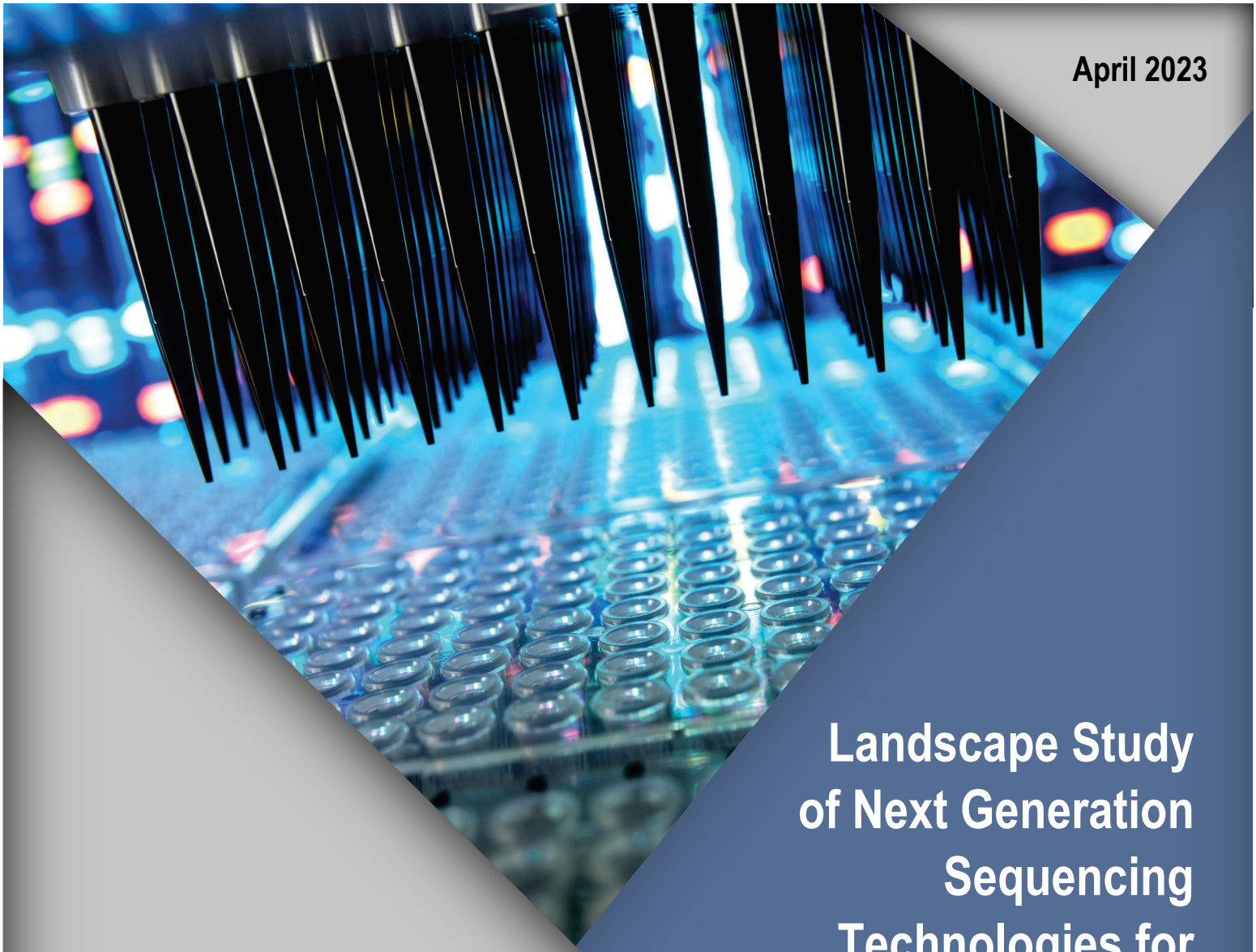

April 2023

Landscape Study of Next Generation Sequencing Technologies for Forensic Applications

FTCOE Contact

Jeri Ropero-Miller, PhD, F-ABFT
Principal Scientist
Project Director, FTCOE
JeriMiller@rti.org

NIJ Contact

Jennifer Love, PhD, D-ABFA
Physical Scientist
Office of Investigative and Forensic Sciences
Jennifer.love@usdoj.gov

Technical Contacts

Mikalaa Martin, BS
mmmartin@rti.org

Shannon Krauss, PhD
skrauss@rti.org

Acknowledgments

The Forensic Technology Center of Excellence (FTCOE) would like to thank the following individuals from Duquesne University's Master of Science in Forensic Science and Law program for providing critical support and collaboration in the development of this landscape study: Lyndsie Ferrara, PhD; Erin Estes, MS; and Ann DiFrank, BS.

Public Domain Notice

Suggested Citation

Martin, Mikalaa, Erin Estes, Shannon Krauss, Cody Sorrell, Rebecca Shute, Lyndsie Ferrara, Tracey Johnson, and Patricia Melton. "Landscape Study of Next Generation Sequencing Technologies for Forensic Applications." Research Triangle Park, NC: RTI International, April 2023.

Image Credits

Cover Photo: Andrew Brookes, "Scientific samples in microplate being prepared for research in laboratory," digital image, Getty Images, accessed March 23, 2023. https://www.gettyimages.com/detail/photo/scientific-samples-in-microplate-being-prepared-for-royalty-free-image/1455924862?phrase=dna%20laboratory&adppopup=true.

Table of Contents

Report Overview

The National Institute of Justice's (NIJ's) Forensic Technology Center of Excellence (FTCOE), led by RTI International, provides valuable resources that promote the use of novel and innovative technologies in the forensic community. Next Generation Sequencing (NGS) is a massively parallel sequencing technology that analyzes DNA samples based on size and sequence. This method enables Forensic Science Service Providers (FSSPs) to obtain additional information from an increased number of target regions by generating more comprehensive and robust data in less time than traditional sequencing methods. Beyond human identification, NGS technology has been researched as a technique to improve, for example, kinship testing and postmortem interval (PMI) estimation, which could benefit forensic casework. Although the use of NGS is increasing, FSSPs have been slow to adopt this technology for several reasons, including lack of funding, staffing, and laboratory space; the need for additional staff training; the time and effort required to restructure analytical reports; the potential need to align with accreditation requirements; and FSSPs' underdeveloped working knowledge of bioinformatics and the software used to analyze the data generated. This landscape study provides an overview of NGS technology, its application to various forensic science disciplines, and information on currently available NGS products specific to forensic science. Furthermore, this study summarizes considerations that impact product procurement and implementation within FSSPs. ***This landscape study focuses on the implementation of forensic science–specific, commercially available products for NGS and lessons learned from both publicly funded and private FSSPs that are early adopters of this technology.***

Landscape Objectives

This landscape study provides FSSP directors, laboratory personnel, decision-makers, and end-users with the following:

- Background information on the application of NGS technology to forensic science casework and research initiatives.
- A product landscape of commercially available NGS technologies, including instrumentation, library preparation kits, software, and automation solutions.
- User profiles from forensic practitioners and forensic researchers using NGS technologies in both FSSP and forensic science research settings.
- Benefits, limitations, and implementation considerations for currently available NGS products tailored to forensic science applications.

Landscape Methodology

To conduct this landscape study, the FTCOE used a two-step process.

1. Researched primary and secondary sources, including industry literature and journal articles, to obtain information related to NGS product capabilities, applications, future directions, and procurement and implementation considerations.
2. Discussed NGS technology, products, and applications with subject matter experts, including FSSP practitioners, decision-makers, researchers, educators, and vendor representatives.

Subject Matter Experts and Collaborators

We would like to thank the various forensic science community practitioners, decision-makers, researchers, and educators who offered insight and helped inform the development of this landscape study.

Adam Garver, MFS
CODIS Forensic Scientist and MPS Technical Leader
Ohio Bureau of Criminal Investigation

Lindsay Loughner Kotchey, MSFS
Forensic Scientist II (NGS Analyst)
The Armed Forces Medical Examiner System – Armed Forces DNA Identification Laboratory

Michael Marciano, PhD
Research Associate Professor and Director of Research
Syracuse University

Nicole Novroski, PhD
Assistant Professor and Forensic Genetics Researcher
University of Toronto Mississauga

Rachel Oefelein, MSc
Chief Scientific Officer
DNA Labs International

Jeanette Wallin, MPH
Criminalist Supervisor
California Department of Justice

Bill Hudlow, MS
Criminalist Supervisor and Technical Leader
California Department of Justice

Michelle Halsing, BS
Criminalist Supervisor and MPS Casework Analyst
California Department of Justice

Steven Myers, MS
Senior Criminalist
California Department of Justice

Daniela Cuenca, MS
Senior Criminalist
California Department of Justice

We would also like to thank the vendor representatives who supplied information about the commercially available NGS products discussed in this landscape study.

Danny Hall
Business Development Director
Verogen

Laura Ascroft
Field Application Scientist, Human Identification Division
Thermo Fisher Scientific

Glossary of Commonly Used Words and Phrases

For this document, the FTCOE has defined the following terms:

Adapter: A unique DNA sequence that caps the ends of the DNA fragments in each sample library to allow them to hybridize to a surface, providing a priming location for primers and barcoding for multiplexing.

Barcode: Unique index sequences added to each sample library during adapter ligation, which allows the user to distinguish between sample libraries.

Cluster: A sequencing reaction of each sample library that results in millions of copies of single-stranded DNA.

Capillary Electrophoresis (CE): An analytical technique that uses an applied voltage to separate DNA fragments within a sample.

Combined DNA Index System (CODIS): The Federal Bureau of Investigation's program of support for criminal justice DNA databases and associated software.[1]

Control Region (CR): The mitochondrial genome region consisting of non-coding DNA.

Coverage: The average number of reads that cover a particular target region; coverage indicates the relation between the number of reads that align to a reference (i.e., a reference standard of known sequence).

Flow cell or sequencing chip: The vessel onto which samples are loaded and where the sequencing reaction occurs.

Indel: An insertion or deletion of nucleotides in a genome.

Index: A barcode sequence added to a DNA fragment during library preparation that allows samples to be uniquely identified and demultiplexed after sequencing.

Instrument: Refers to a device or tool (i.e., sequencer, software).

Ion semiconductor sequencing: A technique that uses hydrogen ions to detect the sequence of the cluster during the sequencing reaction.

Isoalleles: Alleles that have the same length determination but that vary in sequence.

Library: A collection of similarly sized DNA fragments with known adapter sequences added to the 3' and 5' ends; a library corresponds to a single sample.

Ligation: The addition of indices and adapters during library preparation.

Next Generation Sequencing (NGS): A massively parallel sequencing technology that analyzes entire genomes or targeted regions of DNA or RNA based on size and sequence.

Polymerase Chain Reaction (PCR): An enzymatic process in which hundreds to billions of copies of a specific DNA sequence are produced depending on the number of cycles to which the sequence is subjected.

Product: Refers to the whole package, including the instrument, software, and library preparation kit.

Read Length: The number of base pairs sequenced from a DNA fragment; after sequencing, regions of overlap between reads are used to assemble and align the reads to a reference and reconstruct the DNA sequence.

Reads: The output that presents the genomic sequence of a sample subjected to NGS.

Sanger Sequencing: The use of electrophoresis to determine nucleotide sequences of DNA.

Sequencing by Synthesis (SBS): Step-by-step incorporation of modified nucleotides for sequencing.

Short Tandem Repeats (STR): Repeats of segments that are two to seven base pairs in length.

Single Nucleotide Polymorphism (SNP): Locations in the genome where organisms differ by one base pair.

Variant calling: Alignment of a sample with a reference to identify differences including SNPs and indels.

