry into CODIS: CSF1PO, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, FGA, TH01, TPOX, and vWA. Data generated from these loci also satisfy the recommendations of the European Network of Forensic Science Institutes (ENFSI) and Interpol organizations. Two additional tetranucleotide loci, D2S1338 and D19S433, provide concordance with the AmpFlSTR® SGM Plus® PCR Amplification Kit, which was developed in collaboration with the Forensic Science Service (FSS).
>
> "[Figure E-1, below, depicts a] GeneMapper® software electropherogram showing the AmpFlSTR® Identifiler® PCR Amplification Kit results for fifteen STR loci and the Amelogenin locus analyzed on the AB Prism® 3130 Genetic Analyzer. DNA fragments are labeled in 6-FAM™ dye (blue), VIC® dye (green), NED™ dye (yellow, depicted in black), and PET® dye (red). The GeneScan™-500 size standard is labeled with LIZ® dye (orange).

Forensic Paternity Testing and Anthropology." (see: http://www.astm.org/JOURNALS/FORENSIC/PAGES/JFS2004245.htm)

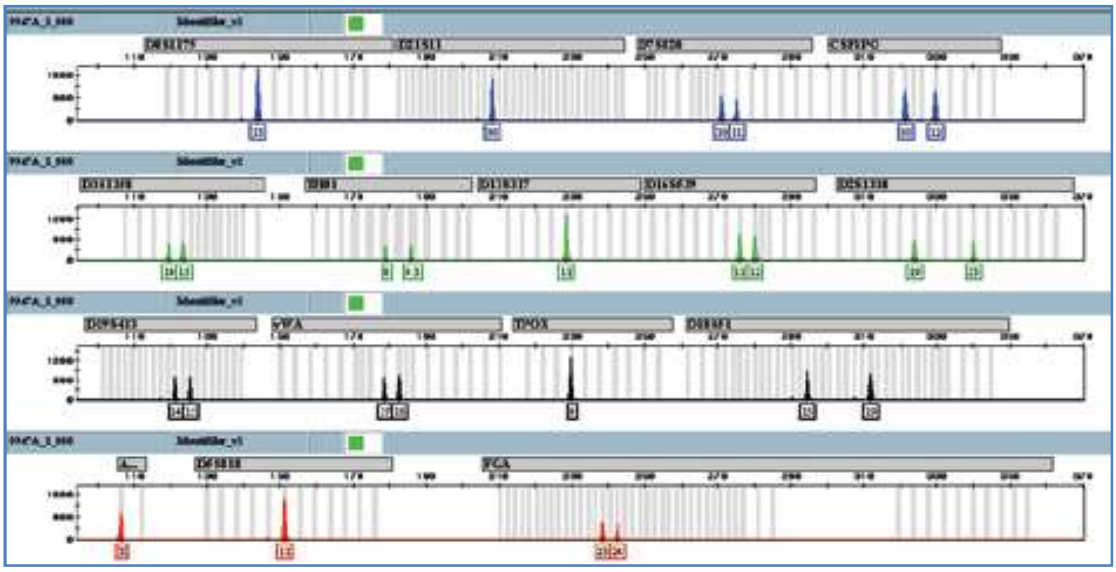


Figure E-1. AmpFlSTR® Identifiler® PCR Amplification Kit Results

"[Figure E-2 depicts a] GeneMapper® *ID* Software plot of the AmpFlSTR® Identifiler® Allelic Ladder analyzed on the 3100 Genetic Analyzer. The new PET® dye (red) used to label PCR products increases in-lane multiplexing throughput capabilities when used along with 6-FAM™, VIC®, and NED™ dyes. The AmpFlSTR® Identifiler® Allelic Ladder includes additional alleles for D18S51, D21S11, FGA, TH01, and vWA loci as compared to the AmpFlSTR® Profiler Plus and Cofiler® Allelic Ladders." (Applied Biosystems, 2006)

