ForenSeq® MainstAY Product Line

Affordable library preps for targeted sequencing of established autosomal and Y-STR markers in a single reaction

Highlights

- *Largest combination of STRs in one amplification (including SE33)*
  Meets European/SWGDAM minimal Y-haplotype, CODIS, Interpol and ESS requirements.
- *Full profiles from 62.5 pg input gDNA*
  Simultaneously sequence 96 samples per run.
- *Superior balance and high call rates*
  Detect alleles from challenging samples.

Introduction

STR (short tandem repeat) loci consisting of a highly variable number of repeats among individuals are an effective tool for human identification. The most common types of STRs used in a forensic context include autosomal STRs (Au-STRs) and Y-STRs. Au-STRs provide data on combined ancestry from both parents, while the strict paternal inheritance of STRs on the Y chromosome makes them useful for paternity and kinship studies.

Methods that leverage differences in the length and number of copies of polymorphic STRs, such as capillary electrophoresis (CE), are limited by the size range in each dye channel. This limitation severely restricts the number of STRs that can be analyzed in a single reaction, forcing labs to split precious probative samples across separate STR workflows. This increases the consumption of samples and the overall cost of labor and reagents (1, 2, 3).

Next-generation sequencing (NGS) supports the simultaneous typing of various categories of STRs in a single reaction, enabling the recovery of the maximum information from a forensic sample. NGS also combines the discriminatory power of STRs with additional sequence-level variation, which can resolve isoalleles within an STR (4, 5).

The ForenSeq MainstAY product line, in conjunction with the MiSeq FGx® Reagent Micro Kit on the MiSeq FGx Sequencing System, generates Au-STRs and Y-STR profiles with a single amplification and sequencing run. You can choose between two geographically relevant options. The NDIS-approved ForenSeq MainstAY Kit is purpose-built to support a majority of national database STR uploads. The ForenSeq MainstAY SE Kit targets the same autosomal and Y-STRs while also including the highly polymorphic SE33 marker.

By focusing on established STRs used in routine human identification applications and maintaining compatibility with global databases, the ForenSeq MainstAY product line eliminates the need to choose between different kits by combining multiple workflows. The familiar and proven ForenSeq workflow generates more comprehensive DNA profiles for up to 96 evidence samples with less than 2 hours of hands-on time. The ForenSeq MainstAY product line is the most accessible way to generate highly discriminating data for mainstream casework and forensic investigative genetic genealogy confirmatory testing. Combining two common but disparate forensic workflows into a single streamlined workflow improves a laboratory's efficiency and reduces total cost of their operations.

The MainstAY Analysis Module on the Universal Analysis Software (UAS) enables guided exploration, rich visualization with project and sample views and meticulous reviews of STR allele calls with extensive filtering and sorting capabilities. It also generates reports that are easy to read and can be exported in multiple formats. This integrated workflow provides a cost-effective entry point to laboratories considering NGS for forensic applications.

Jointly amplify the largest collection of established STRs with a familiar workflow

To support locus requirements for many national and international DNA databases, the ForenSeq MainstAY product line contains 53 standard loci (54 with SE33) for use in forensic analysis, relationship testing and research. The selected STRs are broadly accepted as informative by the forensic community and by globally recognized databases, thereby reducing the burden of analyzing and validating supplementary markers. The combination of 27 Au-STRs (28 with SE33), 25 Y-STRs, and Amelogenin in a single reaction provides a higher expected likelihood ratio than most commercially available autosomal STR kits. Direct comparisons of single-source profiles are possible (e.g., in database searches and international data sharing) regardless of the assay used to type the comparative sample. The genetic profile generated by ForenSeq MainstAY enables casework searches with Au-STRs as well as haplotype searches with a Y-STR database. The Y-STRs produce a haplotype profile from male DNA which makes them extremely useful when working with male/female mixtures. In the presence of only one male contributor, it can be used to exclude potential male contributors (6, 7, 8) (Table 1).