Key Takeaways

- NGS technology can provide comprehensive information from a DNA sample and improve analysis of highly degraded or compromised samples. The capabilities of NGS technology and products may save cost, labor, and time depending on a balance of multiple considerations, including the number of samples run on the sequencer at a given time, sample size, desired coverage, target regions, target type, and read length.[2]
- Forensic applications of NGS, as shown by emerging research, include human identification using nuclear and mitochondrial DNA; externally visible characteristics prediction and kinship testing using discriminatory single nucleotide polymorphisms and short tandem repeats; microbiome analysis using targeted microbiome markers; and age estimation using DNA methylation status.
- Despite the benefits of this technology, few FSSPs have implemented NGS. Adoption is currently limited by high-resource investments (e.g., funding, staffing, training, laboratory space). FSSPs must also consider validation requirements and whether they need to align with accreditation requirements before implementation.
- FSSPs are primarily implementing two instruments—the Ion GeneStudio™ S5 System by Thermo Fisher Scientific, Inc. and the MiSeq FGx© Sequencing System by Verogen, Inc. Both vendors sell library preparation kits, software, and automation solutions compatible with their respective workflows. Promega also offers library preparation kits compatible with the MiSeq FGx.
- Researchers and practitioners working with developers help drive improved FSSP understanding, use, and implementation of NGS.
- Implementing NGS requires investments in sequencing systems, library preparation kits, consumables, set-up, validation, training, and ongoing maintenance and may be more practical for FSSPs with higher case volumes and processing requests.
- Forensic NGS vendors provide resources for systems validation and training to ease the transition from an FSSP's original sequencing method to NGS.

Introduction

DNA has individually distinguishing characteristics used for human identification in criminal investigations. Traditional DNA analysis focuses on repeating units of DNA referred to as short tandem repeats (STRs), which are highly polymorphic regions of non-coding DNA. Alternative markers such as single nucleotide polymorphisms (SNPs) and mitochondrial DNA (mtDNA) can be applied to forensic applications such as identity, ancestry, lineage, and phenotypic trait prediction. SNPs are bi-allelic and therefore have a lower power of discrimination than STRs. As a result, more SNPs are needed to provide the same power of discrimination as STRs. mtDNA, which provides maternal lineage–based genetic information for both the nuclear family (i.e., mother and child) and extended family reference (e.g., grandmother, maternal cousins), can provide useful information in casework samples of highly degraded or compromised nuclear DNA and is often applied in cases of missing persons, unidentified human remains, and mass disaster victim identification efforts.

Currently, most Forensic Science Service Providers (FSSPs) perform forensic DNA analysis through the use of polymerase chain reaction coupled with capillary electrophoresis (PCR-CE),[3] which is limited to amplifying and labeling specific STRs for separation and detection according to size. With the advent of Next Generation Sequencing (NGS), a massively parallel sequencing process, FSSPs are now able to obtain information by analyzing STRs, SNPs, and other forensically relevant identity-informative markers. ***Exhibit 1*** offers a brief overview of PCR-CE and NGS workflows for STRs. Furthermore, NGS processes DNA samples based on size and sequence thereby increasing the discriminatory power for human identification. The ability for NGS to produce size- and sequence-based information from an increased number of DNA target regions provides FSSPs with more data, which traditionally required multiple techniques, technologies, and methods over multiple days to acquire. This has led to a need for FSSPs to become familiar with bioinformatics to understand the software they interface with when interpreting the data generated.[5]

The Forensic Laboratory Needs Technology Working Group (FLN-TWG) developed "Implementation Strategies: Next Generation Sequencing for DNA Analysis," as a primer to understand NGS technology and its practical implementation into FSSPs. This white paper provides an overview of key workflow differences between traditional CE-based methods and NGS.

Exhibit 1. Comparison of PCR-CE and NGS approaches for STRs.

Approach	PCR-CE	NGS
Workflow	Workflow includes linear process of sample collection, extraction,* quantification, PCR, CE, and data output	Mirrors the PCR-CE workflow but requires additional steps (i.e., library preparation) directly following PCR to prepare samples for sequencing
Data	Electropherogram denoting relative fluorescent intensity and size of each fragment	FASTQ file providing the number of reads and a run summary, including quality scores
Output	Allele table (size- and relative fluorescent intensity–based data)	Allele table (size-based data, read depth, and sequence-based data)

**Depending on the sample, DNA isolation, purification, and concentration may be required after extraction and before quantification.*

Although this landscape study primarily focuses on NGS applications for DNA, microbiome analysis (an RNA approach) is mentioned as a means for assisting in human identification efforts.

Across forensic disciplines, implementing NGS may provide several benefits:

- **Increased power of discrimination.** Because NGS generates both size- and sequence-based information, it can exploit isoalleles. The increased sensitivity of NGS products generates more data across the 20 Combined DNA Index System (CODIS) Core Loci, thereby increasing the statistical power for excluding an individual as a DNA donor or generating investigative leads.[6]
- **Improved performance with degraded or compromised DNA.** NGS allows primers to be placed closer to the target regions to create shorter target region fragments.
- **Ability to infer phenotype, ancestry, genealogy, and parentage.** NGS technology can analyze identity-, ancestry-, phenotype-, and lineage-informative SNPs and a constantly expanding panel of additional identity-informative markers within a single reaction.[7]
- **High-throughput capabilities.** NGS allows FSSPs to assess multiple samples or a large number of target regions in a single assay and to sequence multiple marker types in parallel, resulting in increased output efficiency.[8] The ability of NGS to achieve maximum efficiency depends on a balance of multiple considerations, including the number of samples run on the sequencer at a given time, sample size, desired coverage, target regions, target type, and read length.
- **Time savings.** NGS technology allows FSSPs to gather information (e.g., autosomal STRs, X-STRs, Y-STRs, mini-STRs, SNPs, identity-informative markers, and indels) through a single assay and in a single day, saving labor hours compared with running various sequential analyses.[9]

Despite the key benefits of NGS, FSSPs have been slow to adopt this technology.[4] Barriers to widespread implementation include the following:

- **Upfront and continual investment** requirements for resources, including funding and staffing.
- **The need for physical laboratory** space to accommodate instrumentation and digital infrastructure to support the increase in data generation.
- **Training**, whether in house or external, to support FSSPs who will be using NGS technology and to advance FSSPs' knowledge in bioinformatics.
- **The potential need to align with validation and accreditation requirements** for the instrument, library preparation kits, software, and optional automation solutions if replacing an FSSP's traditional STR sequencing method for entry into CODIS with NGS.
- **The time-intensive process** of integrating NGS workflows and reporting structures with standard operating procedures (SOPs) and laboratory information management systems (LIMS).
- **The possible need to procure and validate automated instruments,** such as liquid handlers to meet high case volumes and increasing processing requests.
- **Restrictions associated with the current CODIS upload framework** prevent all data generated by NGS technology from being populated within DNA databases (e.g., sequence-based data).
- **Early-stage implementation within forensic science** for criminal investigations may be accompanied by technology challenges such as fewer colleagues and peers available for troubleshooting, more time required for training, and minimal use in criminal proceedings.

Application of NGS to Forensic Science

For many forensic science disciplines, human identification is an essential component of the investigative process. NGS offers benefits that can overcome challenges related to forensic DNA samples, including insufficient sample quantity or quality because of degradation or inhibition. **Recent advances in forensic research have demonstrated that NGS is a viable approach to generate investigative leads (*Exhibit 2*).**[a]

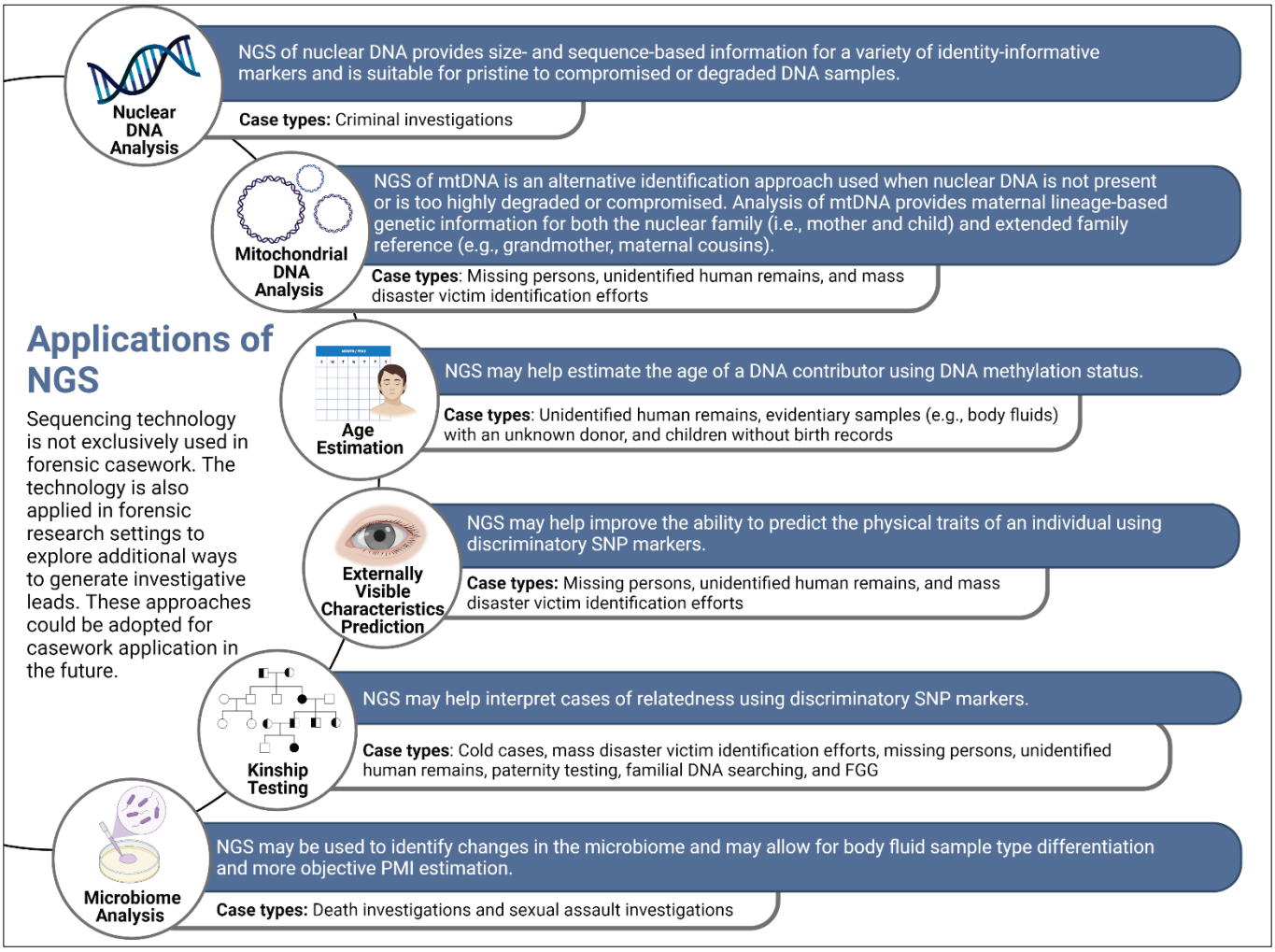

Exhibit 2. Various applications of NGS to forensic science.

