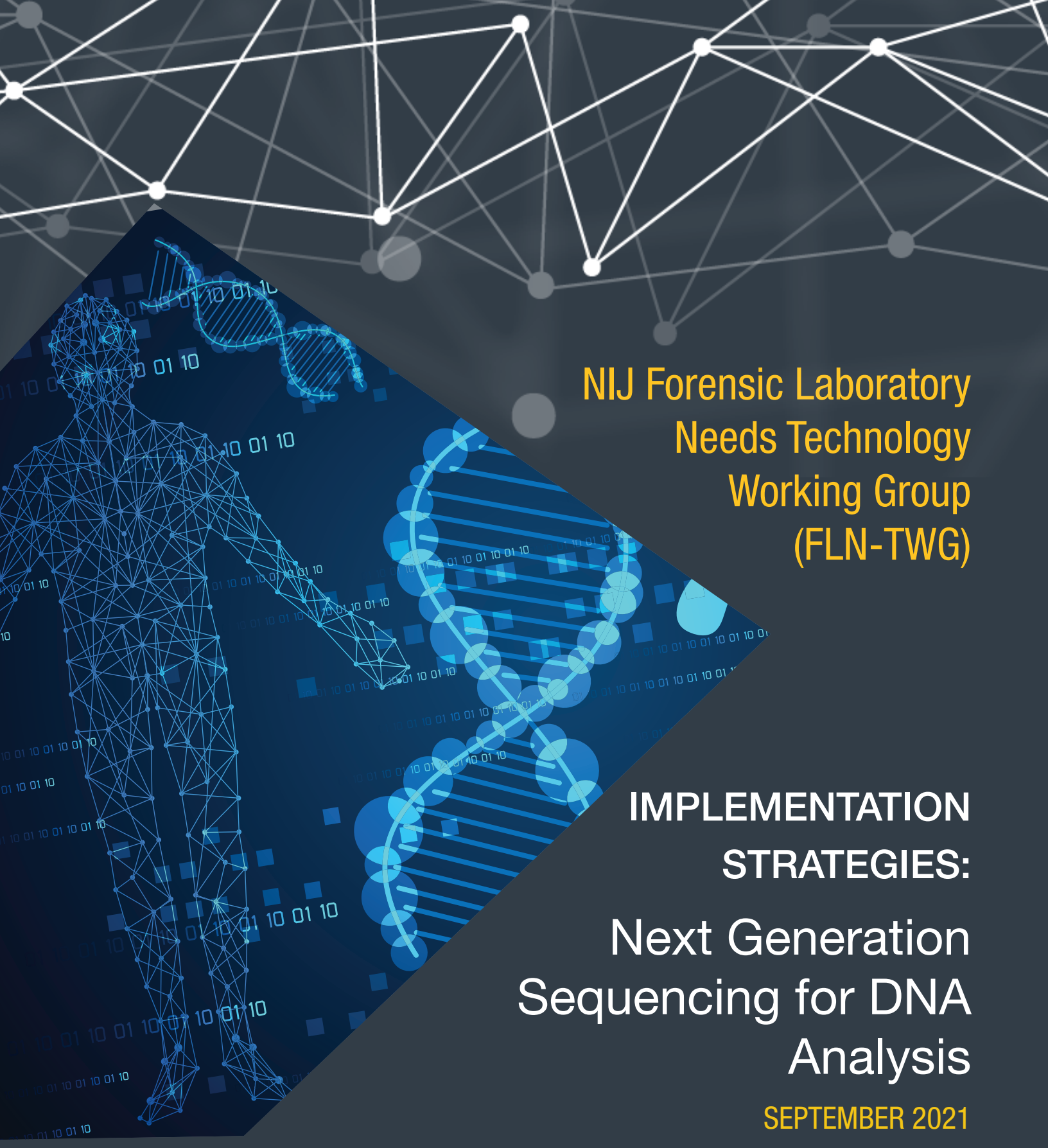


NIJ Forensic Laboratory Needs Technology Working Group (FLN-TWG)

IMPLEMENTATION STRATEGIES:

Next Generation Sequencing for DNA Analysis

SEPTEMBER 2021

The Forensic Laboratory Needs Technology Working Group (FLN-TWG) developed this Implementation Strategy. The FLN-TWG is an activity administered under the National Institute of Justice (NIJ) Forensic Technology Center of Excellence (FTCoE) program. RTI International leads the FTCoE, which is supported through an NIJ Cooperative Agreement (2016-MU-BX-K110), Office of Justice Programs, U.S. Department of Justice (DOJ). Any opinions or points of view expressed in this white paper are those of the FLN-TWG and do not necessarily reflect the official position or policies of NIJ or the DOJ.

Table of Contents

Tables

Figures

Executive Summary

Modern sequencing technology offers several enhancements over traditional DNA analysis methods: parallel analysis of multiple marker types; improved performance with degraded DNA; increased discrimination power and mixture resolution capability through Short Tandem Repeats (STR) sequence variant detection; capability to infer phenotype, ancestry, genealogy, or parentage; and other applications. The massive sequence output of next generation sequencing (NGS) instruments may be divided among many samples for routine testing (i.e., databasing) or focused to analyze a small number of challenging evidence samples for comprehensively analyzing many marker types. Two platforms have kits that have received acceptance for use in National DNA Index System (NDIS) submissions and thus dominate the forensic DNA landscape: the Ion Torrent S5 (Thermo Fisher Scientific) and the MiSeq FGx (Verogen, Inc.). Capital equipment acquisition cost is approximately $150,000 with annual maintenance agreement costs of approximately $15,000. Increased complexity, reagent costs, and labor input may limit NGS use to challenging cases where augmentation of current STR typing methods is needed. Higher throughput usage may necessitate automation of sample preparation and add requirements for large-scale data storage infrastructure.

Description and Overview

DNA analysis is an invaluable component of forensic investigations. Currently, capillary electrophoresis (CE) is the most widely used method to genotype STRs for investigative purposes. STRs are short DNA sequences that contain repeated sequence motifs 2–7 nucleotides in length. The number of repeats may vary among individuals, allowing matching or exclusion of identity in forensic investigations. Currently, the U.S. Federal Bureau of Investigation (FBI) has defined 20 core STR loci accepted for use in NDIS for investigative DNA analysis.[1] STR genotyping uses polymerase chain reaction (PCR) to amplify a DNA region inclusive of the STR. DNA is first extracted and purified from either a crime scene sample (questioned sample) or reference sample (known sample). The content in the extract is then quantified using human-specific quantitative PCR. A specific amount of DNA, usually 0.5 or 1.0 ng, is added to a PCR reaction containing DNA primers for multiple STR loci, with one primer in each primer pair fluorescently labeled to allow detection by the CE instrument and increased multiplexing of STR loci with different fluorescent color channels. The resulting PCR fragment is measured using CE to determine fragment length, which is then correlated with the number of tandem repeats during the interpretation of the results. Typically, standards with known repeat number (allelic) values are included in the analysis, and positive and negative controls are included to verify proper function of the method and instrumentation.

[1] Hares, D. R. (2015, Jul). Selection and implementation of expanded CODIS core loci in the United States. *Forensic Sci Int Genet, 17*, 33-34. https://doi.org/10.1016/j.fsigen.2015.03.006. PMID: 25797140.

Discrimination between individuals by current STR CE typing yields high statistical confidence estimates, and the number of STR profiles contained in national databases ensures that current state-of-the-art methodology will remain in mainstream use. However, the continually increasing use of DNA-based analysis in casework presents many challenging scenarios involving trace amounts of biological evidence, mixed-source evidence, contamination, degraded DNA, and complex analyses related to parentage and ancestry, missing persons and mass disaster victim identification, monozygotic twins, and multiple-contributor samples. These difficult applications highlight some limitations in current STR analysis methods. Improvements have been made to CE-based methods, such as increased numbers of fluorescent channels and chemistry for evaluating single nucleotide polymorphisms (SNPs). Ultimately, CE is limited in the number of loci that can fit into a window of about 400 base pairs (bp), with the most advanced instruments capable of detecting of six to eight dye channels. A key consideration is that spatial separation on CE necessitates PCR amplicon design include some flanking sequence beyond the STR region. This causes the telltale "ski slope" effect when investigating degraded DNA, where larger amplicons perform less well than smaller sized ones. In the past decade, however, new technology has become available that can overcome this limitation while expanding the repertoire and value of forensic DNA analysis.

Contemporary DNA sequencing technology, referred to as NGS or massively parallel sequencing, encompasses both an approach and a host of instruments that enable simultaneous analysis of large amounts of genetic information. Although a large NGS instrument can obtain enough sequence to cover several human genomes, smaller bench-top models have capacity to measure hundreds to thousands of informative sites for several samples (usually 10 to 96 simultaneously). Markers may include STRs (plus any sequence variation within or nearby the repeat region) and SNPs in nuclear and mitochondrial genomes. NGS-based characterization of sequence information from multiple marker types increases discrimination power from a single analysis and provides new capabilities to augment current DNA methodology and to provide investigative intelligence (see Appendix A for a listing of marker types and applications). This paper provides guidance for implementing and using DNA sequencing technology, and the focus will remain on human-specific DNA analysis applications.