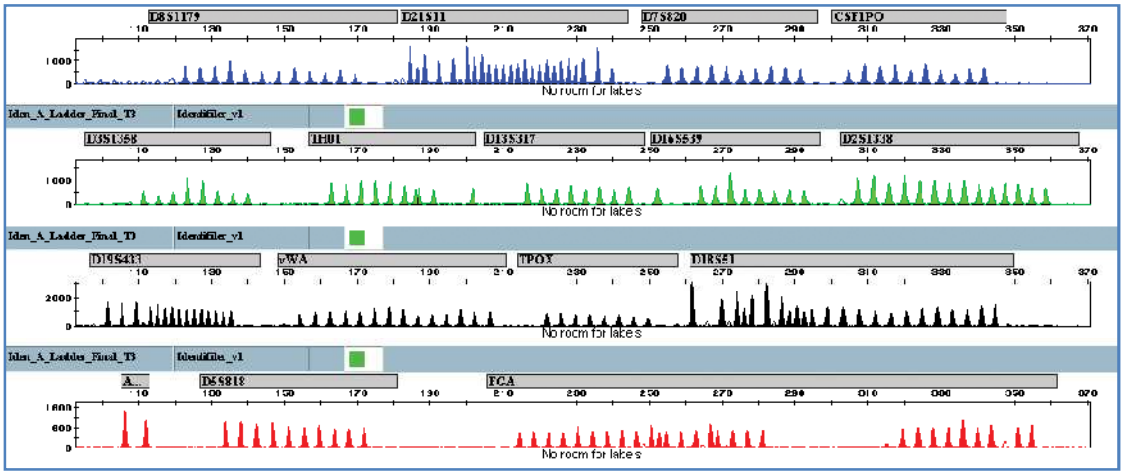


Figure E-2. AmpFlSTR® Identifiler® Allelic Ladder

The results of this analysis are then digitized to generate the DNA profile.

E.5.2 Text/Digital Representation of Sequences

Common digital representation of DNA sequences allows automated searching and is the basis for NDIS and international sharing of DNA profiles under law enforcement treaties or disaster victim identification and recovery efforts.

Originally, STRs were named after the genes in which they were found. However, the STRs are now given "D" segment identifiers that are chromosome-specific identifiers in the form of "D[Chromosome]S[Number]" by the Genome Database; for example "D3S1358" is an STR on chromosome "3" (GDB, 2008) (see Table 4-1). Generally, STR alleles are named after the number of full repeats that they contain. When an allele does not contain a standard repeat (referred to as a "variant" or an "off-ladder" allele), it is designated by the number of complete repeats and the number of base pairs of the partial repeat (SWGDAM, 2000). These two values should be separated by a decimal point; e.g., the "HUMTH01 STR 9.3" allele is relatively common in Caucasians (Puers, Hammond, Jin, Caskey, & Schumm, 1993). The "HUMTH01" is a tetranucleotide locus containing repeats of AATG. As such, the 9.3 allele is a base pair shorter than 10 repeat units. It has been shown that this is due to the loss of an adenine in the 7th repeat unit; i.e., [AATG]6[-ATG]1[AATG]3. Note that without sequencing the locus, the loss of a base pair could occur anywhere in the locus, resulting in the allele being called 9.3; for example [AATG]3[-ATG]1[AATG]6 is called 9.3. If the allele falls above the largest or below the smallest allele of the allelic ladder[13], then the allele designation should be greater than (>) or less than (<) the corresponding ladder allele.

The mtDNA sequences are compared to the rCRS using the nomenclature system set forth by the International Union of Pure and Applied Chemistry for DNA base call designation(Andrews, Kubacka, Chinnery, Lightowlers, Turnbull, & Howell, 1999; SWGDAM, 2003; Parson & Dür, 2007). Differences between the rCRS and the sample will be noted as polymorphisms with the nucleotide position and the DNA base difference from the rCRS (e.g., 16147 T; is a change in "C" to a "T" at position 16147 in the HVI region, see Figure 2-2). Insertions are described by noting the site immediately 3' (3 prime) to the insertion with respect to the light strand[14] followed by a point and sequential number for each inserted base starting at 1 (e.g., 315.1 C; this is a "C" inserted after nucleotide at 315).

---

[13] Allelic ladders are mixtures of the common alleles present in the human population for a particular STR marker and are used for accurate genotype determinations (Reviewed in (Butler(a), 2005)).

[14] mtDNA has two strands, a guanine-rich heavy strand and a cytosine-rich light strand. The light strand is the "inside" strand (the 3' to 5' strand, not the numbered 5' to 3' strand).

More complex variants, such as complex insertions/deletions (indels), should be coded in accordance with the guidelines proposed by Wilson et al (Wilson, Allard, Monson, Miller, & Budowle, 2002). However, these guidelines do not cover all possible variants and new guidelines are needed to capture these variants.