Nuclear DNA Analysis

Nuclear DNA is unique to every individual, with the exception of identical twins, making it useful for human identification purposes and significant for forensic analysis. Recent advances in NGS technology offer FSSPs the ability to detect new levels of variation among STR alleles by providing both size- and sequence-based allele information.[10] NGS allows primers to be placed closer to the target regions to produce shorter fragments, which

[a] NGS may have other relevant applications to forensic science as outlined in subsequent sections. However, it is important to note that some of these methods (e.g., Externally Visible Characteristic [EVC] prediction) would typically be applied after FSSPs have exhausted all other investigative and forensic testing methods and approaches. Although NGS technology provides previously unexploited identity information, the FSSP must be cautious in what type of information is generated and reported as there are ethical implications of using DNA to review coding regions of a DNA donor.

can be advantageous for challenging (e.g., degraded, inhibited, compromised) DNA samples. Furthermore, NGS allows FSSPs to sequence autosomal STRs, X-STRs, Y-STRs, along with identity-, ancestry-, phenotype-, and lineage-informative SNPs using a single assay.

mtDNA Analysis

mtDNA can be used as an alternative forensic identification approach when nuclear DNA is nonexistent, highly degraded, inhibited, or compromised, which is common with challenging sample types such as hair roots, teeth, and bones. Traditional mtDNA analysis focuses on the control region (CR), or the non-coding region of the DNA. Many individuals remain indistinguishable across the CR; therefore, FSSPs assess the entire mitochondrial genome (mtGenome). The use of traditional Sanger sequencing for this assessment can be time-consuming, labor-intensive, and costly.[11] Using NGS to sequence the mtGenome can overcome the limitations of traditional Sanger sequencing because this method can sequence large amounts of the mtGenome in one sequencing run, producing 20–60 times more CR data than traditional mtDNA analysis approaches.[12] Additionally, NGS can be effective on samples containing as low as 100 pg, as is common with degraded samples in which mtDNA analysis is employed.[13]

Age Estimation

Aging is associated with many factors, including various molecular changes to cells that occur throughout an individual's lifetime (e.g., gene expression, chemical changes, structural changes). Historically, age is often estimated using bone morphology, which can have limited accuracy because of individual variation in the skeletal aging process, but NGS offers promising ways to estimate the age of individuals by using specific epigenetic markers that measure DNA methylation status (i.e., a modification that can influence gene expression and developmental processes), which has been identified as a biomarker for age.[14] Specific DNA methylation patterns have proven useful in modeling chronological age from multiple sample types, including blood and saliva.[15] Several age estimation models based on DNA methylation status have been developed with the use of Illumina's MiSeq System, which can quantify methylation status in targeted sites of blood samples.[16]

Externally Visible Characteristic Prediction

Forensic DNA phenotyping allows FSSPs to predict a DNA donor's appearance. EVCs include physical traits such as eye color, skin tone, and hair color, which are determined by environmental factors and the individual's genotype. Assessing these traits could offer guidance to investigators in cases where there are no persons of interest and where no associations have been made in a DNA database to develop investigative leads. Various NGS phenotyping panel assays using phenotype-, ancestry-, and lineage-informative SNPs have been developed to assess hair color, eye color, and skin tone, with studies suggesting these assays have high sensitivity and accuracy for EVC prediction of a DNA donor.[17, 18, 7, 19]

Kinship Testing

Kinship testing helps determine the relationship between two potentially related individuals. One application of kinship testing in forensic casework is Forensic Genetic Genealogy (FGG), which is exclusively an NGS technique. FGG combines genetic and traditional genealogical methods to generate investigative leads for law enforcement entities investigating violent crimes and identifying human remains.[20] FGG uses genetic information uploaded by users to direct-to-consumer companies and other public genealogy databases that allow law enforcement to search forensic uploads to these databases (e.g., FamilyTreeDNA and GEDmatch PRO) to identify unknown perpetrators or unidentified victims. To help FSSPs integrate this technique, there are commercially available forensic kinship library preparation kits with paired integrated workflows (e.g., Verogen's ForenSeq Kintelligence) that target forensically

relevant SNPs in a single assay.[21] The ForenSeq Kintelligence workflow can produce GEDmatch PRO- and FamilyTreeDNA-compatible data, simplifying the FGG approach to generate investigative leads.

Microbiome Analysis

The microbiome refers to the microorganisms and their combined genetic material located in a particular environment (e.g., on and within the body or parts of the body, tissues, body fluids), which varies between individuals.[22] The microbial structures of major body fluid types can be used to classify and identify body fluids (e.g., urine, saliva, blood, and feces).[23] Additionally, the microbiome may be used for human identification by examining the bacterial communities that are prevalent on and within all individuals.[24] Generally, microbial human identification relies on characterizing the present taxonomic groups and their relative abundance, as identified through RNA sequencing; however, the microbial communities of one individual can vary across site and time, which may pose limitations for investigative lead generation.[25] Targeted RNA sequencing, an NGS-based gene expression profiling method, is increasingly used to analyze microbial communities for the purposes of PMI estimation[26] and to identify tissues[27] and body fluids.[28-30, 24, 31, 32] When used for applications outside of human identification (e.g., PMI estimation), targeted RNA sequencing improves upon traditional analysis approaches because it allows for comprehensive testing and results through a single assay. In contrast, traditional analysis approaches require multiple tests or forensic approaches that can be time and resource intensive.

NGS Workflow and Methodologies

The NGS workflow is a series of steps and procedures that converts DNA samples into functional information that FSSPs can use for forensic science applications. Forensic DNA sequencing uses targeted sequencing, and the workflow can be separated into four phases (***Exhibit 3***).[b] ***The sections that follow provide information on the NGS workflow used specifically for targeted DNA sequencing.***

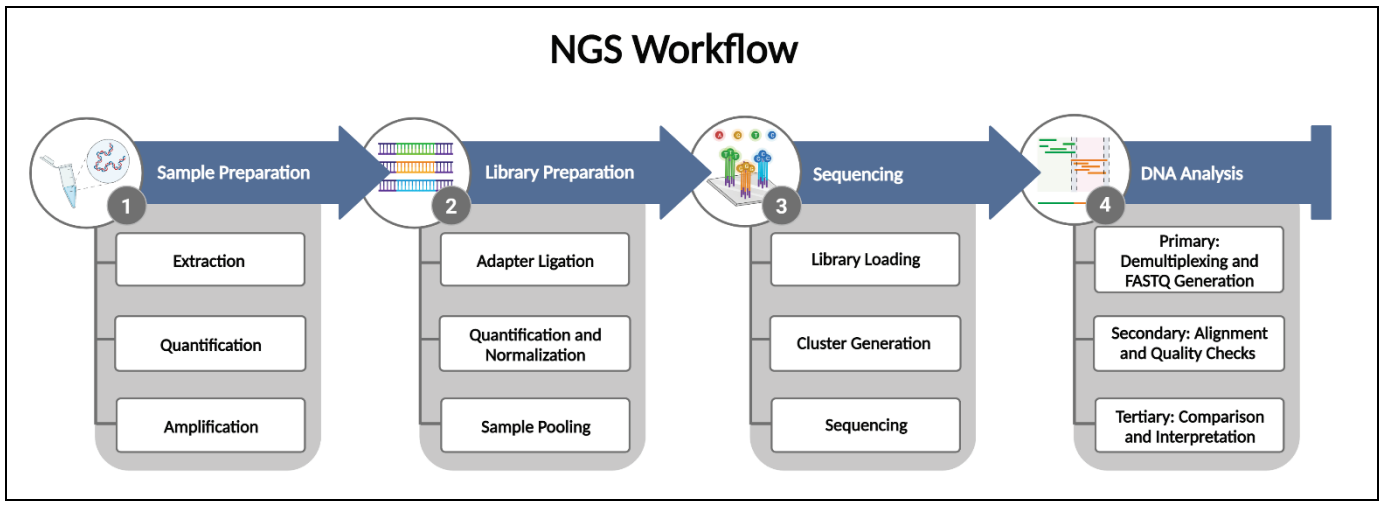

Exhibit 3. The four main phases and associated steps that comprise an NGS workflow.

1. Sample Preparation

The first phase of the NGS workflow is to prepare samples for downstream analysis. DNA is first extracted and isolated from suspended cellular material and then purified to separate DNA from non-DNA cellular debris. The amount of extracted DNA is quantified to determine if there is adequate DNA to proceed to the library preparation phase. An optimal input DNA quantity is typically 1 ng but may vary based on the library preparation kit and the quality of the starting material. If necessary (e.g., if a DNA extract is not of optimal input quantity), an amplification step is performed to tag and amplify target regions (e.g., STRs, SNPs, identity-informative markers) before library preparation.

FTCOE: Understanding NGS Workflows Through a Simulation Tool

The FTCOE's MPS Workflow Through Simulation Tool was designed to teach users the best practices for implementing NGS technologies in FSSPs. It allows users to interact with and navigate through a virtual laboratory and provides a mock SOP to augment the user's experience. Additionally, this simulation illustrates the capabilities of various NGS instruments.

2. Library Preparation

The second phase of the NGS workflow is the library preparation phase. During this phase, the suitability of the DNA sample is evaluated to ensure it is compatible with the sequencing technique and instrumentation. Each library consists of a collection of similarly sized DNA fragments from a single source with known adapter sequences ligated at each end. The choice of appropriate library preparation kit depends on (1) the information to be derived from

[b] Whole genome sequencing, such as is used for sequencing of RNA for microbiome analysis, undergoes a similar process, with the primary difference being the choice of library preparation kit used for creating the library. Whole genome sequencing may require additional sequencing support and varies in workflow and is not within the scope of this landscape study.

the sample, (2) the quantity and quality of the sample, and (3) the amount of space available for sequencing on board the instrument.

Library preparation begins with target sequences pieced or sized to a desired length, although this is often performed during the amplification step within the sample preparation phase for a targeted sequencing workflow. Once the fragments are sized, libraries are prepared for sequencing through target enrichment. The two most common methods of target enrichment are PCR-based enrichment, which uses amplicon sequencing, and hybrid capture–based enrichment (with subsequent PCR amplification), which uses single-stranded oligonucleotides (i.e., probes). Hybrid capture–based enrichment can be advantageous for challenging samples. Specialized adapters are added to each end of the target fragments to bind the fragments to the flow cell (when using Verogen instruments) or sequencing chip (when using Thermo Fisher Scientific instruments). Samples are quantified and then undergo a process called normalization to equalize the concentration of each DNA sample being run. This process is important, especially when running samples with different anticipated DNA yields at a single time. Different quantities of input DNA can lead to inconsistent data quality and output per sample. Once samples are normalized, multiple libraries can be pooled and sequenced together within a single run using a process called multiplexing. A unique index system, or barcode, is added to each library during adapter ligation that can be used during data analysis to demultiplex and distinguish between libraries. Lastly, libraries are diluted and denatured for sequencing.

Quality control of prepared libraries is an important part of the library preparation process. To enable downstream quality control checks, positive and negative controls are prepared during library preparation in addition to the sample, which can help when troubleshooting or if vendor assistance is needed after the sequencing run. Each sample library is checked for quantity and average fragment length before pooling and sequencing. These steps ensure equal and adequate sampling of all libraries to allow for the production of balanced sequencing data. A sequencing reaction may yield uninformative or neutral results if there is low-quantity or low-quality DNA.[c]

Library preparation uses kits specific to sample type, assay, technology, target regions, and desired coverage. Many developers offer both predesigned panels and custom panels. Predesigned panels target variants associated with the target regions. Custom panels are made to order for the user's target regions. Currently, Verogen, Thermo Fisher Scientific, and Promega offer kits for mtDNA and nuclear DNA sequencing.

Additionally, these vendors may offer library preparation and sample loading automation solutions. Thermo Fisher Scientific currently offers the Ion Chef System for use with the Ion GeneStudio S5, and Verogen offers the PrepStation for use with the MiSeq FGx. Instruments may also be compatible with third-party robotics (e.g., automated liquid handlers) to enable a more streamlined workflow.

3. Sequencing

The third phase of the NGS workflow is sequencing. To begin a sequencing run, sample libraries are loaded onto a flow cell or sequencing chip within the sequencing instrument. The sample is then hybridized to the flow cell or sequencing chip to prepare the sample for on-board amplification. On the instrument interface, users select the library preparation kit, assign indexes for each sample, specify parameters required for the sequencing run, load associated reagents into the instrument, and begin the sequencing run.

Although each NGS instrument is based on the same general workflow, the underlying sequencing reactions vary. Verogen's MiSeq FGx uses sequencing by synthesis (SBS) technology, where cluster generation and amplification occur on the flow cell in a bridge format. As the sequencing strand grows, fluorescently labeled nucleotides become incorporated into the sequencing strand and are interpreted by the instrument as different base pairs during

[c] A high-quality library is one that has a diverse set of DNA fragments and minimal duplicate fragments. Duplicate fragments can cause bias during sequencing because repetition can lead to a sequence being over-represented in the resulting data.

detection. Alternatively, Thermo Fisher Scientific's Ion GeneStudio S5 uses ion semiconductor sequencing technology, which detect changes in pH that are converted to an electrical signal while nucleotides are incorporated into the sequencing strand. Each sequencing method has strengths and weaknesses (e.g., the cycle-based sequencing of the MiSeq FGx is more susceptible to base pair substitutions, whereas the Ion GeneStudio S5 generally has an overall higher error rate and is more prone to making insertions and deletions).[34] The MiSeq FGx has a longer run time than the Ion GeneStudio S5 because it lacks optical scanning and cycle-based sequencing.[33]

Product-specific information about the library preparation kits, sequencing instruments, and optional automated library preparation systems for forensic applications can be found in the vendor profiles.