Considerations for NGS implementation and use as a complementary method with improved performance over conventional CE-based DNA analysis in the aforementioned areas are presented below.

NGS—How it Works and Advantages for DNA Analysis

NGS refers to advancements in sequencing techniques beyond the method of dideoxy sequencing developed by Sanger et al. in 1977.[2] Many improvements have been achieved for Sanger sequencing, including multi-color fluorescent dye-terminator dideoxy nucleotide

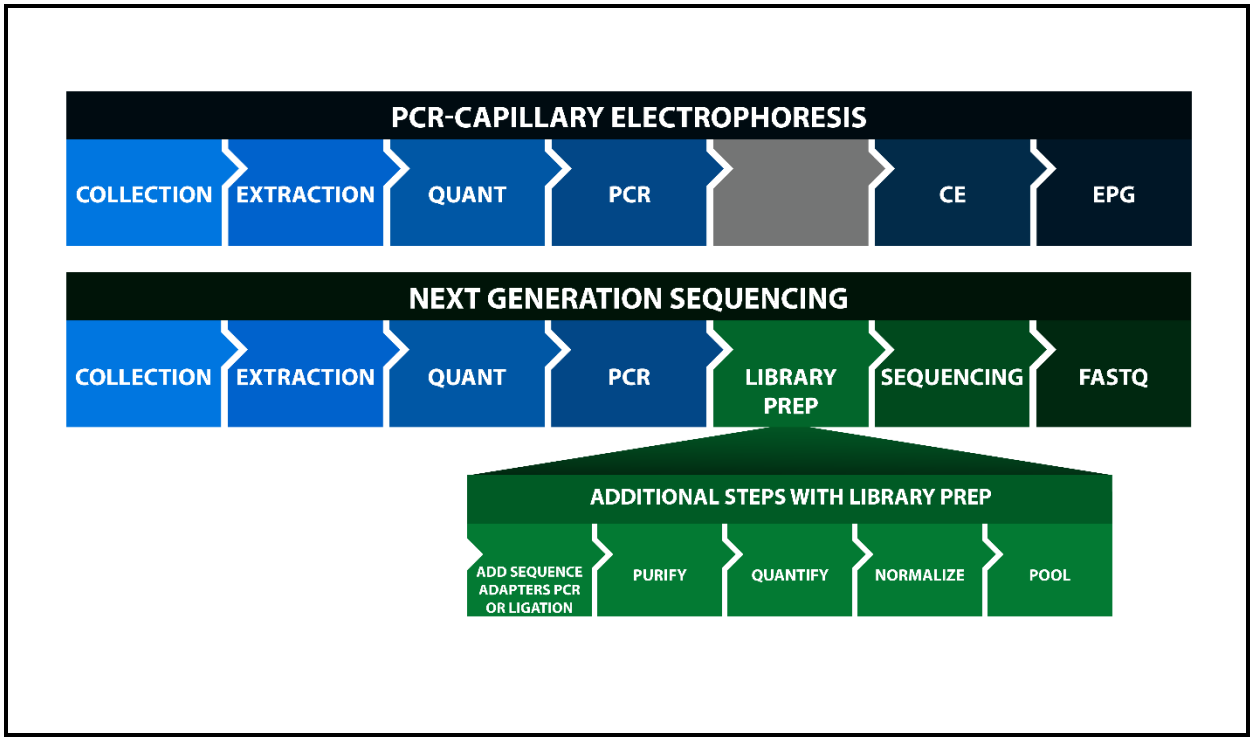
[2] Sanger, F., Nicklen, S., & Coulson, A. R. (1977, Dec). DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A, 74*(12), 5463-5467. https://doi.org/10.1073/pnas.74.12.5463. PMCID: PMC431765. PMID: 271968.

chemistry and automated CE. However, the method remains labor intensive and relatively low throughput.

The terminology of "next generation" was introduced with the emergence of higher throughput and lower cost sequencing technologies based on parallel analysis. Several second-generation (and now third-generation) technologies have been commercialized, and each employ different sequencing techniques, such as pyrosequencing, sequencing by ligation, pH-mediated semiconductor sequencing, and reversible dye-terminator chemistry. Currently, bench-top second-generation technologies from Illumina (MiSeq), Thermo Fisher Scientific (Ion Torrent S5), and Verogen (MiSeq FGx) are the most purchased systems by forensic laboratories because of the availability of compatible human identity testing kits, appropriate scale of sequence output, and relatively low reagent costs. This is a rapidly growing field, and although few U.S. agencies use NGS technology, advancements resulting in third-generation technologies may one day find a niche in forensics. Instruments from Pacific Biosciences and Oxford Nanopore are focused on single-molecule sequencing methods that do not require DNA amplification. The Oxford Nanopore MinIon is about the size of a large USB drive, and its use has been demonstrated in the field, in the arctic, and on the International Space Station.

The overall mechanism for NGS may seem mystifying considering the numerous sequencing methods and technologies available. Platform technology considerations aside, forensic DNA sequencing uses a targeted sequencing approach with PCR amplification or hybridization capture probes (followed by PCR amplification) for the selective analysis of DNA regions of interest. This is advantageous for throughput and cost factors, relative to whole-genome sequencing. Because PCR is predominantly used for enrichment, targeted resequencing is similar in overall sensitivity to DNA input and has similar stutter artifacts as CE-based methods. PCR-based enrichment is considered more sensitive than hybridization capture probe-based methods. Hybridization capture, however, has advantages when analyzing extremely degraded and contaminated sample types. NGS requires a few extra steps beyond the CE-based analysis workflow currently used in forensic DNA labs (see Figure 1).

Figure 1. Comparison of CE and NGS Workflow Diagrams of CE and NGS



Extra steps are highlighted in the green "library prep" inset. Sequencing requires a slightly more complex procedure. Resulting data output format and analysis software are also different.

Following (or during) PCR enrichment, amplicon fragments are used to generate a "library," which refers to DNA molecules with sequences attached to make the DNA compatible with the sequencing technology being used. Libraries can be constructed by different methods (e.g., PCR or ligation) to incorporate multiple sequencing elements (e.g., target sequence, sample barcode, key sequence for calibration, and sequencing primer sites). Before sequencing, libraries may be evaluated for quality and concentration, though bead-based normalization procedures are available to simplify the workflow. During sequencing template preparation, library molecules undergo "clonal" amplification to generate a DNA cluster that originated from a single library molecule. This happens on a glass slide inside a flow cell on a MiSeq/MiSeq FGx or, with the Ion Torrent system, a microscopic "Ion Sphere Particle" (ISP); both are coated with forward and reverse oligonucleotides that are complementary to the library molecule adaptors. Once the library molecules are hybridized to the oligos on the solid substrate, they are amplified by either bridge or emulsion PCR (Illumina/Verogen and Ion Torrent methods, respectively) to generate sequencing template DNA. Interrogation of each cluster is performed in parallel for millions of clusters or ISPs, using methodology specific to the NGS technology chosen. With Illumina/Verogen's method, fluorescently labeled reversible dye-terminator nucleotides are incorporated into the sequencing strand, then imaged under laser illumination to generate four colors of fluorescence data, one for each DNA base. In Ion Torrent's S5 platform, ISPs are situated on a semiconductor chip that detects pH change when a nucleotide is incorporated into the sequencing strand and transduces it into an electrical signal interpreted by the instrument as a

base-incorporation signal, depending on which nucleotide is being introduced to the sequencing reaction at the time.

In both cases, the signal produced is ultimately decoded into a string of sequence data by the instrument, with each sequence corresponding to a single molecule from the library. The raw sequence output is stored in "fastq" format (see Figure 2), which encodes a quality estimate for each base in the string. A typical output folder from a MiSeq FGx run contains files occupying 10 to 20 GB or more of storage. Frequent instrument usage may require large-scale storage. Retention of just the fastq output may reduce storage needs, requiring about 3 GB per run. The Ion Torrent S5 instrument relies on a UNIX-based (Ubuntu OS) server through which run information is input to the instrument and output data can be managed. Files may be exported in fastq format or other formats, requiring between 2 and 10 GB of storage per run, depending on format.

Figure 2 shows a sample of one "read" from a fastq file. The first line is a unique identifier for the read, and the second line is a text string of nucleotide base calls (A, G, C, or T). The third line ("+"), and the fourth line is the quality score for each of the bases on the second line. Quality scores are encoded by their ASCII characters (ASCII code #73 encodes a perfect Q-score of 40 whereas code #33 represents the lowest Q-score of 0), with letters I through A (shown here) representing the highest quality and numbers or characters are lower quality. Quality score diminishes corresponding to the decreasing number of the ASCII code. A single fastq file may contain hundreds of thousands to billions of reads.