Table 1. Markers in the ForenSeq MainstAY product line

Autosomal STRs		Y-STRs	
D1S1656	vWA	DYF387S1	DYS439
TPOX	D12S391	DYS19	DYS448
D2S441	D13S317	DYS385a-b	DYS460
D2S1338	PentaE	DYS389I	DYS481
D3S1358	D16S539	DYS389II	DYS505
D4S2408	D17S1301	DYS390	DYS522
FGA	D18S51	DYS391	DYS533
D5S818	D19S433	DYS392	DYS549
CSF1PO	D20S482	DYS393	DYS570
D6S1043	D21S11	DYS437	DYS576
D7S820	PentaD	DYS438	DYS612
D8S1179	D22S1045	DYS635	Y-GATA-H4
D9S1122	TH01	DYS643	
D10S1248	SE33*		
Amelogenin			

* DPMD only for ForenSeq MainstAY SE.

The ForenSeq MainstAY workflow is built on the familiar chemistry that underpins the ForenSeq library prep portfolio. This sensitive, PCR-based assay efficiently amplifies 53 (or 54) markers with integrated unique dual indices (UDIs), generating short amplicons that increase the likelihood of detecting alleles from degraded DNA. A DNA extract input volume of 8 µl supports flexibility for low concentration samples and dilution of inhibited samples.

The workflow includes five safe stopping points and a pre-mixed adapter plate that increases library preparation efficiency and ease. The workflow is compatible with a variety of common sources of forensic DNA, such as extracted gDNA, crude lysates and storage media card punches, such as FTA®.

The products are fully kitted solutions that include a positive amplification control for easy purchasing and calibration. The kits support sequencing and analysis of up to 96 samples on the MiSeq FGx Reagent Micro Kit using the long paired-end read capability of the MiSeq FGx Sequencing System. This maximizes sample throughput and minimizes cost per sample. A low gDNA input recommendation of 1 ng delivers reliable and reproducible recovery of full profiles from high-quality single-source samples all the way down to 62.5 pg for low input and difficult samples. Table 2 provides a complete list of product line specifications.

Table 2. Specifications for the ForenSeq MainstAY product line

Specification	Value
Sample type	Extracted gDNA from bones, blood and teeth, crude cellular lysates (e.g., from buccal swabs, blood), punches from storage cards (FTA)
Recommended input	1 ng per sample (gDNA) 1.2 mm per samples (FTA card punch)
Kit configuration	96 reactions and 384 reactions
Amplicon size	Mean: 235 bp; Maximum: 481 bp
Recommended multiplexing	8 to 96 samples on a MiSeq FGx Reagent Micro kit
Library prep time	7 hours 15 minutes (total) 1 hour 30 minutes (hands on)
Shelf life	6 months

Uniform high read coverage and well-balanced amplicons reduce allelic dropout

The maximum number of libraries that can be simultaneously sequenced depends upon the number of reads or the depth of coverage (DoC) desired per locus. Balanced read coverage across amplicons is also critical to locus and allele recovery. The ForenSeq MainstAY product line has a small number of markers multiplexed in a single sequencing reaction, resulting in high overall coverage.

Interlocus balance was determined using MainstAY libraries from 1 ng of input DNA from 46 male Coriell samples, including a positive amplification control and 49 female Coriell samples (9). Average coverage across the Au-STRs and Y-STRs for male samples was 669 and 469 reads, respectively. Average coverage for the Au-STRs for female samples was 1065 reads. The fold difference between the Au-STRs with highest and lowest depth of coverage was 15x, while the fold difference between the Y-STRs with highest and lowest reads was 4.6x. (Figure 1).