4. Data Analysis

NGS generates large volumes of data that can lead to complex and time-consuming analyses. NGS files may reach up to hundreds of gigabytes (GB) in size depending on the coverage, sample size, target regions, and selected target type and read length. Vendor analysis software perform various data processing steps, including data quality checks. Additionally, these software can help the user understand and interpret the results and remove the need to develop complex bioinformatics pipelines.

After completing an NGS run, the software will perform three main data processing steps, which include demultiplexing pooled samples and creating a FASTQ file containing the raw data and quality score; performing quality checks, alignment, and variant calling; and generating a genome variant call format (gVCF) file that is used to compare and interpret results.

Commercially Available Data Analysis Software

Vendors such as Verogen and Thermo Fisher Scientific offer data analysis software specifically designed for compatibility with their instrumentation. Some software, including Verogen's Universal Analysis Software (UAS) and Thermo Fisher Scientific's Converge Software, are designed to offer support across an entire workflow. Various third-party vendors offer additional software analysis options that support specific phases of the data analysis process, several of which are open source. SoftGenetics LLC, for example, offers a commonly used software called GeneMarkerHTS that possesses different tools for analyzing forensic markers, including mtDNA, STRs, and Y-STRs.

The primary data analysis step includes demultiplexing the pooled samples using the indices added during library preparation to allow the sequencing reads to be separated and analyzed individually. Once samples are demultiplexed, the software automatically calls base pairs and computes sequencing quality scores into a FASTQ file output. The FASTQ file contains information such as raw reads, index sequences, and a sequencing quality score for each base.[34]

The secondary data analysis step involves aligning reads within a Binary Alignment Map file. This is accomplished by assigning each read to a specific STR or SNP locus followed by assigning an allele value to each locus. Following read alignment, variant calling is completed using the assigned allele values, which allows the user to identify how the sample differs from the reference. The output of variant calling is a gVCF file that is used for extracting information, annotation, interpretation, and comparison.

During the mapping (i.e., realigning) of sequencing reads to the reference, sequence artifacts may be observed. NGS data may exhibit stutter and noise (i.e., incorporation errors). Within NGS data, noise results from the misincorporation of one or two bases, which leads to alleles of the same size presenting a different sequence.

Because NGS generates sequence-based information, stutter can more easily be filtered out because it shares the same sequence as the parent allele.

The tertiary data analysis step compares the sequencing reads to the reference to calculate match statistics (e.g., likelihood ratios) for reporting. Currently, efforts are being made to make population statistics for NGS data, specifically for sequence-based alleles, publicly available.[35-37] The final data format for NGS is an allele chart that presents coverage versus allele value. Additionally, sequence-based allele information can be displayed and exported as needed. Data presentation will differ based on the software used.

Vendor Profiles

The vendor profiles that follow provide a product landscape of commercially available NGS technologies, including instrumentation, library preparation kits, software, and automation solutions.

VEROGEN Verogen offers sequencing instruments and library preparation kits that analyze forensic data.

Verogen was established in 2017 as the first company dedicated to sequencing technology for forensic applications. Verogen currently offers instrumentation, library preparation kits, software, and automation solutions for the identification and analysis of forensic material, including STRs, SNPs, and mtDNA.

Instrumentation

Verogen's NGS sequencing instrument, the MiSeq FGx, was originally developed by Illumina, Inc., and uses their SBS technology featuring Short Read Synthesis for STR, mtDNA, and whole genome applications found in their original MiSeq System (***Exhibit 4***). The MiSeq FGx differs from the original MiSeq System (Illumina, Inc.) because it incorporates the Universal Analysis Software (UAS), which is used to visualize results specifically useful for forensic applications. The MiSeq FGx is National DNA Index System (NDIS)-approved and was the first instrument used to gather NGS evidence for a criminal conviction.[38] The MiSeq FGx offers a Forensic Genomics mode for casework and a Research Use Only mode that allows users to alter configurations and parameters to explore the full capabilities of this system.

Model	MiSeq FGx© Sequencing System
Technical Method	Sequencing by synthesis (SBS)
Throughput	5 GB
Reads Per Run	12.5 million
Run Time	30 hr
Dimensions (W x D x H)	68.6 cm × 56.5 cm × 52.3 cm
Weight	54.5 kg
Ideal Operating Environment	19°C–25°C; 30%–75% Humidity
Power Requirements	100–240 V AC @ 50/60 Hz, 10 A, 400 W

Exhibit 4. Image of the MiSeq FGx Sequencing System offered by Verogen, Inc. (left) and technical specifications (right). Photographs and content used courtesy of Verogen; copying prohibited.

Library Preparation Kits

Verogen supports multiple library preparation kits designed for compatibility with the MiSeq FGx (***Exhibit 5***). The ForenSeq mtDNA Whole Genome Kit and the ForenSeq mtDNA Control Region Kit are both capable of amplifying and sequencing mtDNA with a recommended 100 pg and 50 pg of starting material, respectively. Amplification of nuclear DNA can be performed using the following kits:

- *ForenSeq DNA Signature Prep Kit*: Capable of generating DNA profiles using up to 230 genetic markers, this kit offers two different primer mixes. DNA Primer Mix A (DPMA) contains primer pairs for autosomal, X-, and Y-chromosome STR targets while DNA Primer Mix B (DPMB) includes all the primer pairs in DPMA in addition to primer pairs for biogeographical ancestry-informative SNPs. This kit is NDIS-approved.

- *ForenSeq Kintelligence Kit*: This kit targets 10,230 SNP markers, including markers in degraded samples with an average amplicon size of less than 150 bp, from genomic DNA that are aimed specifically for forensic kinship applications. This kit allows for single SNP multiplexing that works across global populations and is the only sequencing-based assay available for generating GEDmatch PRO– and FamilyTreeDNA-compatible data for FGG uploads and searching through the associated UAS preconfigured ForenSeq Kintelligence Analysis Module.[21]
- *ForenSeq MainstAY Kit*: This kit targets 27 autosomal STRs and 25 Y-STRs markers for up to 96 samples per run. This kit is best used for general casework.
- *ForenSeq Imagen Kit*: This kit targets 111 SNP markers for up to 96 samples per run and offers two different primer mixes. DNA Primer Mix E (DPME) contains SNP and Y-SNP primer pairs for skin, hair, and eye color while DNA Primer Mix F (DPMF) includes all the primer pairs in DPME in addition to primer pairs for SNP targets used for the prediction of biogeographical ancestry. This kit generates preconfigured reports through the associated UAS module that can be uploaded to third-party analytical tools (e.g., HIrisPlex-S DNA Phenotyping Webtool that allows for additional EVC predictions such as skin, hair, and eye color).[39]

Verogen offers an optional PrepStation liquid handler validated for semi-automated preparation of the ForenSeq kits, and allows for a simplified workflow between the MiSeq FGx and UAS. At the time of this landscape study, only the ForenSeq MainstAY Kit has a PrepStation script available, with all other ForenSeq kit scripts still in the development stage.[40]

Exhibit 5. Verogen library preparation kits available for NGS analysis.

Technical Specifications	mtDNA		Nuclear DNA			
	ForenSeq mtDNA Whole Genome Kit©	ForenSeq mtDNA Control Region Kit©	ForenSeq DNA Signature Prep Kit© Primer Mix A/B	ForenSeq Kintelligence Kit©	ForenSeq MainstAY Kit©	ForenSeq Imagen Kit© Primer Mix E/F
DNA Source	Mitochondrial	Mitochondrial	Nuclear	Nuclear	Nuclear	Nuclear
Flow Cell Type	MiSeq FGx Reagent Kit	MiSeq FGx Reagent Micro Kit	MiSeq FGx Reagent Kit and Reagent Micro Kit	MiSeq FGx Reagent Kit	MiSeq FGx Reagent Micro Kit	MiSeq FGx Reagent Micro Kit
Libraries per Flow Cell	16	3–48	8–96	3	8–96	8–96
Amplicon Detection	131 bp (average)	118 bp (average)	≥ 65bp	< 150 bp	235 bp (average)	112 (DPME) 111 (DPMF)
NDIS Approval	No	No	Yes	Not Required	Submitted for NDIS review	Not Required
Recommended Input	100 pg	100 pg	1 ng	1 ng	1 ng	1 ng
Amplicons	245	18	—	—	—	55 (DPME) 111 (DPMF)
Target Size	16,569 bp	1,200 bp	—	—	—	—
Y-STRs	—	—	24	—	25	—
X-STRs	—	—	7	—	—	—

Technical Specifications	mtDNA		Nuclear DNA			
	ForenSeq mtDNA Whole Genome Kit©	ForenSeq mtDNA Control Region Kit©	ForenSeq DNA Signature Prep Kit© Primer Mix A/B	ForenSeq Kintelligence Kit©	ForenSeq MainstAY Kit©	ForenSeq Imagen Kit© Primer Mix E/F
Identity SNPs	—	—	94	94	—	—
Phenotypic SNPs	—	—	— (DPMA) 22 (DPMB)	22	—	41
Biogeographical SNPs	—	—	— (DPMA) 56 (DPMB)	56	—	56
Kinship SNPs	—	—	—	9,867	—	—
X-SNPs	—	—	—	106	—	—
Y-SNPs	—	—	—	85	—	14
Total Markers	—	—	152 (DPMA) 230 (DPMB)	10,230	52	107
Global Autosomal STRs	—	—	27	—	27	—

All kits are compatible with the MiSeq FGx Sequencing System. Dashed information indicates information not available or not applicable at the time of this landscape study. Note that the flow cell type used alters the number of libraries per flow cell.

Analysis

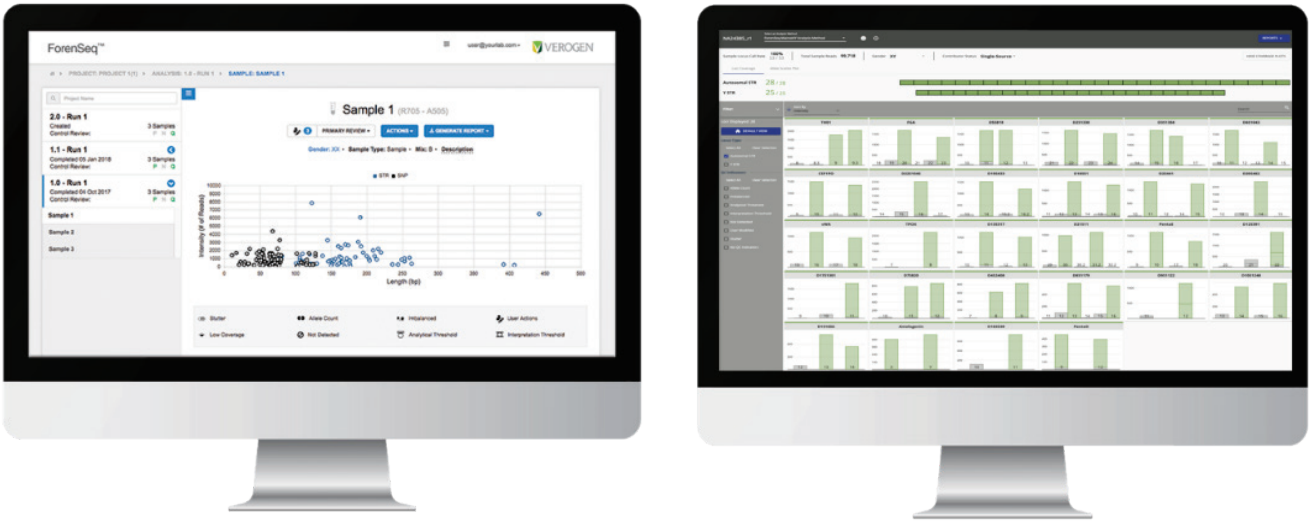

Exhibit 6. A representation of the user interface for the preconfigured UAS modules for the ForenSeq DNA Signature Prep Kit (left) and the ForenSeq MainstAY Kit (right).

The UAS is available for visualizing and interpreting NGS data generated through the use of all current ForenSeq kits and workflows.[41] Sequencing quality and feedback can be acquired post-run and generate an appropriate FASTQ file when operated in either the Forensic Genomic or Research Use Only modes. For the Forensic Genomic mode, each library preparation kit is analyzed with a specific plug-in within the UAS that best visualizes the type of

amplified genetic data. The UAS is locally hosted, so data does not need to be stored in a cloud environment. UAS modules can create multiple report types for visualizing sample analysis related to kinship, ancestry, and EVC prediction. In Research Use Only mode, generated FASTQ files can be exported and analyzed using several third-party software or custom scripts. Verogen continuously releases version control updates to expand the software to accommodate new applications and update nomenclature conventions. ***Exhibit 6*** below shows a representative image of the analysis pipeline for two different UAS modules.