Figure 2. Sample of One "Read" from a Fastq File

```
@MDKNP:03689:01140
CCCGAGTAGATTACAGGCGTGCACCACCACGCCTGGCCAATTTTTGT
+
?A>ABBBBBBB@CCCB@BCBBBAA?AA=AA;;6:B>B@B59999);?
```

Image provided courtesy of Kevin Kiesler of the National Institute of Standards and Technology

Following the collection of the sequence data, specialized software is used to interpret the data into STR allele calls. These software packages may be provided by the instrument or assay manufacturer, with third-party and open-source software also available. (See Appendix B for examples of NGS analysis software with graphical user interfaces.) Determining the number of STR repeats from sequence-based data allows comparison with CE-based STR genotypes, including those stored in NDIS.

However, STR analysis at the DNA sequence level increases discrimination power through characterization of underlying sequence variations, or polymorphisms, a capability not possible with size-based CE analysis. The International Society of Forensic Genetics is currently working to provide best practice guidelines on nomenclature standardization for sequence-based STR alleles. Foundational research to determine the appropriate population-specific frequencies to

apply to sequence-based STR analysis for weight-of-evidence statistical calculations is ongoing. Efforts are being made to catalogue observed sequence variants within and adjacent to (termed "flanking SNPs") forensically relevant STR loci to allow end users to validate STR alleles and to facilitate population frequency research and nomenclature development.[3]

The Scientific Working Group on DNA Analysis Methods (SWGDAM) has produced a guidance document for laboratories seeking to implement NGS. SWGDAM's guidance document states that the number of reads present for any given locus is analogous to fluorescence intensity in CE typing. Although this is an indirect comparison, each read from an NGS instrument originates from a single molecule from the library whereas in CE, each template DNA molecule contributes a portion of the fluorescent signal in the final PCR product. One advantage of NGS is that fluorescent separation and size separation constraints seen in CE are removed; therefore, amplicons can be shortened to include the minimal STR region, or in the case of SNPs, may be less than 100 bp, which improves performance with degraded DNA. The only limitation on the number of markers typed relates to the desired sequencing coverage (which may be relevant to mixture analysis) and how many samples may be typed in a sequencing run. In practical terms, this means that 10 to 100 samples may be genotyped for 100 to 200 markers, or a few samples may be interrogated at thousands of marker sites enabling simultaneous analysis of STRs, SNPs, and lineage markers found on the Y-chromosome and mitochondrial DNA (mtDNA).

The ability to characterize multiple marker types enables new capabilities; SNPs are single-base variations that can be used to provide valuable presumptive information related to global ancestry (e.g., geographic or cultural group discernment) and phenotypic traits (e.g., eye, hair, or skin color). Because sample enrichment uses short PCR amplicons, there is increased likelihood of success when interrogating degraded DNA samples, which makes NGS well suited to missing persons, missing or killed in action remains, mass grave remains, and mass disaster victim identification. mtDNA analysis for degraded evidence is routinely achieved by Sanger sequencing, and the emergence of NGS greatly simplifies sequencing of mtDNA (including the entire mitochondrial genome) with improved sensitivity to detect heteroplasmy (a mixture of mtDNA sequence within an individual), with decreased labor input, faster analysis time, and reduced cost relative to Sanger-type sequencing. Because of the increased labor, cost, and turnaround time with NGS methods, laboratories may choose to focus only on specific types of cases to supplement mainstream CE analysis. A non-exhaustive list of cases where NGS may be advantageous over CE is found below:

- Degraded DNA sample
- One-shot, limited sample

[3] Gettings, K. B., Borsuk, L. A., Ballard, D., Bodner, M., Budowle, B., Devesse, L., King, J., Parson, W., Phillips, C., & Vallone, P. M. (2017, Nov). STRSeq: A catalog of sequence diversity at human identification Short Tandem Repeat loci. *Forensic Sci Int Genet, 31*, 111-117. https://doi.org/10.1016/j.fsigen.2017.08.017. PMCID: PMC7304526. PMID: 28888135.

- Mixtures with high levels of shared alleles
- Male minor contributor (e.g., total human:male ratio > 20:1)
- Cold case evidence or unidentified remains
- Dead-end or no suspect cases (e.g., no Combined DNA Index System [CODIS] hit)
- Tissue source identification
- Biological age estimation

Cost–Benefit Analysis

Estimated Cost of Instrumentation and Facility Requirements

As an entry point to NGS, most practicing laboratories focus on STR analysis for reverse compatibility with CE and the advantages offered by sequence variant alleles in mixture analysis and improved performance with degraded DNA. The following section will focus mainly on existing commercial NGS/STR kits. Additional kits for SNP typing for individual identification and biogeographic ancestry estimation are available from Thermo Fisher Scientific and are compatible with the S5 system. These SNP markers are included in Verogen's ForenSeq DNA Signature kit. Mitochondrial DNA assays are available from Promega, Thermo Fisher, Qiagen, and Verogen.

Instrumentation and associated equipment: The MiSeq/MiSeq FGx and Ion Torrent S5 are currently the most used instrument platforms in forensics. The MiSeq FGx can run in a mode that emulates the Illumina MiSeq and is compatible with any assay for that system. However, to use Verogen's ForenSeq DNA Signature kit, a MiSeq FGx is required for instrument control and downstream analysis with the Universal Analysis Software (UAS) and server.

The choice of platform and chemistry will dictate specific implementation requirements, including automation and analysis software needs. Performance characteristics are out of scope of this guide, and although commercially available equipment is specifically mentioned herein, in no way does this guide endorse any commercial product and no recommendation will be made as to which system is best suited to the needs of an individual laboratory.

For simplicity, the information presented here will be organized into three use cases (*Tables 1–3*) according to STR analysis chemistry used. The laboratory costs for NGS vary depending on the number of samples processed. This results from single-use sequencing reagent, which is independent of the number of samples processed; thus, the cost per sample will decrease as the number of samples increases. An example of reagent cost to analyze 96 samples is presented for comparison.

Currently available STR analysis kits (with the number of loci included):

- Promega—PowerSeq 46GY System
  - Sex determination marker (1)
  - Autosomal STRs (22)
  - Y-chromosome STRs (23)

- Thermo Fisher Scientific—GlobalFiler NGS V2 Kit
  - Sex determination markers (4)
  - Autosomal STRs (30)

- Verogen—ForenSeq DNA Signature Prep Kit
  - DNA Primer Mix A (DPMA)
    - Sex determination marker (1)
    - Autosomal STRs (27)
    - Y-chromosome STRs (24)
    - X-chromosome STRs (7)
    - Identity SNPs (94)
  - DNA Primer Mix B (DPMB)
    - Loci in DPMA
    - Phenotype SNPs (22)
    - Biogeographic Ancestry SNPs (56)

Table 1. Use Case 1: Promega PowerSeq 46GY System

Capital equipment: Illumina MiSeq ($100,000)
Analysis software: SoftGenetics GeneMarker HTS ($3,500)
Reagent cost per 96 samples:

Part #	Name	List Price (approx.)	# Samples/ Kit	Cost per 96 samples
PS4600	PowerSeq 46GY System	$4,500.00	96	$4,500.00
PS5000	PowerSeq Quant MS System	$700.00	500	$134.40
20015963	TruSeq DNA PCR Free High Throughput (Illumina)	$2,900.00	96	$2,900.00
20022370	IDT for Illumina TruSeq DNA UD Indexes (Illumina)	$700.00	96	$700.00
MS-102-3003	MiSeq Reagent Kit V3 (600 cycle) (Illumina)	$1,700.00	96	$1,700.00
			Total	**$9,934.40**

Table 2. Use Case 2: Verogen ForenSeq

Capital equipment: Verogen MiSeq FGx ($140,000)
Analysis software: UAS (included)
Reagent cost per 96 samples (using DNA Primer Mix A—STRs & Identity SNPs):

Name	List Price	No. Samples/ Kit	Cost per 96 samples
ForenSeq DNA Signature Prep Kit	≈$17,500.00	384	$4,375.00
MiSeq FGx Reagent Kit	≈$1,600.00	96	$1,600.00
		Total	**$5,975.00**

Table 3. Use Case 3: Thermo Fisher Scientific—Ion Torrent S5

Capital equipment: Ion Torrent S5 ($70,000) and Ion Chef Instrument ($60,000)
Analysis software: Converge ($10,000 for 3-year license)
Reagent cost per 96 samples (using manual library preparation):

Name	List Price	No. Samples/ Kit	Cost per 96 samples
Precision ID Library Kit (Manual)	≈$8,700.00	96	$8,700.00
GlobalFiler NGS STR Panel V2	≈$3,600.00	96	$3,600.00
Ion Xpress Barcode Adapters 1–96	≈$10,400.00	768	$1,300.00
Ion Library TaqMan Quantitation Kit	≈$1,500.00	250	$576.00
Ion 530 Chip Kit (4 pack)	≈$2,500.00	128	$1,875.00
Ion S5 Precision Id Chef & Sequencing Kit (2 runs/initialization)	≈$5,700.00	256	$2,137.50
		Total	**$18,188.50**

Table 4. Reagent Cost per 96 Samples (Using DNA Primer Mix B—Loci in DPMA + Ancestry and Phenotype SNPs)

Name	List Price	No. Samples/Kit	Cost per 96 samples
ForenSeq DNA Signature Prep Kit	≈$17,500.00	384	$4,375.00
MiSeq FGx Reagent Kit	≈$1,600.00	32	$4,800.00
		Total	**$9,175.00**