E.6 DNA Profile Comparison

DNA profiling is a well-established forensic tool that has held up to the Daubert principles for use in court. In the U.S., there are several approaches used to determine the probability of a match between a subject's sample profile and a profile stored in a database (reviewed in(NRC, 1996; NIJ, 2000; FBI, 2000). A summary of methods derived from population genetics to estimate how individuality arguments are constructed is provided below, with a brief overview of some of the general statistical procedures and issues. These procedures are for simple cases (i.e., a single-contributor case, statistically negligible population substructure, and not a mixture or relative samples).

E.6.1 Population Databases

The primary goal of generating a population database is to reliably estimate the frequency of occurrence of the common alleles in a population by sampling a portion of the population. It is generally accepted that a population sample size of 100-150 is usually sufficient to make projections about a larger population(Butler(a), 2005), chapter 20). It is important that loci are inherited independently so that allele frequencies can be multiplied in the form of the product rule generating match probabilities(Butler(a), 2005, p. 485).

Originally, alleles in populations were assumed to be in random proportions (independent of each other), which is known as Hardy-Weinberg equilibrium (HWE). Under HWE, it is assumed that the population randomly mates, which does not exist in real human populations. In general, departures from HWE are very minor and not statistically significant (NIJ, 2000).

Linkage equilibrium (LE) occurs when loci are sorted independently (randomly) from one another; the frequency of the DNA complete profile should be the product of the frequencies (product rule) at each individual locus. STRs show good agreement with being in LE and are suitable for practical forensics work (NIJ, 2000).

Since the U.S. population is not a single randomly-mating unit, matings are more likely to occur within a subpopulation than between subpopulations. This results in an increase in the frequency of homozygotes of one or more specific alleles with a decrease in the frequency of heterozygotes,

generating a substructure within the population. Because of this, populations are divided into subpopulations, each in HWE. A measure of the population subdivision is then empirically determined[15], usually expressed as "θ."

E.6.2 Minimum Allele Frequency

To ensure that the allele is sampled sufficiently for reliable use in statistical tests, a conservative minimum allele frequency (MAF) is used. The National Research Council report (1996) states that an estimate of an allele frequency may be inaccurate if the allele is so rare that it is represented only once or a few times in a database; some rare alleles might not be represented at all (NRC, 1996, p. 148). Thus, it is recommended that each allele should be observed at least five times for inclusion in reliable statistical calculations. The MAF for autosomal markers is thus 5/2N, where N is the number of individuals sampled from a population (making 2N the number of chromosomes counted). For example, when sampling 200 individuals from a population, the MAF would be 5/(2*200)=5/400, which equals 0.0125. If an allele is observed empirically at 0.009, the MAF of 0.0125 should be used for calculations instead. Collecting more samples usually only improves the precision of the result rather than the accuracy of the allele count (Butler(a), 2005, p. 477).

E.6.3 Profile Probability

Under the profile probability procedure, the probability that a given person having this allelic profile, chosen at random, has the same profile of the evidentiary DNA profile (E) is given. This approach has been used since the earliest days of DNA evidence for criminal cases (NIJ, 2000). Population databases are used for calculating the probability that a person's profile chosen at random from the relevant population has the same profile of the reference sample.

For a single-contributor sample and under the assumption that the profiles are from unrelated individuals, this probability is the frequency of occurrence of the profile(FBI, 2000), calculated from the allelic frequencies found in a population database. Generally, for a heterozygous individual (an individual having both alleles at a particular locus or gene), if the two alleles have frequencies of p and q in a population, the probability (P) of an individual of having both alleles at a single locus is:[16]

---

[15] NRC recommends for the U.S. population (general population) a conservative value of 0.01 for θ and a value of 0.03 for small isolated populations (See Recommendation 4.1 (NRC, 1996) and (Butler(b), 2005, p. 488).

[16] This is also expressed as $P_{ij} = 2p_ip_j$, $i \neq j$ where i and j are the alleles p and q.

$$P=2pq$$

Equation E.6.3-1: Heterozygous Probability

If an individual is homozygous (a person with a single allele at a particular locus or gene), for an allele with a frequency of p, the probability (P) of the genotype is.[17]

$$P=p^2$$

Equation E.6.3-2: Homozygous Probability

If the allele frequencies in the database are shown to be independently inherited, then the probability (P) for a DNA profile is the product of the probability (P1, P2, ... Pn) for each individual locus. Thus,

$$\text{Profile Probability} = \prod_{i=1}^{n} P_i$$

Equation E.6.3-3: Profile Probability

Note that these frequencies, and thus probabilities, are generated against the general population sampling, not against the frequencies of specific full profiles in the database.