Implementation

Verogen provides installation and on-site training with the purchase of the MiSeq FGx products. Training covers the use of library preparation kits, how to execute runs with the MiSeq FGx, and data analysis via the UAS. The training is designed to teach essential techniques for success and provide experience to new users. The 3-day structured training is included with every instrument purchase, and additional customizable training is available to meet each FSSP's specific needs. For long-term support, multiple internal validation plans specifically suited for FSSP casework are available upon request through Verogen. There are three levels of validation services available:

- **Planned**: FSSP personnel perform validation studies with reagents provided by Verogen.
- **Assisted**: FSSP personnel plan and perform validation studies with the help of Verogen technical experts for analysis.
- **Full**: Verogen technical experts plan and perform the validation studies independent of FSSP personnel.

Thermo Fisher Scientific enables NGS sequencing using their standard Ion GeneStudio S5 System and specialized nuclear and mtDNA library preparation kits.

Thermo Fisher Scientific's Applied Biosystems brand focuses on integrated systems for genetic analysis, including the Ion GeneStudio S5 System. This instrument is optimized for FSSPs, and automated workflows may be set up using the Ion Chef System.

Instrumentation

The Ion GeneStudio S5 is an ion semiconductor-based instrument using long-read sequencing technology focused on enabling simplified and targeted workflows (***Exhibit 7***). The Ion GeneStudio S5 offers cartridge-based reagents, meaning all reagents are provided in a single cartridge for simplicity and ease. This instrument can sequence and analyze SNPs, mtDNA, and STRs useful for forensic analysis and human identification applications.

Model	Ion GeneStudio™ S5 System
Technical Method	Ion semiconductor sequencing
Throughput	15 GB
Reads Per Run	80 million
Run Time	19 hr
Dimensions (W x D x H)	54.2 cm × 80.6 cm × 50.9 cm
Weight	63.5 kg
Ideal Operating Environment	20°C–30°C; 40%–60% Humidity
Power Requirements	100–240 V AC @ 50/60 Hz, 14A

Exhibit 7. Image of the Ion GeneStudio S5 System and Precision ID mtDNA library preparation kit offered by Thermo Fisher Scientific (left) and technical specifications (right). Photographs and content used courtesy of Thermo Fisher Scientific; copying prohibited.

Library Preparation Kits

Thermo Fisher Scientific offers a variety of panels specifically designed for forensic applications (***Exhibit 8***). All library preparation kits are compatible with the Ion GeneStudio S5 and the optional Ion Chef System for automated library preparation. The Precision ID mtDNA Whole Genome Panel and Precision ID mtDNA Control Region Panel are both capable of amplifying and sequencing mtDNA and are ideal for highly compromised or degraded samples. The CR of the Precision ID mtDNA Whole Genome Panel is NDIS-approved and provides robust performance across diverse population samples. Amplification of nuclear DNA can be performed using the following kits:

- *Precision ID Identity Panel*: This panel includes 124 autosomal markers useful for discriminating similar DNA donor profiles and uses SNPs for human identification of degraded and trace DNA.
- *Precision ID Ancestry Panel*: This panel includes 165 autosomal markers for biogeographical ancestry. The panel was developed to amplify small amplicon sizes (e.g., less than 130 bp), making it optimal for degraded DNA samples.
- *Precision ID GlobalFiler NGS STR Panel v2*: This panel includes 35 markers for preparing libraries of forensically relevant STRs, including the 20 CODIS Core Loci, nine multiallelic STR markers, four markers for sex determination, and two pentanucleotide STRs.
- *Community panels for human identification*: Thermo Fisher Scientific maintains a catalog of community panels that researchers have built for custom applications and specific target regions. These are not verified by Thermo Fisher Scientific, but the relevant publication information is linked on their website.

Exhibit 8. Thermo Fisher Scientific library preparation kits available for NGS analysis.

Technical Specifications	mtDNA		Nuclear DNA		
	Precision ID mtDNA Whole Genome Panel™	Precision ID mtDNA Control Region Panel™	Precision ID Identity Panel™	Precision ID Ancestry Panel™	Precision ID GlobalFiler™ NGS STR Panel v2
DNA Source	Mitochondrial	Mitochondrial	Nuclear	Nuclear	Nuclear
Panel Primer Pools	2	2	1	1	1
Samples per 530 Chip (Ion Chef Prep/Manual Prep)	32	32/56 (520 Chip recommended)	32/384	32/362	32
Average Amplicon Size	163 bp	153 bp	138 bp	127 bp	129–303 bp
NDIS Approval	Yes (CR Only)	No	No	No	No
Recommended Input	0.3 ng (gDNA)	0.3 ng (gDNA)	1 ng	1 ng	1 ng
Total Targets	162	14	—	—	—
Target Size	16,569 bp	1,200 bp	—	—	—
Y-STRs	—	—	—	—	4
Identity SNPs	—	—	90	165	—
Y-SNPs	—	—	34	—	—
Total Markers	—	—	124	165	35
Global Autosomal STRs	—	—	—	—	31
X-STRs	—	—	—	—	1

All kits are compatible with the Ion GeneStudio S5 and Ion Chef Systems. Dashed information indicates information not available or not applicable at the time of this landscape study.

Analysis

Thermo Fisher Scientific offers the Converge™ Software, which specializes in the analysis of genomic data from Thermo Fisher Precision ID Panels. NGS analysis includes information on STR allele calling, STR sequence motifs, SNPs in flanking regions, and the ability to distinguish isometric heterozygotes. The software contains modules for kinship and paternity testing and information can be integrated into an existing LIMS.

Implementation

Thermo Fisher Scientific offers a variety of specialized training courses and educational services taught by their application scientists. Training may be included as part of an FSSP's validation package, depending on their needs. The training aims to provide the FSSP with knowledge on best practices in experimental design, workflow, and instrument troubleshooting. Validation services are also offered through their Human Identification Professional Services team. Human Identification Professional Services is designed to help FSSPs integrate, validate, and implement their new instrumentation. HPS offers the following services:

- Validation/performance checks and accreditation support
- SOP development and project management
- Consulting and operational efficiency programs
- Advanced human identification training
- Laboratory relocation and renovation
- Robotic script verification and validation
- LIMS integration and validation
- Turnkey solutions and bespoke package

Promega currently sells library preparation kits for sequencing mtDNA and nuclear DNA sources that are compatible with the MiSeq FGx.

Promega manufactures enzymes and reagents for biotechnology and molecular biology assays.

Library Preparation Kits

Promega has developed two library preparation kits for forensic science sequencing applications that are both compatible exclusively with Verogen's MiSeq and MiSeq FGx. Sequencing data generated from either kit below can be analyzed using the GeneMarker® HTS software (SoftGenetics, LLC).

- *PowerSeq® 46GY System*: This system contains reagents to amplify autosomal and Y-STR loci using Illumina® TruSeq® chemistry to prepare libraries and generate sequencing data. The PowerSeq 46GY System is designed for nuclear DNA analysis of 22 autosomal loci and 23 Y-STR loci used for human identification applications, including forensic analysis and kinship testing.
- *PowerSeq® CRM Nested System*: This custom, NDIS-approved system generates 10 small amplicons of the mtGenome for CR analysis. The protocol consists of a single PCR step to amplify the target amplicons and incorporate indexed sequencing adapters. This nested amplification protocol greatly reduces the number of steps and time required to produce libraries ready for sequencing.

Implementation and Product Procurement Considerations

FSSP decision-makers should consider their existing laboratory policies, procedures, workflows, and resources before purchasing an NGS product. During this process, decision-makers should consider the following.

Identifying whether NGS is practical for your laboratory.

FSSPs can choose to use this technology for casework by either purchasing an NGS system or using a contracted vendor. The number of processing requests a FSSP receives may be the primary determinant—an FSSP with high case volumes and processing requests will most likely recoup the investment of implementing NGS. In the user profiles below, some FSSPs mentioned the added benefit of automation technologies, such as robots and liquid handlers, to further decrease turnaround times.

Planning for costs beyond the NGS instrument.

A sequencing instrument, which may cost more than $140,000, is a significant investment for an FSSP. Additionally, implementing NGS may also require consumables (e.g., 96-well plates, pipettes, centrifuge tubes, microcentrifuge tubes, PCR plates, ethanol) and available laboratory equipment (e.g., thermal cycler, microcentrifuge, microplate centrifuge, vortexer, magnetic stand) specific to NGS. This upfront investment will also include validation, technical labor and training costs, preparing and expanding digital infrastructure for data storage, preparing and expanding facilities for laboratory bench space, instrument support, and preventative maintenance. FSSPs may consider federal, state, local, or private funding to help defray some of these costs.[42-46]

The FLN-TWG report, "**Implementation Strategies: Next Generation Sequencing for DNA Analysis**,"* provides comprehensive cost breakdowns (totals and cost per sample) for the products mentioned in this landscape study.

**The pricing for products in this white paper was current as of September 2021. Please contact vendors directly for updated pricing.*

Understanding the potentially long road to implementation and available resources.

Interviews with early NGS adopters indicated that successful NGS implementation is a multi-step process and a significant time, resource, and monetary investment. FSSPs need to be aware of, and prepared for, a long road to adoption, procurement, and implementation. Beyond initial procurement of the NGS product, FSSPs should also consider the time investment for the following:

- Setting up and validating the NGS instrument.
- Validating the NGS data analysis software.
- Validating the data analysis and interpretation that leads to reporting.
- Integrating the system with a LIMS.
- Training FSSPs on the NGS workflow and data analysis (which may require taking FSSPs away from casework).
- Ensuring that the technology implementation still meets and aligns with accreditation standards.

Bioinformatics and Data Analysis Considerations for NGS

NGS leads to a large amount of data generated for a single sequencing run. Because of the amount and type of data generated by NGS, it is useful for FSSPs to be trained in foundational bioinformatics principles to understand and interpret the data generated. Software can help users of varying technical familiarity with bioinformatics understand, interpret, and draw insights from the data.

Approaches for NGS data analysis can generally be broken down into two main categories:

1. ***Commercial Off-The-Shelf (COTS) software***—Commercial NGS sequencing data analysis programs are readily available and often associated with specific vendor workflows (e.g., Verogen's ForenSeq DNA Signature Prep workflow). COTS software options eliminate the need for FSSPs to read, script, and code complex bioinformatics pipelines because they are preconfigured to the workflow (including the instrumentation and library preparation kits) selected for use. Additionally, some COTS software (e.g., Verogen's UAS for the ForenSeq Kintelligence Kit or ForenSeq Imagen Kit) enables the NGS data generated to be compatible with third-party analytical tools, allowing for further simplified workflows and data use. These software often analyze data on a cloud server, reducing the need for physical hardware to analyze and store generated data.
2. ***Software developed in-house***—Some FSSPs may implement custom panels to analyze specific target regions that are not coupled with a preconfigured software. As a result, these FSSPs may require development of an in-house software that can analyze the resulting data and convert multiple file types between the primary, secondary, and tertiary data analysis steps (see ***Section 4. Data Analysis***). This approach requires FSSPs to possess working knowledge in bioinformatics principles to read, script, and code bioinformatics pipelines, which could be a barrier to adoption and implementation. These software may require data to be analyzed on a local server, which will require additional hardware when developing a digital infrastructure to support generated data.

FSSPs should consider the following questions regarding software adoption and implementation:

- Is there a COTS software that suits our laboratory's needs?
  - If not, is our laboratory equipped to designate individuals who possess foundational bioinformatics knowledge to develop an in-house software?
- Does our NGS instrument and library preparation kit(s) allow for in-house developed software to be integrated within our pre-defined workflow?
- Does our laboratory have existing relationships with NGS vendors that offer support in developing or adapting custom built software?
- Do our laboratory's IT staff possess the necessary digital infrastructure and security expertise to store NGS data on a cloud or local server?

Knowing the current limitations of NGS data use in CODIS.