Table 5. Reagent Cost per 96 Samples (Using Automated Library Preparation on Ion Chef)

Name	List Price	No. Samples/Kit	Cost per 96 samples
Ion Ampliseq Kit for Chef DL8	≈$4,600.00	32	$13,800.00
GlobalFiler NGS STR Panel V2	≈$3,600.00	32	$10,800.00
Ion Library TaqMan Quantitation Kit	≈$1,500.00	250	$576.00
Ion 530 Chip Kit (4 pack)	≈$2,500.00	128	$1,875.00
Ion S5 Precision ID Chef & Sequencing Kit (2 runs/initialization)	≈$5,700.00	256	$2,137.50
		Total	**$29,188.50**

Common laboratory equipment:

- Bench-top centrifuge (≈$3,000)
- Bench-top vortexer (≈$400)
- Ice machine (≈$4,000)
- Plate shaker (≈$2,000)
- Magnetic stand for 96-well plate (≈$500)
- Multichannel pipettes (≈$3,000)
- Thermal cycler (≈$7,000)

Automation for library preparation: Intensive, high-throughput usage is labor intensive, and automation generally improves reproducibility. Various third-party vendors may offer instrumentation for library preparation. Systems range in price from ≈$100,000 to >$200,000 with many options. These systems are complex and will require service and support contracts:

- VERSA (Aurora Biomed Inc., Vancouver BC, Canada)
- Bravo NGS (Agilent, Santa Clara, CA, USA)
- NGS STAR (Hamilton, Reno, VN, USA)

Analysis software: To use Thermo Fisher Scientific's GlobalFiler NGS V2 kit on the Ion Torrent S5 instrument, downstream analysis of STRs will require Converge software (Thermo Fisher Scientific, Waltham MA, USA) with a Case Management and NGS Data Analysis License ($10,000 for 3-year license, server included). Promega's PowerSeq 46GY System does not include software for STR analysis. Promega recommends using GeneMarker HTS (Softgenetics, $3,500) for genotyping. Verogen's MiSeq FGx includes UAS plus the server to run UAS. Alternatives exist for STR genotyping, including the open-source software STRait Razor[4] (University of North Texas, Fort Worth TX, USA) and Exact ID 2.0 (Battelle, Columbus

[4] Woerner, A. E., King, J. L., & Budowle, B. (2017, Sep). Fast STR allele identification with STRait Razor 3.0. *Forensic Sci Int Genet, 30*, 18-23. https://doi.org/10.1016/j.fsigen.2017.05.008. PMID: 28605651.

OH, USA), among others. Open-source software may present unique challenges to implementation. Agency information technology policy may preclude the use of open-source software and may present challenges to updates to operating systems and software updates.

Facility considerations: Separate pre-PCR and post-PCR areas are recommended (e.g., post-PCR co-located with PCR thermal cyclers or CE instruments) to keep amplified product separate from amplification reagents, reducing contamination risk. If possible, the sequencing instrument should be physically separated from the PCR set-up or library construction locations to prevent contamination of sample processing areas with amplified libraries from the sequencing instrument. Instrument placement should fit the laboratory workflow; available counter space should be assessed to determine whether the instrument fits with adequate access for instrument function or maintenance, such as door clearance, adequate airflow, or removal of outer coverings. Availability of electrical supply, access to laboratory grade water, and liquid waste disposal options (e.g., formamide or other environmentally hazardous waste) must also be considered. Additional airflow is typically not required because normal ventilation should be enough for the total thermal output of the instrument. Pre-installation checklists are available upon request from vendors and can differ marginally based on NGS technology (vendor specific). Additional considerations are as follows:

- Network connection may be required to operate instrumentation:
  - Gigabit connection between the NGS instrument and data management system (e.g., computer). Connection can be made directly or through a network switch.
  - Turn off the automatic operating system updates and virus software scans on the computer so as to not interrupt a run. Manually run these updates; vendor recommendation suggests waiting at least 1 month after a new operating system release before performing an update.

- Bench space:
  - Instrument: ≈75 cm^3 (30.0 in^3),
  - Liquid handling robot: ≈75 to 150 cm^3 (30.0 to 60 in^3)
    - Vibration reduction: immobilized lab bench without castors; remove shaker, vortexer, centrifuge, heavy fans, and other potential vibrational equipment from the same bench-top
    - Electrical: 110/120 VAC (220/240 VAC); 9 A (max); 50/60 Hz; 400 W
    - Uninterrupted Power Supply: APC Back-UPS Pro #BR1500G or other similar uninterrupted power supply
    - Laboratory temperature: 19°C–25°C
    - Laboratory humidity: 30%–75%

Maintenance costs: After the initial purchase of the instrument, reagents, and other needs associated with the validation of NGS technology, future budgets should plan for ongoing costs to use this technology. Continued costs can include vendor-led staff training sessions, proficiency testing, validation of any new or emerging reagent kits, analysis software upgrades or additional licenses, routine instrument preventative maintenance and repair servicing through the vendor, replacement parts not covered by service agreements, and additional external hard drives or data storage solutions.

Reagents and consumables: Based on the scope of internal validation, enough reagents should be identified and on hand to include specific evaluations and sub-studies based on the desired genetic markers for instrument calibration and control data. NGS reagents and kits will vary depending on the instrumentation being used. These reagents may be included in the scope and cost of the contract if an external vendor will be performing the internal validation. The number of reagents and kits procured should be in-line with expected throughput to avoid materials expiring.

- Consumables (e.g., pipette tips, 96-well PCR plates, gloves, alcohol cleaning solution, lab tissue, lens paper, microcentrifuge tubes, sodium hypochlorite, Tween 20, tweezers, nuclease-free water, disposable gloves): ≈$6,000/year
- Library and sequencing reagent kits: cost varies

Digital data storage: Instrument platforms can store several runs worth of data before hard drive space must be cleared for more files. Higher throughput usage will necessitate additional storage infrastructure. There are several different options available. The appropriate scale and mechanism for each laboratory should be researched early in the implementation planning process. Resulting data can be several hundred MB or larger, so multiple TB storage is recommended for long term data storage. Automated backup for off-site data copies offer protection against fire, flood, or other disaster.

Storage options are external hard drives, cloud storage, and network attached storage (NAS). External hard drives can be vulnerable to damage or misplacement and do not allow for remote access and network storage. If an external hard drive approach is desired, a secondary duplicate drive should be considered for data backup (e.g., computer and external hard drive, two external hard drives, or network- or cloud-based storage and external hard drive). Alternatively, cloud and NAS allow for remote access. However, these storage options may be incompatible with agency policy and can also have potential vulnerabilities and threats from undesired hackers. Cloud services are rented off-site storage from a provider, whereas NAS allows for remote use on a networked local device that could reside within the laboratory or building. Established agency servers should be considered because these are often routinely monitored by the agency for potential threats and vulnerabilities.

- External hard drive for data storage: ≈$1,000 (24 TB)
- NAS storage: ≈$1,000 (24 TB)

- Vendor-specific options:
  - Torrent Storage NAS Device: 48 TB (≈$39,000) or 110 TB (≈$58,000)
  - Thermo Fisher Cloud: free (1 TB storage per instrument)

Personnel Requirements

Appropriate personnel should be assigned to this project. The number and time commitment of assigned personnel depend on implementation timeline and planned usage. At minimum, two dedicated forensics scientists should be trained and available for NGS operation and data analysis (to accommodate vacations schedules, illnesses, increased caseloads, or turnover). Although NGS operation involves common molecular biological techniques (e.g., PCR, bead-based cleanup, quantitative PCR, instrument maintenance), staff should possess qualifications recommended by the Federal Bureau of Investigation (FBI) Quality Assurance Standards (for accredited labs, a bachelor's degree), with additional recent training or coursework in modern genomics methods highly recommended. Continuing education should be focused on maintaining expertise in NGS (e.g., theory, troubleshooting, and advanced data analysis) for legal testimony.

The laboratory's Technical Leader should be available to oversee validation, training, competency assessments, maintenance, control data, troubleshooting, and data analysis. External vendors may offer validation services.

Training: Specialized training on hardware, software, and routine maintenance is necessary and is vendor specific. Therefore, training and technical support from the vendor will be required during the initial NGS implementation phase. Although some vendors may provide free online training opportunities, on-site vendor training before initial use is highly recommended. Vendor-based training can last between 1 and 2 days, with options available for both instrument and bioinformatics/data analysis training. Using commercial software solutions provides validated analysis and interpretation capability with ease of use. Open-source software is available for alternative analyses but only with prerequisite skill in bioinformatics.

A good approach for operational training includes the following steps: read to understand, observe the process, perform the process under supervision, then perform the process independently.[5] Prior to initial implementation or new user hands-on training, personnel should review NGS user manuals for operation and analysis and any additional standard operating procedures provided by the laboratory. Subsequent hands-on training steps can be performed with a vendor upon initial purchase, and it may be helpful to have multiple staff (at least two and up to the maximum allowed by the vendor) participate in the vendor-led on-site training.