E.6.4 Likelihood Ratio

Under the likelihood ratio (LR) procedure, the probability of the evidence is tested under two competing hypotheses. In a simple case, the two hypotheses are the prosecution hypothesis (H_p) and the defense hypothesis (H_d) about the source(s) of the profile(FBI, 2000).

$$LR = \frac{\Pr(E \mid H_p)}{\Pr(E \mid H_d)}$$

Equation E.6.4-4: Likelihood Ratio

For example, H_p is the hypothesis that the two profiles are from the same person (this is the profile probability from above), whereas H_d is the hypothesis that the two profiles are from unrelated persons. Using the above equation (Equation 1-4: Likelihood Ratio), this leads to the probability of the suspect profile given a crime sample profile and the two are from the same person over the

[17] Generally, for populations with substructure $P=p^2 + p(1-p)\theta_p$ (NRC, 1996), p 102, eq. 4.4a), one can replace θ with the kinship coefficient to measure the degree of relationship two persons. θ is not recommended to be used for heterozygotes (Recommendation 4.1in (NRC, 1996)).

probability of the suspect profile given a crime sample profile and the two are from people who are unrelated. If the LR is greater than one, then the evidence supports H_p, but if it is less than one, it supports H_d. For example, an LR of 100 means that it is 100 times more likely that the suspect is the perpetrator than someone else (e.g., a random person selected from the population) being the perpetrator (NIJ, 2000).

In simple cases when other evidence first implicates the suspect, the only match in the database, and ignores the possibility of relatives and population structure, LR is the same as the reciprocal of the profile probability(Gill, et al., 2006; FBI, 2000). For example, px is the probability of randomly drawing E from the appropriate population, which, in this case, is the same as the profile match from above and the likelihood of the sample coming from the same individual (Hp). The likelihood that the samples are from two different sources is px * px. The LR in this instance is now:

$$LR = \frac{\Pr(E \mid H_p)}{\Pr(E \mid H_d)} = \frac{p_x}{(p_x \times p_x)} = \frac{1}{p_x}$$

Equation E.6.4-5: Likelihood Ratio 2

There are several additional reasons for using LR: 1) it has useful statistical properties; 2) it can be used in complicated cases such as mixed samples; and 3) it can be converted into a probability by using Bayes' Theorem (NIJ, 2000).

E.6.5 Bayesian Inferences

While the match probability and likelihood ratio contain the same information, the likelihood ratio is especially useful given the availability of prior odds that the DNA samples come from the same source; in other words, the odds that the DNA came from the same source, using data other than DNA (NRC, 1996). In this case, Bayes' Theorem, shown below, can be used.

$$\text{Posterior Odds} = \text{Prior Odds x LR}$$

Equation E.6.5-6: Posterior Odds

The posterior odds can be used to directly answer the question in which the court is interested: What is the probability that the sample from the suspect and the evidence sample come from the same person? The odds are converted into a probability as follows.

$$probabilit\,y = \frac{odds}{odds - 1}$$

Equation E.6.5-7: Bayesian Inference

While used extensively for paternity testing and genetic counseling, Bayesian inferences are not often used in criminal cases. This is due to the fact that the courts are hesitant to ask juries to assign the prior odds based on the evidence, and that even a wide range of prior odds has no significant affect on the outcome. A four or five locus match yields an LR high enough such that a

realistic set of prior odds will generate a set of posterior odds arriving at the same conclusion (NRC, 1996).

E.6.6 Probability of Exclusion

Under the probability of exclusion (P_E) procedure, the probability is given that a randomly chosen person from a given population can be excluded as a contributor of a given mixture profile. In other words, this probability estimates the portion of the population that exhibits a genotype that contains at least one allele absent in the observed profile. For a population substructure that is statistically negligible, P_E is calculated as follows.

$$P_E = 1 - (P1 + P2 + \ldots Pn)^2$$

Equation E.6.6-8: Probability of Exclusion (PE)

While P_E is useful for complex mixtures, as it does not require any assumptions about the number of contributors to a given mixture, it does not use all available genetic information. If this information is available, other methods, such as LR methods, are preferred.

E.6.7 Counting Method

Under the counting method, the number of times a profile is in the database is counted. A frequency is generated from that count and confidence intervals are assigned (see (Butler(a), 2005, p. 271) Box 10.3 for an overview). This is usually applied to mtDNA or Y chromosome STRs (lineage markers), which are generally transferred directly from generation to generation unchanged(Butler(a), 2005, p. 515; Parson, Niederstätter, Lindinger, & Gill, 2007). This method is limited by the size of the database.