FSSPs search DNA profiles against criminal justice databases, most notably CODIS. CODIS allows participating FSSPs to upload DNA profiles to the database for comparison against DNA profiles obtained from unidentified human remains, forensic evidence, relatives of missing persons, convicted offenders, and arrestees (if applicable, based on state legislation).[1]

NGS technology generates size-based data for STR analysis, which can be uploaded to CODIS. However, NGS also generates sequence-based data that have not been implemented or approved for upload into the current CODIS framework. As a result, data are restricted to fit the limitations of CODIS; therefore, CODIS cannot be used to store, maintain, or search information such as isoalleles; identity-, ancestry-, or phenotype-, or lineage-informative SNPs; or other forensic markers. Only information from the 20 CODIS Core Loci may be entered into CODIS.[1]

Peer FSSPs and institutions can aid in the decision-making and implementation process.

FSSPs may lean on NGS technology vendors to provide training opportunities. They may also seek help from academic institutions that have previously implemented these systems. The University of Toronto Mississauga and Syracuse University are examples of resources that FSSPs may contact for implementation and validation guidance.

Curriculum and Research Support

The **University of Toronto Mississauga** offers classes and research opportunities using NGS for students in their forensic science program. Within this program, Dr. Nicole Novroski leads the Novroski Forensic Genetics Research Laboratory, an investigative genetics research laboratory that uses the MiSeq FGx Sequencing System (Verogen) as their primary NGS instrument. One of the main objectives of Dr. Novroski's research pertains to the exploration of previously uncharacterized genomic markers for improved DNA mixture deconvolution. Beyond research initiatives, students within the Forensic Anthropology and Forensic Biology Specialist Degree programs can gain hands-on exposure to NGS through specialized coursework.

For FSSPs interested in NGS support during NGS implementation, please contact Dr. Nicole Novroski, Assistant Professor and Forensic Genetics Researcher at Nicole.novroski@utoronto.ca directly, or reach out to forensiccoe@rti.org.

For more information on the University of Toronto Mississauga's NGS curriculum and research, please visit UTM Forensic Science News & Updates.

Syracuse University offers two courses related to NGS and Forensic Applications:

- FSC 474/674L Forensic DNA Analysis, which covers the basics of NGS technology.
- FSC 645 Biochemical Analysis, a hands-on course where students can perform sequencing and analysis using NGS instrumentation.

User Profiles

Subject matter experts shared insights gained from implementing, validating, and using NGS products.

When FSSPs apply NGS technology successfully, they gain insight into procurement and implementation considerations for decision-makers. The user profiles that follow provide examples of successful NGS technology implementation in forensic research and casework settings to illustrate the potential benefits and challenges associated with procuring and implementing NGS technology.

To create these profiles, the FTCOE interviewed a variety of NGS users to determine how the technology has impacted their workflow and capabilities and to understand lessons learned related to procurement, implementation, validation, and use.

The Armed Forces Medical Examiner System – Armed Forces DNA Identification Laboratory implemented Verogen's MiSeq FGx Sequencing System for the identification of human remains.

Lindsay Loughner Kotchey, MSFS, Forensic Scientist II (NGS Analyst)

The Armed Forces Medical Examiner System – Armed Forces DNA Identification Laboratory (AFMES-AFDIL) was the first ANSI National Accreditation Board (ANAB) ISO 17025 and Federal Bureau of Investigation (FBI) Quality Assurance Standards for Forensic DNA Testing Laboratories accredited laboratory to bring a validated NGS method online. However, because of the space demands for NGS sample preparation work and digital infrastructure for data storage (e.g., physical servers), AFMES-AFDIL is still reconfiguring their physical laboratory space. AFMES-AFDIL has implemented five MiSeq FGx Sequencing Systems (Verogen) for use in their Past Accounting Section, which conducts DNA analysis using NGS, Sanger sequencing, auSTR, and Y-STR testing. Samples submitted to the Past Accounting Section are obtained from war memorials or dig sites and may have been exposed to the environment for over 75 years. Samples sent to AFMES-AFDIL undergo initial testing through the application of NGS or Sanger sequencing. Typically, only chemically treated samples are automatically sent for NGS testing, although this may depend on the disinterment from the sample recovery site.

AFMES-AFDIL purchased a MiSeq System (Illumina) in 2012 and dedicated the following 3 years to the research and development of an appropriate NGS method, which included (1) testing commercially available kits and (2) developing in-house methods to process samples when the non-human DNA far exceeds the amount of human DNA (i.e., challenging samples for commercial kits to process by routine methods). In 2013, AFMES-AFDIL purchased a Fragment Analyzer (Agilent) to help determine the average amount of DNA fragments present within sample extractions, which helped optimize the in-house mtDNA capture method. AFMES-AFDIL completed developing the method in 2015, and in September 2015, AFMES-AFDIL purchased three MiSeq FGx instruments that underwent validation and performance checks. Meanwhile, AFMES-AFDIL upgraded two of the pre-existing MiSeq instruments to MiSeq FGx instruments. The validation used a combination of NDIS developmental validation and NDIS internal validation experiments and was completed in December 2016, with the first sample reported in February 2017. AFMES-AFDIL uses the NEBNext Ultra DNA Library Preparation Kit for Illumina (New England Biolabs) followed by the myBaits Custom RNA Probe Array kit (Daicel Arbor Biosciences).

One important piece of the library preparation kit is uracil-specific excision reagent (USER). USER acts as a damage repair tool to remove uracil, a byproduct caused by cytosine deamination, present in samples because of exposure to the environment for extended periods. AFMES-AFDIL chose to develop their own capture assay using RNA probes with a capture hybridization step to enrich human DNA before amplification. Despite the large recovery of information from using the validated NGS method, AFMES-AFDIL's reporting of casework processed by Sanger sequencing and NGS remains the same. Only the CR is reported for privacy considerations and for consistency between NGS and Sanger methods. In the future, AFMES-AFDIL plans to validate methods for nuclear SNP assays for autosomal identification, which is optimal for highly degraded samples.

> *"If you think of CE technology as an optical microscope, NGS would be a scanning electron microscope. It provides you an increase of information, but you have to know how to use it appropriately."*
>
> —Lindsay Loughner Kotchey, Forensic Scientist II (NGS Analyst)

DNA analysts trained in Sanger sequencing required approximately 1 year to be cross-trained to perform NGS; these analysts continued processing casework during their training. Laboratory staff members who assisted in the MiSeq validation process underwent a more streamlined training process because they gained hands-on experience during the validation. However, AFMES-AFDIL has more recently begun to hire DNA analysts directly into NGS analyst roles. NGS technicians responsible for capture events and perform sequencing can be trained in approximately 8 months. Additionally, an NGS analyst responsible for front-end preparation work (e.g., extractions and library preparation) and data analysis, can be trained in approximately 7 months. AFMES-AFDIL's training occurs in house and includes analysis of high-quality and low-quality samples, lectures, a literature review, and a competency examination.

Key Considerations:

- Assessing the laboratory space needed for NGS mtDNA analysis before implementation is beneficial.
- Planning for an increase in data can help gauge digital infrastructure requirements before product procurement.
- Possessing bioinformatics knowledge to understand and interpret data generated by NGS can help streamline validation.

Ohio Bureau of Criminal Investigation implemented Verogen's MiSeq FGx Sequencing System for mtDNA casework.

Adam Garver, MFS, CODIS Forensic Scientist and MPS Technical Leader

The Ohio Bureau of Criminal Investigation (BCI) was one of the first forensic laboratories to use NGS for casework. In 2016, Ohio BCI participated in a technology transfer partnership with Battelle (Columbus, OH), a nonprofit global research and development institute, for the procurement of a MiSeq FGx Sequencing System (Verogen). Ohio BCI selected the MiSeq FGx because Battelle had experience with this instrument and its compatibility with external, non–vendor-specific library preparation kits. In addition to the MiSeq FGx, Ohio BCI purchased supporting laboratory equipment (an Agilent 2100 Bioanalyzer System and a Thermo Fisher Scientific Qubit Fluorometer) to help troubleshoot during validation and implementation.

A team of two Ohio BCI analysts worked collaboratively with two Battelle analysts to validate the MiSeq FGx for use with two library preparation kits, the PowerSeq Nested CRM System (Promega) and the ForenSeq DNA Signature Prep Kit (Verogen). The analysts validated the ForenSeq kit by adapting a validation plan provided by Verogen. Validation took approximately 6–8 months, with ~2 months dedicated to completing wet laboratory work. The PowerSeq kit validation, using a plan developed in house, took approximately 1 year (with ~4 months dedicated to completing wet laboratory work). Subsequently, the validation data for both kits were submitted to NDIS and were approved for casework. Despite having validated the ForenSeq kit for STR analysis, Ohio BCI is currently only using

the PowerSeq kit, which was brought online in 2020, for mtDNA casework. However, Ohio BCI has a strong interest in implementing NGS for STRs in the future.

Ohio BCI is using the PowerSeq kit for missing persons and unidentified human remains casework. Before implementing the MiSeq FGx, Ohio BCI was unable to perform mtDNA testing. As such, unidentified human remains cases were routinely outsourced to the University of North Texas (UNT). Upon bringing NGS online, Ohio BCI saw a decrease in the turnaround times for these cases and received an influx in additional requests because Ohio BCI now had in-house NGS capabilities and did not depend on outsourcing. Currently, NGS is only being used for missing persons and unidentified human remains cases, and Ohio BCI analysts have yet to testify on any NGS processed case samples.

> *"The PowerSeq kit was easy to implement because it offered compatibility with standard extractions, one amplification step, and simple library clean up steps."*
>
> —Adam Garver,
> CODIS Forensic Scientist and MPS Technical Leader

When Ohio BCI first procured the MiSeq FGx, limited educational and training resources were readily available. Ohio BCI frequently used vendor contacts (i.e., Verogen and Promega) and partners and colleagues (e.g., Battelle, California Department of Justice [DOJ], UNT, and the FBI) for assistance during instrument and kit validation and implementation. One challenge encountered transitioning from CE to NGS was the large increase in data generated. Analysts had to learn which part of the data was important, emphasizing the value of an understanding and knowledge in bioinformatics.

Key Considerations:

- Work with IT to develop an infrastructure for backing up and storing data prior to NGS procurement.
- Automation, such as robotics and automated liquid handlers, can be used to increase sample throughput or manage influxes of processing requests.
- Reaching out to experienced vendors, collaborators, and colleagues can greatly assist with validation and implementation.

California Department of Justice implemented Thermo Fisher Scientific's Ion GeneStudio S5 System with the Ion Chef System for mtDNA casework.

Jeanette Wallin, MPH, Criminalist Supervisor, Method Development Group; Bill Hudlow, MS Criminalist Supervisor and Technical Leader, Missing Persons DNA Program; Michelle Halsing, BS, Criminalist Supervisor and MPS Casework Analyst, Missing Persons DNA Program; Steven Myers, MS, Senior Criminalist, Method Development Group; Daniela Cuenca, MS, Senior Criminalist, Method Development Group

The California DOJ was an early adopter of NGS technology. In 2014, California DOJ purchased a MiSeq System (Illumina) for STR and SNP testing and an Ion GeneStudio S5 with the Ion Chef System (Thermo Fisher Scientific) for mtDNA analysis in 2016. In 2017, California DOJ procured a MiSeq FGx Sequencing System (Verogen) to upgrade to the latest available technology and system capabilities.

California DOJ has a dedicated Method Development Group responsible for validating new DNA methods and instrumentation for criminal casework and for the Missing Persons DNA Program (MPDP) and Data Bank programs, which are statewide and overseen by the California DOJ. Because NGS validation guidelines were not available at the time of procurement, the Method Development Group determined what was necessary for validation by convening conversations with other forensic scientists, including the Scientific Working Group on DNA Analysis Methods NGS Working Group.

The Ion GeneStudio S5 with the Ion Chef System was validated with the Precision ID mtDNA Whole Genome Panel (Thermo Fisher Scientific) as a fully automated system applied to MPDP casework, which took slightly over 2 years. During validation, California DOJ tested both challenging mock samples and non-probative samples from the MPDP. The implementation of a clean-up procedure following extraction improved the recovery of mtDNA from inhibited samples, and the smaller amplicon size enabled enhanced information recovery compared with previous methods. Because California DOJ was an early NGS adopter, parts of the validation were repeated because reagent formulations changed and software was being developed simultaneously. Although the ForenSeq DNA Signature Prep Kit (Verogen) validation with the MiSeq FGx remains ongoing while California DOJ seeks to implement robotics and validate statistical analysis software for SNPs, the wet laboratory validation work required 2 years to complete. The implementation of this kit at California DOJ is contingent upon the development of kinship testing software that incorporates substructure, meiotic mutation, and linkage analyses for MPDP use and improved methods for mixture deconvolution for criminal casework.