[5] Hutchins, R. J., Phan, K. L., Saboor, A., Miller, J. D., & Muehlenbachs, A. (2019, Aug). Practical Guidance to Implementing Quality Management Systems in Public Health Laboratories Performing Next-Generation Sequencing: Personnel, Equipment, and Process Management (Phase 1). *J Clin Microbiol, 57*(8). https://doi.org/10.1128/jcm.00261-19. PMCID: PMC6663893. PMID: 31142608.

Protocol simulation tools are available on the Forensic Technology Center of Excellence (FTCoE) website (https://forensiccoe.org/workshop/mps-workflow-through-simulation/).

Vendors should provide routine maintenance recommendations during training for daily, weekly, and annual or semi-annual maintenance. Tasks may include cleaning, routinely washing the instrument, replacing wash solutions, and replacing user serviceable parts. It is recommended to keep a maintenance log and schedule preventative maintenance service through the vendor yearly. Example logs and procedures that may be helpful can be found on the Centers for Disease Control and Prevention (CDC) website.[6]

Potential Funding Sources

Funding acquisition may be achieved via numerous grant opportunities and local/state funding; for example, NGS testing is currently an allowable expense under the DNA Capacity Enhancement and Backlog Reduction (CEBR) grant program. Thought may be given to setting up "Centers of Excellence," where multiple jurisdictions pool funds and resources to provide this instrumental platform to several laboratories.

Pros and Cons of NGS

- **Advantages**
  - SWGDAM and other standards are in place.[7,8]
  - ForenSeq™ DNA Signature Prep Kit and MiSeq FGx (Verogen, San Diego, CA, USA) have received NDIS approval, allowing NGS-generated STR profiles to be searched at NDIS[9]
  - NGS shows promise for improved mixture deconvolution with sequence-based alleles.
    - National Institute of Standards and Technology (NIST) reported a substantial increase in unique alleles detected with NGS compared with length-based CE technology.[10]

[6] Centers for Disease Control and Prevention. (n.d.). *QMS tools and resources. The Next Generation Sequencing Quality Initiative.* Centers for Disease Control and Prevention. Retrieved August 27, 2020, from https://www.cdc.gov/labquality/qms-tools-and-resources.html

[7] Scientific Working Group DNA Analysis Methods. (2019). Addendum to the SWGDAM interpretation guidelines for autosomal STR typing by forensic DNA testing laboratories to address next generation sequencing. https://1ecb9588-ea6f-4feb-971a-73265dbf079c.filesusr.com/ugd/4344b0_91f2b89538844575a9f51867def7be85.pdf.

[8] Scientific Working Group DNA Analysis Methods. (2020). The guidance document for the FBI quality assurance standards for forensic DNA testing and DNA databasing laboratories, Effective 07/01/2020. https://1ecb9588-ea6f-4feb-971a-73265dbf079c.filesusr.com/ugd/4344b0_2bce9398b6a640fdb626063469939151.pdf

[9] FBI Laboratory. (2021). *National DNA Index System (NDIS) operational procedures manual. Version 10. Effective 07/01/2021*. https://www.fbi.gov/file-repository/ndis-operational-procedures-manual.pdf

[10] Steffen, C. R., Gettings, K. B., Kiesler, K. M., Borsuk, L. A., & Vallone, P. M. (2019, 2019/12/01). Sequence variation observed in 27 Y-STR markers with U.S. population samples. *Forensic Science International: Genetics Supplement Series, 7*(1), 520-521. https://doi.org/https://doi.org/10.1016/j.fsigss.2019.10.074.

  - Increased multiplexing in a single assay offers more types of information about a sample.
  - NGS can be used to analyze samples that are too degraded for CE/STR analysis.
  - NGS can be used to examine evidence from cold cases or post-conviction cases.
  - NGS can be used to determine the number of contributors to a DNA mixture better.
  - NGS provides an open platform for development of new markers

- **Disadvantages**

  - The technology is currently early in adoption in forensics.
  - NGS may face legal challenges as it is adopted.
  - There may be a potential need for additional personnel resources.
  - Limited automation solutions are currently available.
  - NGS data are storage intensive.
  - The high cost per sample.
  - Probabilistic genotyping solutions for mixture interpretation are not fully developed.
  - CODIS can only be searched by length-based alleles.

    - This reduces the advantage of sequence information.

  - NGS data artifacts need to be studied to determine their effect on estimating the number of contributors to a mixture.
  - NGS is only one-half of the forensic genetic genealogy equation; only a handful of available genealogical DNA databases.

Implications for Current Casework

Timeline: The timeline for implementing NGS can be expected to vary widely. Time savings can be realized through using an external vendor to perform the internal validation, avoiding increased workload on existing personnel. As an estimate, the implementation can could be accomplished in less than one year using an external vendor; an estimate of one to two years can be anticipated if the validation is done internally. The use of external resources will still require personnel assigned to NGS, and work to develop standard operating procedures (SOPs) and laboratory parameters can be carried out concurrently. Example SOPs that may be helpful can be found on the CDC website.[11]

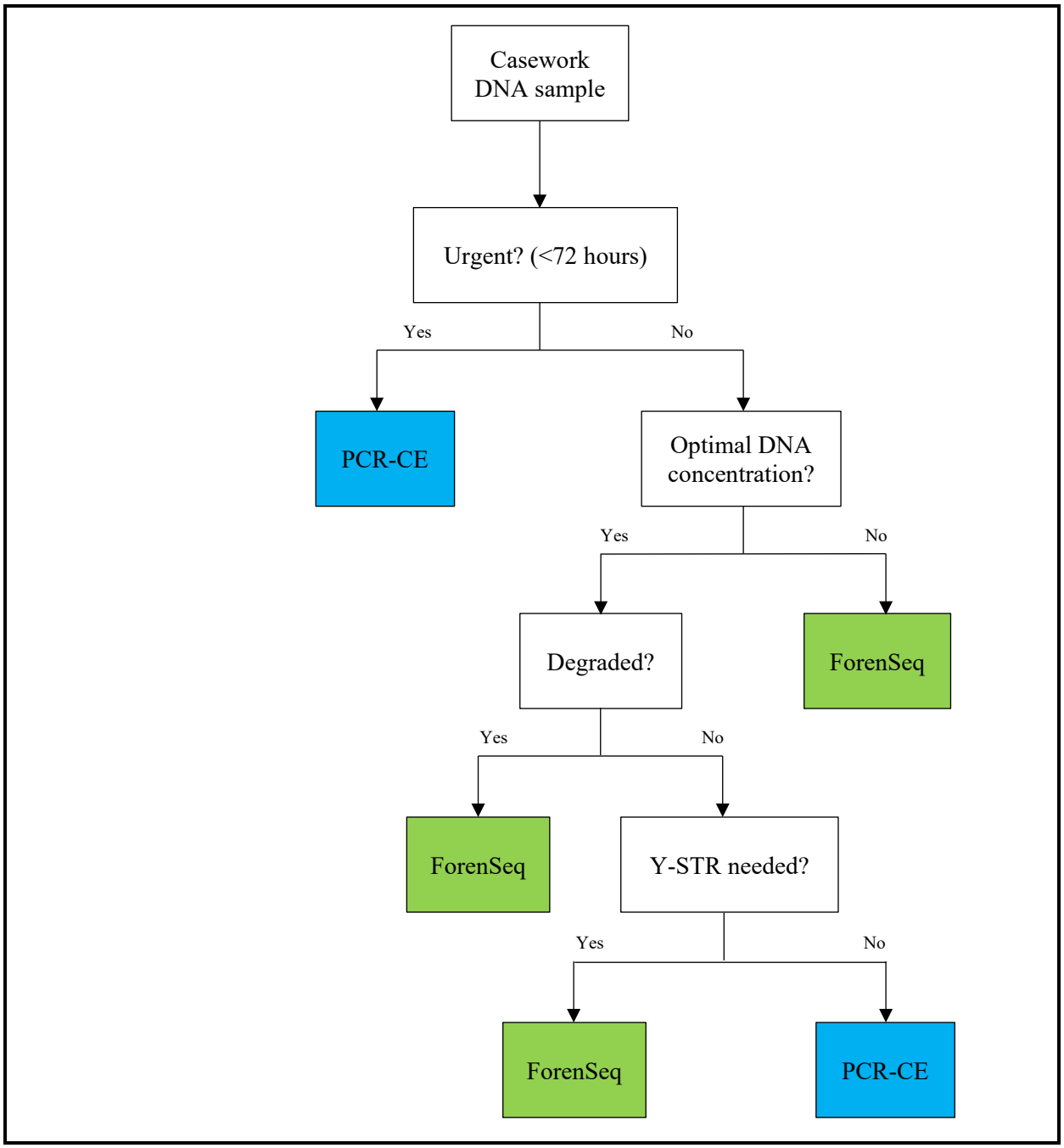
Determine the scope of the project: A flow chart may be helpful to anticipate where NGS technology will be inserted into the case flow (identity testing vs. investigative intelligence).[12]

[11] Centers for Disease Control and Prevention. (n.d.). *QMS tools and resources. The Next Generation Sequencing Quality Initiative.* Centers for Disease Control and Prevention. Retrieved August 27, 2020, from https://www.cdc.gov/labquality/qms-tools-and-resources.html

[12]Laurent, F.-X. (n.d.). *Routine use of ForenSeq® solution on casework samples: feedback one year after lab implementation.* Institut National de Police Scientifique. Retrieved July 9, 2021, from https://verogenevents.com/wp-content/uploads/2018/11/App-Spot-I_Francois-LAURENT.pdf

See Figure 3 for an example flow chart used by the Institut National de Police Scientifique on implementing NGS into casework and triaging samples. Once use cases are determined, further considerations impacting the implementation timeline can be addressed. This should include integration with the laboratory's Laboratory Information Management System (LIMS), training of existing personnel, expected NGS caseload, and additional personnel resources needed (e.g. automation specialists or bioinformaticists). Depending on the size of the laboratory and volume of cases undergoing NGS, it is not likely that all analysts will need to be proficient to keep up with caseload demands. However, all analysts should be aware of the capabilities and limitations of NGS to properly identify and route eligible cases for analysis.