Appendix F DNA Databases

Selected additional DNA database initiatives are presented in this appendix to highlight the tremendous growth in commercial personal genetics, biopharmaceutical, and other areas not currently used for law enforcement and general forensics purposes. The technologies implemented in these rapidly expanding marketplaces should be monitored to determine technological opportunities to leverage capabilities within the federal government arena.

F.1 DNA Ancestry Project

The DNA Ancestry Project (http://www.dnaancestryproject.com) is a genetic genealogy project that uses Genebase as their Web 2.0 hosting site to link an individual's DNA profile to other subscriber profiles and their ancestors, similar in concept to LinkedIn (http://www.linkedin.com/). Genebase is owned and operated by Genetrack Biolabs Inc. (http://www.genetrack.bc.ca), a Canadian ISO17025 (International Organization for Standardization) and American Association of Blood Banks accredited laboratory. In joining Genebase a subscriber must agree to the following: "You are responsible for making the decision to upload your personal information to the Genebase website. Once you upload your information to Genebase's public database, it becomes accessible by all persons accessing Genebase for comparison. Genebase will not have any liability or obligation in connection with the publishing of any information by you."

Genebase contains both Y-DNA and mtDNA information. They use 61 markers, "making it the most comprehensive database available"[18]. This service uses 44 Y-STR loci, mtDNA[19], and seems to cover all the CODIS STR loci. Genebase does not sell a test for the CODIS loci, but one could order it from another company[20].

Genebase has 937,980 active members (27 January 2009). In addition, Genebase supports loading genealogies, currently having 3,532,432ancestors[21]. The number of DNA samples

[18] http://www.genebase.com/my/dnaMarkerDetail.php

[19] The order sheet only mentions hypervariable region I, but they sequence the both regions I and II.

[20] Such as the Bode Child ID Kit; http://www.dna-identity.com/kit.

[21] Between February 2008 and January 2009, Genebase added over 300,000 users and 1.6 million ancestors.

submitted is not explicitly stated on their site. However, by examining the frequencies for several Y-STR markers, it is estimated that Genebase has about 48,000 samples from the active members[22]. No frequencies are given for the mtDNA samples[23].

F.2 23andMe

23andMe, Inc., (www.23AndMe.com) is a privately-held personal genetics company based in Mountain View, CA and founded in 2006. This company offers DNA analysis to trace maternal and paternal lineage, and relatedness to ten world regions having more than 50 populations. This web-based service, launched November 2007, provides genotyping tools and features to report each customer's personal data within the context of environmental and other factors that contribute to variation in human traits and conditions (23AndMe, 2008).

Subscribers are sent a kit to collect saliva in a bar-coded tube, which is returned to the 23andMe's contracted laboratory for DNA extraction. The sample is exposed to 23andMe's customized version of the HumanHap550+ Genotyping BeadChip from Illumina that analyzes half a million locations in the autosomal DNA and more than 2000 SNPs across the entire mitochondrial genome.

The results of the analysis are available online via 23andMe's secure website. Users also have access to Ancestry.com for the analysis of their history. Investors include Genentech, Inc., Google Inc. and New Enterprise Associates (23AndMe, 2008).

F.3 deCODE Genetics

deCODE Genetics, Inc., (www.decode.com) is a biopharmaceutical company based in Reykjavík, Iceland. The company was founded in 1996 to research and isolate key human genes related to common diseases such as cancer, schizophrenia, and cardiovascular disease based on population studies. The results of this research are used to develop candidate drugs to fight these diseases.

In addition, deCODE offers contract services for human SNP and microsatellite genotyping, citing their peak capacity as 30 million microsatellite genotypes and 3 billion SNP genotypes per month. “deCODE now offers the first comprehensive solution for copy number variation

---

[22] Data current as of 27 January 2009 for DYS19 with allele 14 and 15 being the majority.

[23] mtDNA differences are relate to the CRS not the rCRS that EMPOP and the FBI use.

(CNV) analysis. The CNV genotyping platform covers over 13,000 novel regions and all known regions that can be expected to show common and rare CNVs, using more than 55,000 SNPs or probes. The CNV platform is available as a stand-alone product or in combination with Illumina's genome wide association genotyping array chips (370K and 1M SNPs)" (deCode, 2008).

Similar to 23andMe, deCODE also offers DNA analysis services to the general public to trace ancestry or determine ethnic background, to calculate the risk of contacting specific diseases based on the individual's genetic variants, or determine genetic traits. The analysis includes more than one million SNPs; this service is called deCODEme (https://www.decodeme.com).