"With more recent availability of recommendations, resources, and support from other users, laboratories looking to implement NGS can now undergo a more streamlined process."

—Jeanette Wallin,
Criminalist Supervisor

The transition from Sanger sequencing to NGS for mtDNA testing in the MPDP allowed for more streamlined sample processing with less time spent in the laboratory. Additionally, this technology provided a conducive workflow that allowed staggered work schedules and remote analyses, which was advantageous during the COVID-19 pandemic. California DOJ continues to evaluate other potential applications for NGS, including FGG, mixture analysis, and familial searches of the convicted offender data bank, although challenges with mixture interpretation and probabilistic genotyping need to be addressed because commercially available products do not currently exist.

Key Considerations:

- Understanding laboratory priorities can help determine the most suitable NGS system for procurement.
- Dedicating ample resources to NGS validation and implementation can help overcome the technology transformation hurdles.
- Using validation samples that are representative of casework is key to exploring the capabilities and limitations of a given method.

DNA Labs International implemented Verogen's MiSeq FGx Sequencing System for application to human identification and criminal investigation casework.

DNA LABS INTERNATIONAL
ISO/IEC 17025:2017 and FBI QAS DNA Laboratory

Rachel Oefelein, MSc, Chief Scientific Officer

DNA Labs International is a private forensic DNA laboratory, accredited under ANAB ISO 17025, that services agencies across the United States and Caribbean Nations. DNA Labs International purchased a MiSeq FGx Sequencing System (Verogen) in January 2020 after reviewing NGS instruments available on the market and surveying products in use by publicly funded laboratories. The flexibility offered by the MiSeq FGx for sequencing options (e.g., from whole genome sequencing to SNPs) and the compatibility with library preparation kits offered by other vendors helped inform the purchasing decision.

DNA Labs International purchased three Verogen library preparation kits, including the ForenSeq Kintelligence Kit for FGG application to violent crime cold cases; the ForenSeq DNA Signature Prep Kit/Primer Mix B to recover identity-, phenotypic-, and biogeographical ancestry–informative SNPs for criminal investigation casework; and the ForenSeq mtDNA Whole Genome Kit for missing persons and unidentified human remains casework. The validations

were completed by the DNA Labs International with assistance from Verogen. The wet laboratory work for the validations was completed in house and required approximately 3 months for each kit, with each kit-specific report taking approximately 2 months to complete. Overall, it took 2 years for all components to be validated and audited for compliance with accreditation standards. DNA Labs International intends to purchase and validate the ForenSeq MainstAY Kit (Verogen) in the future for autosomal and Y-STR analysis for agencies that are not interested in obtaining SNP information.

Also in 2020, DNA Labs International purchased an Auroa Biomed VERSA 1100 Automated NGS Library Preparation System to support workflow automation. This system was found to be advantageous because it is pre-scripted for NGS application, which reduced the time needed to incorporate robotic features into the workflow. Both automated and manual processing were internally validated. For data analysis, DNA Labs International uses Verogen's UAS with one module supporting the ForenSeq DNA Signature Prep Kit and another module supporting both the ForenSeq mtDNA Whole Genome Kit and the Kintelligence Kit.

> *"Completing wet laboratory validation work in-house is a great opportunity for staff members to gain firsthand NGS training and experience."*
>
> —Rachel Oefelein,
> Chief Scientific Officer

Once the NGS implementation was made public, DNA Labs International saw a large influx of casework processing requests. Complete sample processing requires approximately 10 hours of hands-on time for sample preparation (including thermocycling) and 24 hours of runtime on the MiSeq FGx to generate sequence data. There is a large cost associated with this runtime; however, more information (including autosomal, X-STRs, Y-STRs, and SNPs) can be obtained from a limited amount of DNA compared with traditional CE methods by using the ForenSeq DNA Signature Prep Kit.

Key Considerations:

- Performing a cost-benefit analysis of instruments and library preparation kits before procurement can be advantageous.
- Considering automated systems for sample preparation and pre-scripted robotics can reduce implementation time.
- Adhering to all accreditation standards when implementing NGS technology can support future acceptance of NGS technology into court.

MiSeq FGx Sequencing System to further explore mixture interpretation.

Michael Marciano, PhD, Research Associate Professor and Director of Research, Forensic & National Security Sciences Institute

Syracuse University College of Arts & Sciences
Forensic & National Security Sciences Institute

Syracuse University adopted the MiSeq FGx Sequencing System (Verogen) in 2016 through the Department of Defense's Defense University Research Instrumentation Program. In 2018, research professor Dr. Michael Marciano and colleague Jonathan Adelman were awarded a 2-year grant from the NIJ (award #2018-DU-BX-0202) to expand on previous research on an automated separation technique for mixture data using CE by examining STR length and sequence variations from NGS data. The research examined an approach for predicting the number of contributors in forensic mixture samples from STR length and sequence variations from NGS data. The goal of the research was to develop a fully continuous machine learning probabilistic approach to predict the number of contributors in forensic mixture samples from STR length and sequence variations from NGS data. This project examined extensive validation data sets provided by Verogen, Promega, and numerous forensic laboratories. Although this project focused more on data interpretation and did not have a wet laboratory component, other projects completed within the research laboratory have used both the MiSeq FGx Forensic Genomics and Research Use Only modes.

Dr. Marciano's research group continues to analyze mixture and single-source human samples under a variety of conditions (e.g., samples with high degradation and inhibition).

When analyzing mixture samples using NGS, it is important to recognize limitations given the size restrictions of the flow cell for sample loading, which can result in decreased sensitivity for core loci as additional markers are amplified. Dr. Marciano's research group found that limiting the number of samples sequenced when analyzing mixture samples helped avoid potential allelic dropout and coverage issues. Additionally, this research found that using an STR-specific library preparation kit may be better suited for complex mixture sample analysis.

> *"Sometimes less can be more—we do not need to target every DNA marker available just to get the most information. We already have a lot of information in current sequencing data that is not being fully utilized."*
>
> —Michael Marciano,
> Research Associate Professor

Dr. Marciano advised that funding support from agencies like the NIJ helps promote the adoption of NGS technology into FSSPs and that although there is still work needed, such as addressing naming conventions and the application of probabilistic genotyping, the use of NGS in forensic science provides many opportunities to advance the field.

Key Considerations:

- Evaluating specific laboratory needs can inform which markers are of interest and can help determine which library preparation kit is best suited to obtain this information.
- Researching the longevity of NGS instruments can inform product decisions.
- Considering instruments that offer automation capabilities can be advantageous when comparing products.

Conclusion

NGS enables FSSPs to assess multiple markers of forensic relevance (e.g., autosomal STRs, X-STRs, Y-STRs, mini-STRs, SNPs) using a single assay within a shorter timeframe than previous sequencing techniques, technologies, and methods, making this technology beneficial for the forensic community in terms of cost, labor, and time savings.[2] NGS increases the power of discrimination by providing both sequence- and size-based information, which exploits isoalleles and can help obtain more data related to the genotype of the DNA donor. NGS also allows primers to be placed closer to target regions, which can be advantageous for challenging (e.g., degraded, inhibited, compromised) samples. Finally, NGS technology can sequence identity-, ancestry-, phenotype-, and lineage-informative SNPs and a constantly expanding panel of additional identity-informative markers that can be used to predict phenotype, ancestry, genealogy, and parentage to develop investigative leads.

A targeted DNA NGS workflow consists of four main phases: sample preparation, library preparation, sequencing, and DNA analysis. The two forensic-specific NGS systems currently commercially available include the Ion GeneStudio S5 System by Thermo Fisher Scientific and the MiSeq FGx Sequencing System by Verogen. Both vendors provide sequencing instrumentation, library preparation kits, software options, and automation options compatible with their respective workflows. Promega offers library preparation kits compatible with the MiSeq FGx. These three vendors supply kits have received NDIS approval. Additionally, both Verogen and Thermo Fisher Scientific offer support for system validation and training to ease workflow transitions and technology implementation.

Although NGS is still in an early adoption phase with FSSPs because of upfront investment, validation, and accreditation requirements, the application of NGS within forensic science has allowed the community to obtain higher levels of discrimination and discover novel and forensically relevant identity-informative markers. As more FSSPs implement NGS, the increase in published validation studies and publicly available information will help ease the transition from PCR-CE to NGS technology. The increase in technology adoption and open-access validation data can in turn lead to general acceptance of NGS within the court system. Given the adolescence of NGS in the forensic community, there is likely unknown potential, demonstrating the importance of research into the capabilities and limitations of this technology both for human identification and alternative forensic applications.

References

1. FBI. "Frequently asked questions on CODIS and NDIS." Accessed Aug 24, 2022. Federal Bureau of Investigation, https://www.fbi.gov/services/laboratory/biometric-analysis/codis/codis-and-ndis-fact-sheet.
2. Børsting, C, and Morling, N. "Use of next-generation sequencing in forensic genetics." (17 April 2017). https://doi.org/10.1002/9780470015902.a0027106.
3. National Institutes of Justice. "What is STR analysis?" Accessed Aug 23, 2022. NIJ, https://nij.ojp.gov/topics/articles/what-str-analysis.
4. Forensic Laboratory Needs Technology Working Group, (FLN-TWG), ,. Implementation strategies: Next generation sequencing for DNA analysis Prefix *Implementation strategies: Next generation sequencing for DNA analysis.* Forensic Technology Center of Excellence. US Department of Justice, National Institute of Justice, Office of Investigative and Forensic Sciences, 2021. https://forensiccoe.org/private/6171719a25bfc.
5. Illumina. "Key differences between next-generation sequencing and Sanger sequencing: Understanding when NGS can be a more effective option." Accessed Aug 23, 2022. Illumina, Inc., https://www.illumina.com/science/technology/next-generation-sequencing/ngs-vs-sanger-sequencing.html.
6. Volk, A, Bintz, B. J, and Grisedale, K. S. "Sensitivity Comparison of the Illumina® ForenSeq™ DNA Signature Prep kit and the Applied Biosystems® GlobalFiler™ kit." Accessed 10 March, 2023. Western Carolina University, https://verogen.com/wp-content/uploads/2018/03/AAFS-Poster_Volk.pdf.
7. Daniel, R., Santos, C., Phillips, C., Fondevila, M., van Oorschot, R. A., Carracedo, A., Lareu, M. V., and McNevin, D. "A SNaPshot of next generation sequencing for forensic SNP analysis." [In eng]. *Forensic Sci Int Genet* 14 (Jan 2015): 50-60. https://doi.org/10.1016/j.fsigen.2014.08.013.
8. Kchouk, Mehdi, Gibrat, Jean-François, and Elloumi, Mourad. "Generations of sequencing technologies: From first to next generation." *Biology and medicine* 9 (2017): 1-8.
9. Yang, Yaran, Xie, Bingbing, and Yan, Jiangwei. "Application of next-generation sequencing technology in forensic science." *Genomics, Proteomics & Bioinformatics* 12, no. 5 (2014/10/01/ 2014): 190-97. https://doi.org/10.1016/j.gpb.2014.09.001.
10. Revoir, A. P., Lancaster, H., and Ames, C. "Initial assessment of NGS as a tool to augment routine CE analysis of STRs." *Forensic Science International: Genetics Supplement Series* 7, no. 1 (2019/12/01/ 2019): 747-49. https://doi.org/10.1016/j.fsigss.2019.10.162.
11. Forsythe, Bethany, Melia, Lisa, and Harbison, SallyAnn. "Methods for the analysis of mitochondrial DNA." *WIREs Forensic Science* 3, no. 1 (2021): e1388. https://doi.org/10.1002/wfs2.1388.
12. Brandhagen, M. D., Just, R. S., and Irwin, J. A. "Validation of NGS for mitochondrial DNA casework at the FBI laboratory." *Forensic Sci Int Genet* 44 (Jan 2020): 102151. https://doi.org/10.1016/j.fsigen.2019.102151.
13. Kim, H., Erlich, H. A., and Calloway, C. D. "Analysis of mixtures using next generation sequencing of mitochondrial DNA hypervariable regions." [In eng]. *Croat Med J* 56, no. 3 (Jun 2015): 208-17. https://doi.org/10.3325/cmj.2015.56.208.
14. Naue, J., Hoefsloot, H. C. J., Mook, O. R. F., Rijlaarsdam-Hoekstra, L., van der Zwalm, M. C. H., Henneman, P., Kloosterman, A. D., and Verschure, P. J. "Chronological age prediction based on DNA methylation: Massive parallel sequencing and random forest regression." *Forensic Sci Int Genet* 31 (Nov 2017): 19-28. https://doi.org/10.1016/j.fsigen.2017.07.015.
15. Aliferi, Anastasia, Ballard, David, Gallidabino, Matteo D., Thurtle, Helen, Barron, Leon, and Syndercombe Court, Denise. "DNA methylation-based age prediction using massively parallel sequencing data and

multiple machine learning models." *Forensic Science International: Genetics* 37 (2018/11/01/ 2018): 215-26. https://doi.org/10.1016/j.fsigen.2018.09.003.