Figure 3. Flowchart for Triaging Samples Suitable for NGS



Establishing a budget for NGS: The startup cost will be different than the budget needed to sustain NGS use. Training by the vendor should be included with the initial purchase of a new system, but the costs of on-site training and integrating the new workflow into an existing LIMS

system should be included in the budget. Updates the LIMS configuration may need to support new NGS workflows include the following:[13]

- Set-up of analysis and batch templates for library preparation and library quantitation,
- Development of files to electronically export massively parallel sequencing sample set-ups and analysis to instrument software,
- Configuration of scripts for the electronic transfer of STR allele and sequence data files generated by the analysis software into the LIMS system, and
- Development of DNA data export files to upload into national DNA databases.

Communicating Return on Investment to Stakeholders

Early implementers will benefit from a formal user's group to communicate return on investment, best practices, challenges, and tips,. As more laboratories work to implement NGS, such peer resources will become more readily available.

Topics to cover when communicating return on investment may include the following:

- What NGS can do (and cannot do)
- Compatibility with existing databases and legacy data
- Benefit of added information relative to cost
- Ability to validate off-ladder, variant, and rare STR alleles from CE
- How attorneys can explain NGS in trial: technology, procedure, statistics
- Dissemination of information via webinars and workshops

Considerations of a Validation Plan

SWGDAM has addressed internal validation of NGS platforms and kits. SWGDAM Validation Guidelines for DNA Analysis Methods were revised to address NGS technologies and were approved by the SWGDAM Executive Board in December 2016. These recommendations can be referenced for the validation parameters/procedures consistent with the FBI's Quality Assurance Standards for Forensic DNA Testing Laboratories, found here:

https://www.fbi.gov/file-repository/qas-audit-for-forensic-dna-testing-laboratories.pdf/

Information derived from the internal validation studies will be used to inform overall data expectations to derive standard operation procedures and interpretation guidelines.

[13] Barrio, P. A., Martín, P., & Alonso, A. (2017, 09/12). LIMS Configuration to fit new Massively Parallel Sequencing Workflows in Forensic Genetics. *Forensic Science International: Genetics Supplement Series, 6*, e104-e106. https://doi.org/10.1016/j.fsigss.2017.09.040

Resources

Laboratories that have implemented NGS technology include the following:

- California Department of Justice
- Armed Forces DNA Identification Laboratory
- Netherlands Forensic Institute
- Institut für Gerichtliche Medizin (Institute for Legal Medicine)
- Institut National de Police Scientifique
- Defense Forensic Science Center
- Ohio Bureau of Criminal Investigation
- FBI
- DC Department of Forensic Sciences

Academic and research institutions that are considering or have implemented the technology include the following:

- University of North Texas Health Science Center
- Sam Houston State University
- Syracuse University
- Kings College London (also using NGS in investigation)
- Penn State University (Dr. Mitch Holland)
- Florida International University
- Rutgers (Dr. Catherine Grgicak)
- NIST (Applied Genetics Group)

Some laboratories may share SOPs, training manuals, and validation plans and summaries on their websites. Additionally, resources can be found at the FTCoE website, including guidance (https://forensiccoe.org/report-massively-parallel-sequencing/) and protocol simulation tools (https://forensiccoe.org/workshop/mps-workflow-through-simulation/).

Further Reading

Ballard, D., Winkler-Galicki, J., & Wesoły, J. (2020). Massive parallel sequencing in forensics: advantages, issues, technicalities, and prospects. *International journal of legal medicine, 134*(4), 1291-1303. https://doi.org/10.1007/s00414-020-02294-0. PMID: 32451905.

Battelle Memorial Institute. (2019, February 9). Next generation sequencing (NGS) feasibility and guidance study for forensic DNA. Washington, DC: National Criminal Justice Reference Service. https://www.ncjrs.gov/pdffiles1/nij/grants/252287.pdf

Børsting, C., & Morling, N. (2015, Sep). Next generation sequencing and its applications in forensic genetics. *Forensic Sci Int Genet, 18*, 78-89. https://doi.org/10.1016/j.fsigen.2015.02.002. PMID: 25704953.

Jäger, A. C., Alvarez, M. L., Davis, C. P., Guzmán, E., Han, Y., Way, L., Walichiewicz, P., Silva, D., Pham, N., Caves, G., Bruand, J., Schlesinger, F., Pond, S. J. K., Varlaro, J., Stephens, K. M., & Holt, C. L. (2017, 2017/05/01/). Developmental validation of the MiSeq FGx Forensic Genomics System for Targeted Next Generation Sequencing in Forensic DNA Casework and Database Laboratories. *Forensic Science International: Genetics, 28*, 52-70. https://doi.org/https://doi.org/10.1016/j.fsigen.2017.01.011.

Scientific Working Group on DNA Analysis Methods (SWGDAM). (n.d.). Addendum to the SWGDAM Interpretation Guidelines of Autosomal STR Typing by Forensic DNA Testing Laboratories to Address Next Generation Sequencing. https://www.swgdam.org/publications

Acknowledgments

We recognize the contributions from FLN-TWG members and many authors who contributed to this white paper and from Kevin Kiesler of the Applied Genetics Group at the National Institute of Standards and Technology (NIST) for review and revision.

Suggested Citation

NIJ Forensic Laboratory Needs Technology Working Group (FLN-TWG). (2021, September). *Implementation Strategies: Next Generation Sequencing for DNA Analysis*. Forensic Technology Center of Excellence. U.S. Department of Justice, National Institute of Justice, Office of Investigative and Forensic Sciences.

Appendix A. Marker Types and Applications

Type	Applications	Comments
Autosomal STR	▪ One-to-one matching ▪ Forensic casework	▪ Workhorse of forensic laboratory DNA testing. ▪ Twenty core STR markers are present in NDIS. ▪ High matching power (e.g., >1 in 1,000,000,000,000). ▪ Has potential for more information in mixed DNA profiles. ▪ Certain STR markers can be more discriminating than others when typed by sequencing.
Y chromosome STR	▪ Missing persons investigations ▪ Paternal lineage ▪ Ancestry ▪ Kinship ▪ Forensic casework ▪ Sexual assault cases	▪ Paternally inherited genetic information. ▪ Only men possess a Y chromosome. ▪ Relatively low matching power (e.g., 1 in 1,000). ▪ Useful in sexual assault cases. ▪ Approximately 30% of forensic labs perform this analysis.[14] ▪ May help determine the number of contributors to a DNA mixture. ▪ Sequencing may provide more discrimination but not to the same level as autosomal STRs.
X chromosome STR	▪ Complex kinship ▪ Lineage	▪ Women inherit from mother and father (XX). ▪ Men inherit from mother (XY). ▪ Matching statistics consider linkage groups on the X chromosome. ▪ Rarely implemented by current CE-based methods.
Mitochondrial genome sequence	▪ Missing persons investigations ▪ Maternal lineage ▪ Ancestry ▪ Kinship ▪ Degraded samples	▪ Maternally inherited genetic information. ▪ Multiple (e.g., 100) copies of the mitochondrial genome are present in a cell making it useful for trace quantity and environmentally compromised samples. ▪ Relatively low matching power (e.g., 1 in 1,000). ▪ Few forensic labs perform this analysis.[14]