16. Vidaki, A., Ballard, D., Aliferi, A., Miller, T. H., Barron, L. P., and Syndercombe Court, D. "DNA methylation-based forensic age prediction using artificial neural networks and next generation sequencing." [In eng]. *Forensic Sci Int Genet* 28 (May 2017): 225-36. https://doi.org/10.1016/j.fsigen.2017.02.009.
17. Turchi, Chiara, Onofri, Valerio, Melchionda, Filomena, Fattorini, Paolo, and Tagliabracci, Adriano. "Development of a forensic DNA phenotyping panel using massive parallel sequencing." *Forensic Science International: Genetics Supplement Series* 7, no. 1 (2019/12/01/ 2019): 177-79. https://doi.org/10.1016/j.fsigss.2019.09.069.
18. Phillips, C., Salas, A., Sanchez, J. J., Fondevila, M., Gomez-Tato, A., Alvarez-Dios, J., Calaza, M.*, et al.* "Inferring ancestral origin using a single multiplex assay of ancestry-informative marker SNPs." *Forensic Sci Int Genet* 1, no. 3-4 (Dec 2007): 273-80. https://doi.org/10.1016/j.fsigen.2007.06.008.
19. Walsh, S., Lindenbergh, A., Zuniga, S. B., Sijen, T., de Knijff, P., Kayser, M., and Ballantyne, K. N. "Developmental validation of the IrisPlex system: determination of blue and brown iris colour for forensic intelligence." *Forensic Sci Int Genet* 5, no. 5 (Nov 2011): 464-71. https://doi.org/10.1016/j.fsigen.2010.09.008.
20. Kling, D., Phillips, C., Kennett, D., and Tillmar, A. "Investigative genetic genealogy: current methods, knowledge and practice." *Forensic Sci Int Genet* 52 (May 2021): 102474. https://doi.org/10.1016/j.fsigen.2021.102474.
21. Verogen. "ForenSeq Kintelligence Kit." Accessed Aug 23, 2022. Verogen, https://verogen.com/products/forenseq-kintelligence-kit/.
22. Johnson, A. J., Vangay, P., Al-Ghalith, G. A., Hillmann, B. M., Ward, T. L., Shields-Cutler, R. R., Kim, A. D.*, et al.* "Daily sampling reveals personalized diet-microbiome associations in humans." [In eng]. *Cell Host Microbe* 25, no. 6 (Jun 12 2019): 789-802.e5. https://doi.org/10.1016/j.chom.2019.05.005.
23. Singh, B., and Publow, A. Forensic body fluid identification using microbiome signature attribution Prefix *Forensic body fluid identification using microbiome signature attribution.* Richmond, VA: Virginia Commonwealth University, 2021. 256087 https://www.ojp.gov/pdffiles1/nij/grants/256087.pdf.
24. Wang, Shuangshuang, Song, Feng, Wang, Yanyun, Huang, Yun, Xie, Bowen, and Luo, Haibo. "High resolution melting analysis (HRM) based on 16SrRNA as a tool for personal identification with the human oral microbiome." *Forensic Science International: Genetics Supplement Series* 7, no. 1 (2019/12/01/ 2019): 161-63. https://doi.org/10.1016/j.fsigss.2019.09.063.
25. Woerner, A. E., Novroski, N. M. M., Wendt, F. R., Ambers, A., Wiley, R., Schmedes, S. E., and Budowle, B. "Forensic human identification with targeted microbiome markers using nearest neighbor classification." *Forensic Sci Int Genet* 38 (Jan 2019): 130-39. https://doi.org/10.1016/j.fsigen.2018.10.003.
26. Wang, Y., Wang, M., Xu, W., Wang, Y., Zhang, Y., and Wang, J. "Estimating the postmortem interval of carcasses in the water using the carrion insect, brain tissue RNA, bacterial biofilm, and algae." [In eng]. *Front Microbiol* 12 (2021): 774276. https://doi.org/10.3389/fmicb.2021.774276.
27. Hanson, E., and Ballantyne, J. "Human organ tissue identification by targeted RNA deep sequencing to aid the investigation of traumatic injury." *Genes (Basel)* 8, no. 11 (Nov 10 2017). https://doi.org/10.3390/genes8110319.
28. Illumina. "NGS for studying the human microbiome." Accessed Aug 24, 2022. Illumina, Inc., https://www.illumina.com/areas-of-interest/microbiology/human-microbiome-analysis.html.
29. Ingold, S., Dørum, G., Hanson, E., Berti, A., Branicki, W., Brito, P., Elsmore, P.*, et al.* "Body fluid identification using a targeted mRNA massively parallel sequencing approach - Results of a EUROFORGEN/EDNAP collaborative exercise." [In eng]. *Forensic Sci Int Genet* 34 (May 2018): 105-15. https://doi.org/10.1016/j.fsigen.2018.01.002.
30. Johnson, Abigail J., Vangay, Pajau, Al-Ghalith, Gabriel A., Hillmann, Benjamin M., Ward, Tonya L., Shields-Cutler, Robin R., Kim, Austin D.*, et al.* "Daily Sampling Reveals Personalized Diet-Microbiome Associations

in Humans." *Cell Host & Microbe* 25, no. 6 (2019/06/12/ 2019): 789-802.e5. https:/doi.org/10.1016/j.chom.2019.05.005.

31. DeBruyn, Jennifer M., and Hauther, Kathleen A. "Postmortem succession of gut microbial communities in deceased human subjects." [In English]. *PeerJ* (2017 Jun 12 2018-01-04 2017). https://doi.org/10.7717/peerj.3437.
32. Johnson, Hunter R., Trinidad, Donovan D., Guzman, Stephania, Khan, Zenab, Parziale, James V., DeBruyn, Jennifer M., and Lents, Nathan H. "A Machine Learning Approach for Using the Postmortem Skin Microbiome to Estimate the Postmortem Interval." *PLOS ONE* 11, no. 12 (2016): e0167370. https://doi.org/10.1371/journal.pone.0167370.
33. Ballard, David, Winkler-Galicki, Jakub, and Wesoły, Joanna. "Massive parallel sequencing in forensics: Advantages, issues, technicalities, and prospects." *International Journal of Legal Medicine* 134, no. 4 (2020/07/01 2020): 1291-303. https://doi.org/10.1007/s00414-020-02294-0.
34. Illumina. "Sequencing quality scores." Accessed Aug 24, 2022. Illumina, Inc., https://www.illumina.com/science/technology/next-generation-sequencing/plan-experiments/quality-scores.html.
35. Young, B., Marciano, M., Crenshaw, K., Duncan, G., Armogida, L., and McCord, B. "Match statistics for sequence-based alleles in profiles from forensic PCR-mps kits." [In eng]. *Electrophoresis* 42, no. 6 (Mar 2021): 756-65. https://doi.org/10.1002/elps.202000087.
36. Gettings, K. B., Borsuk, L. A., Steffen, C. R., Kiesler, K. M., and Vallone, P. M. "Sequence-based U.S. population data for 27 autosomal STR loci." [In eng]. *Forensic Sci Int Genet* 37 (Nov 2018): 106-15. https://doi.org/10.1016/j.fsigen.2018.07.013.
37. Bodner, M., Bastisch, I., Butler, J. M., Fimmers, R., Gill, P., Gusmao, L., Morling, N.*, et al.* "Recommendations of the DNA Commission of the International Society for Forensic Genetics (ISFG) on quality control of autosomal Short Tandem Repeat allele frequency databasing (STRidER)." *Forensic Sci Int Genet* 24 (Sep 2016): 97-102. https://doi.org/10.1016/j.fsigen.2016.06.008.
38. Verogen. "First criminal conviction secured with next-gen forensic DNA technology." (2019). Published electronically Jun 4. https://verogen.com/first-criminal-conviction-with-next-gen-forensic-dna/.
39. Verogen. "ForenSeq Imagen Kit." Accessed Dec 21, 2022. Verogen, Inc., https://verogen.com/products/forenseq-imagen-kit/.
40. Verogen. "Verogen PrepStation." Accessed Dec 21, 2022. Verogen, Inc., https://verogen.com/products/verogen-prepstation/.
41. Verogen. "Universal Analysis Software." Accessed Dec 21, 2022. Verogen, Inc., https://verogen.com/products/universal-analysis-software/.
42. Bureau of Justice Assistance. "National Sexual Assault Kit Initiative (SAKI)." Last modified October 18, 2022. Accessed 10 March, 2023. Bureau of Justice Assistance, https://bja.ojp.gov/program/saki/overview.
43. Bureau of Justice Assistance. "Prosecuting Cold Cases Using DNA." Last modified June 14, 2022. Accessed 10 March, 2023. Bureau of Justice Assistance, https://bja.ojp.gov/program/prosecuting-cold-cases-using-dna/overview.
44. Bureau of Justice Assistance. "Paul Coverdell Forensic Science Improvement Grants Program." Last modified July 15, 2022. Accessed 10 March, 2023. Bureau of Justice Assistance, https://bja.ojp.gov/program/coverdell/overview.
45. Bureau of Justice Assistance. "DNA Capacity Enhancement for Backlog Reduction (CEBR) Program." Last modified February 9, 2023. Accessed 10 March, 2023. Bureau of Justice Assistance, https://bja.ojp.gov/program/dna-cebr/overview.
46. National Institute of Justice. "Forensic Science Research and Development." Last modified 21 Oct. Accessed 10 March, 2023. National Institute of Justice, https://nij.ojp.gov/topics/forensics/forensic-science-research-and-development.

The Forensic Technology Center of Excellence

RTI International (RTI) and its academic and community based-consortium of partnerships, including its Forensic Science Education Programs Accreditation Commission partners, work to meet all tasks and objectives put forward under NIJ's Forensic Technology Center of Excellence (FTCOE). These efforts include determining technology needs; developing technology program plans to address those needs; developing solutions; demonstrating, testing, evaluating, and adopting potential solutions into practice; developing and updating technology guidelines; and building capacity and conducting outreach. The FTCOE is led by RTI, a global research institute dedicated to improving the human condition by turning knowledge into practice. The FTCOE builds on RTI's expertise in forensic science, innovation, technology application, economics, data analytics, statistics, program evaluation, public health, and information science.

NIJ | National Institute of Justice
STRENGTHEN SCIENCE. ADVANCE JUSTICE.

Forensic Technology
CENTER OF EXCELLENCE

RTI INTERNATIONAL

Disclaimer

The FTCOE, led by RTI International, is supported through a Cooperative Agreement from the National Institute of Justice (15PNIJ-21-GK-02192-MUMU; 2016-MU-BX-K110), Office of Justice Programs, U.S. Department of Justice. Neither the U.S. Department of Justice nor any of its components operate, control, are responsible for, or necessarily endorse, this landscape study. Information provided herein is intended to be objective and is based on data collected during primary and secondary research efforts available at the time this report was written. The information provided herein is intended to provide a brief overview of currently available products on the market, as well as those in development; it is not intended as an exhaustive summary. Features or capabilities of additional products identified outside of this landscape may be compared with these product offerings to aid in the information-gathering or decision-making processes.