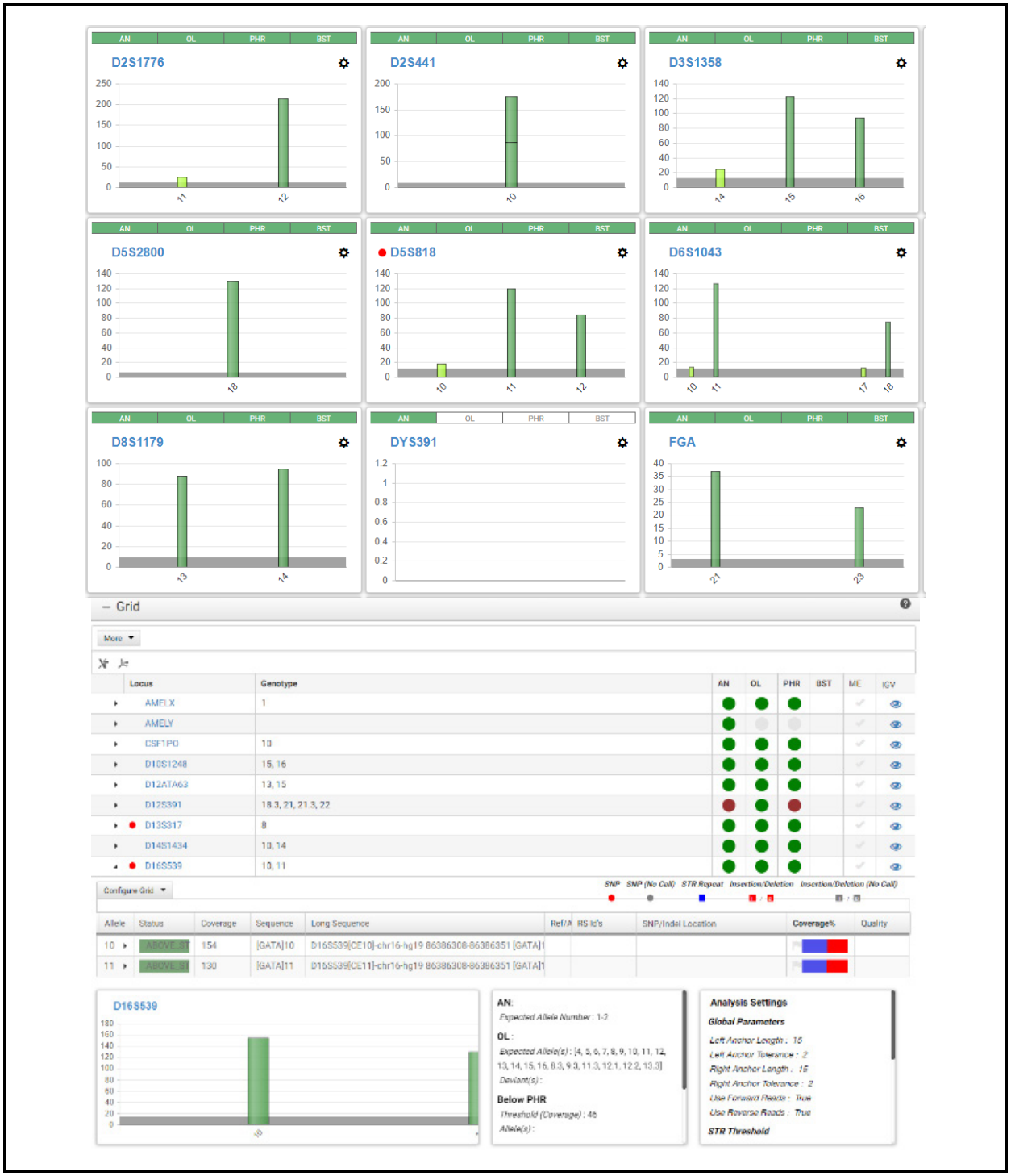
[14] Durose, M. R., Burch, A. M., Walsh, K. A., & Tiry, E. (2016, Nov). *Publicly Funded Forensic Crime Laboratories: Resources and Services, 2014.* U.S. Department of Justice, Office of Justice Programs, Bureau of Justice Statistics. https://bjs.ojp.gov/library/publications/publicly-funded-forensic-crime-laboratories-resources-and-services-2014

Type	Applications	Comments
		▪ NGS provides a high-throughput method to characterize the entire mitochondrial genome. ▪ International database, European DNA Profiling (EDNAP) mtDNA Population Database (EMPOP), can be used to identify populations with similar haplotype.
Identity SNPs	▪ Degraded samples ▪ Forensic casework ▪ Potential mixtures (microhaplotype SNPs)	▪ Generally, less informative than autosomal STRs on a per marker basis. ▪ Shorter PCR products are compatible with highly degraded samples (compared with STRs). ▪ A greater number of SNPs is required to equal the matching power of 20 autosomal STRs (e.g., >40). ▪ Moderate to high matching power. ▪ Not prone to "stutter" artifacts present in STR markers. ▪ Currently, there is no standardized core SNP marker panel in the United States (for submission to NDIS). ▪ Microhaplotypes are closely located SNPs in the genome that increase the number of observed alleles (vs. a single binary SNP).
Ancestry SNPs	▪ Biogeographical ancestry (e.g., Caucasian)	▪ Markers were selected for estimating global ancestry. ▪ Admixed populations (e.g., U.S. Hispanic) may confound ancestry interpretation. ▪ Commercial and open-source tools exist for interpreting and contextualizing results.
Phenotype SNPs	▪ Eye, skin, and hair color	▪ More complex traits like height and facial structure are typically not estimated with these commercial off-the-shelf tests. ▪ Other predictions are on the horizon (e.g., age estimation).

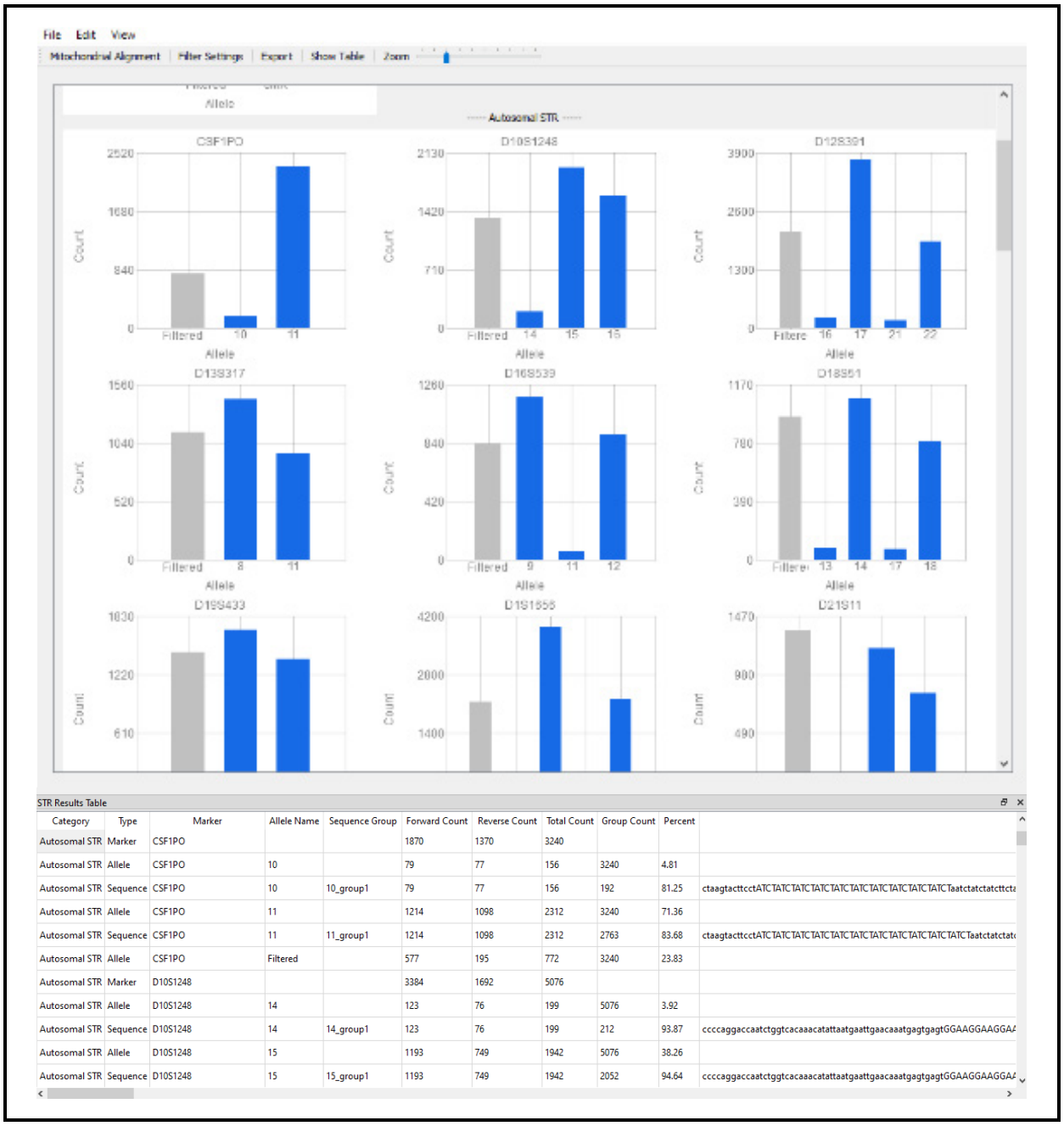
Appendix B. Examples of STR Analysis Software Output

(A) Converge from Thermo Fisher Scientific, (B) GeneMarker HTS from SoftGenetics, and (C) UAS from Verogen. These programs show histograms of the STRs and tables (with various formats) showing STR sequence and CE-equivalent allele calls.

A.

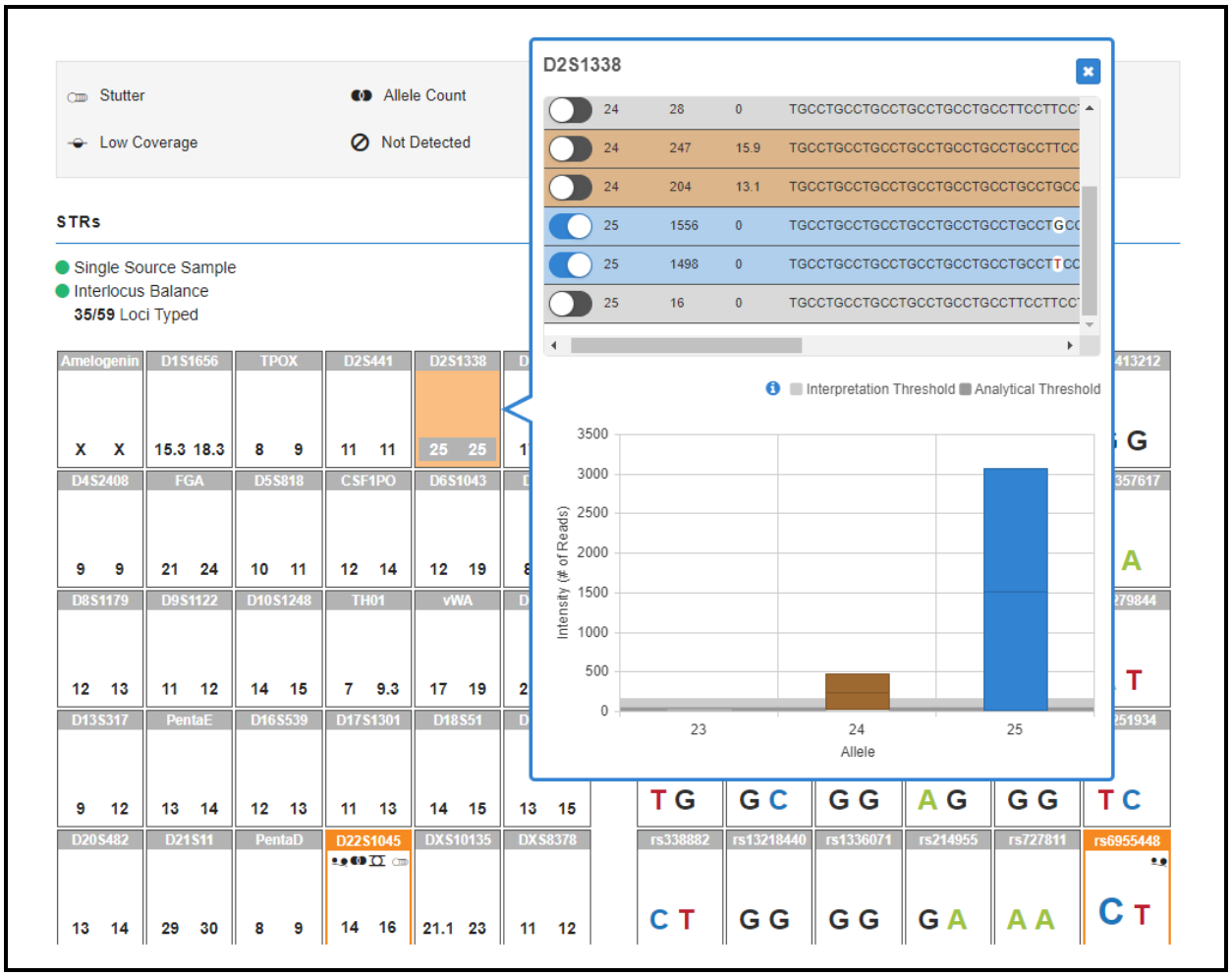


B.



Category	Type	Marker	Allele Name	Sequence Group	Forward Count	Reverse Count	Total Count	Group Count	Percent	
Autosomal STR	Marker	CSF1PO			1870	1370	3240			
Autosomal STR	Allele	CSF1PO	10		79	77	156	3240	4.81	
Autosomal STR	Sequence	CSF1PO	10	10_group1	79	77	156	192	81.25	ctaagtacttcctATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTaatctatctatcttcta
Autosomal STR	Allele	CSF1PO	11		1214	1098	2312	3240	71.36	
Autosomal STR	Sequence	CSF1PO	11	11_group1	1214	1098	2312	2763	83.68	ctaagtacttcctATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTaatctatctatc
Autosomal STR	Allele	CSF1PO	Filtered		577	195	772	3240	23.83	
Autosomal STR	Marker	D10S1248			3384	1692	5076			
Autosomal STR	Allele	D10S1248	14		123	76	199	5076	3.92	
Autosomal STR	Sequence	D10S1248	14	14_group1	123	76	199	212	93.87	ccccaggaccaatctggtcacaaacatattaatgaattgaacaaatgagtgagtGGAAGGAAGGAA
Autosomal STR	Allele	D10S1248	15		1193	749	1942	5076	38.26	
Autosomal STR	Sequence	D10S1248	15	15_group1	1193	749	1942	2052	94.64	ccccaggaccaatctggtcacaaacatattaatgaattgaacaaatgagtgagtGGAAGGAAGGAA

C.


Stutter
Allele Count
Low Coverage
Not Detected
STRs
Single Source Sample
Interlocus Balance
35/59 Loci Typed
D2S1338
24 28 0 TGCCTGCCTGCCTGCCTGCCTGCCTTCCTTCC
24 247 15.9 TGCCTGCCTGCCTGCCTGCCTGCCTTCC
24 204 13.1 TGCCTGCCTGCCTGCCTGCCTGCCTGCC
25 1556 0 TGCCTGCCTGCCTGCCTGCCTGCCTGCC
25 1498 0 TGCCTGCCTGCCTGCCTGCCTGCCTTCC
25 16 0 TGCCTGCCTGCCTGCCTGCCTTCCTTCC
Interpretation Threshold
Analytical Threshold
Intensity (# of Reads)
3500
3000
2500
2000
1500
1000
500
0
23
24
25
Allele
Amelogenin X X
D1S1656 15.3 18.3
TPOX 8 9
D2S441 11 11
D2S1338 25 25
D4S2408 9 9
FGA 21 24
D5S818 10 11
CSF1PO 12 14
D6S1043 12 19
D8S1179 12 13
D9S1122 11 12
D10S1248 14 15
TH01 7 9.3
vWA 17 19
D13S317 9 12
PentaE 13 14
D16S539 12 13
D17S1301 11 13
D18S51 14 15
D20S482 13 14
D21S11 29 30
PentaD 8 9
D22S1045 14 16
DXS10135 21.1 23
DXS8378 11 12
rs338882 C T
rs13218440 G G
rs1336071 G G
rs214955 G A
rs727811 A A
rs6955448 C T

Appendix C. Links to NGS instrument information

Thermo Fisher S5 (NGS platform)
https://www.thermofisher.com/order/catalog/product/A27212

Thermo Fisher S5xl (includes computer analysis system)
https://www.thermofisher.com/order/catalog/product/A27214

Precision ID Ancestry Panel (SNPs)
https://www.thermofisher.com/order/catalog/product/A25642

Precision ID Identity Panel (SNPs)
https://www.thermofisher.com/order/catalog/product/A25643

Precision ID GlobalFiler NGS STR Panel v2
https://www.thermofisher.com/order/catalog/product/A33114?SID=srch-srp-A33114

Precision ID mtDNA Whole Genome Panel
https://www.thermofisher.com/order/catalog/product/A30938

Precision ID mtDNA Control Region Panel
https://www.thermofisher.com/order/catalog/product/A31443?SID=srch-hj-A31443

Illumina MiSeq FGx (NGS platform)
https://www.illumina.com/systems/sequencing-platforms/miseq-fgx.html

Verogen products (ForenSeq Signature Prep Kit, Mitochondrial DNA sequencing)
https://verogen.com/products/

Promega PowerSeq 46GY (Autosomal and Y STR)
https://www.promega.com/Products/Genetic-Identity/mps-workflow/target-amplification-and-library-prep/powerseq-46gy-system/?fq=powerseq%2046gy

PowerSeq CRM Nested System, Custom
https://www.promega.com/products/genetic-identity/mps-workflow/target-amplification-and-library-prep/powerseq-crm-nested-system-custom/?catNum=AX5810

Forensic Technology
CENTER OF EXCELLENCE
A program of the National Institute of Justice

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

NIJ is dedicated to improving knowledge and understanding of crime and justice issues through science. NIJ provides objective and independent knowledge and tools to inform the decision-making of the criminal and juvenile justice communities to reduce crime and advance justice, particularly at the state and local levels. The NIJ Office of Investigative and Forensic Sciences (OIFS) is the federal government's lead agency for forensic science research and development. OIFS's mission is to improve the quality and practice of forensic science through innovative solutions that support research and development, testing and evaluation, technology, information exchange, and the development of training resources for the criminal justice community.