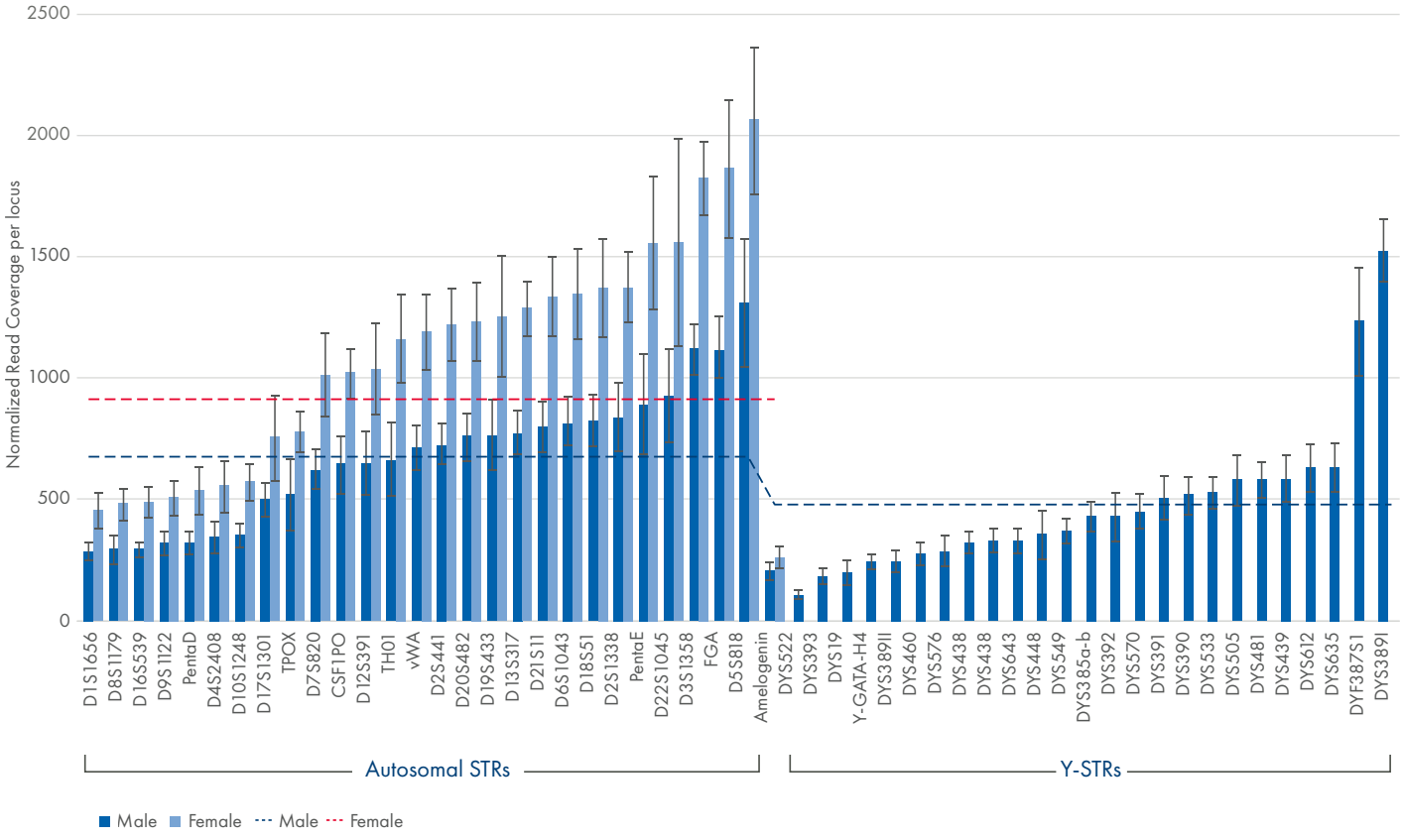

Figure 1.
Average DoC for markers included in the ForenSeq MainstAY Kit. Male samples are in dark blue and female samples are in light blue. The average coverage is shown using dotted lines, for males in dark blue and for females in pink.

The median interlocus balance for male samples, expressed as a percentage of total reads, was 1.9% (Au-STRs: 2.2% and Y-STRs: 1.6%). For female samples, the interlocus balance for the Au-STRs was 3.7% (Figure 2).

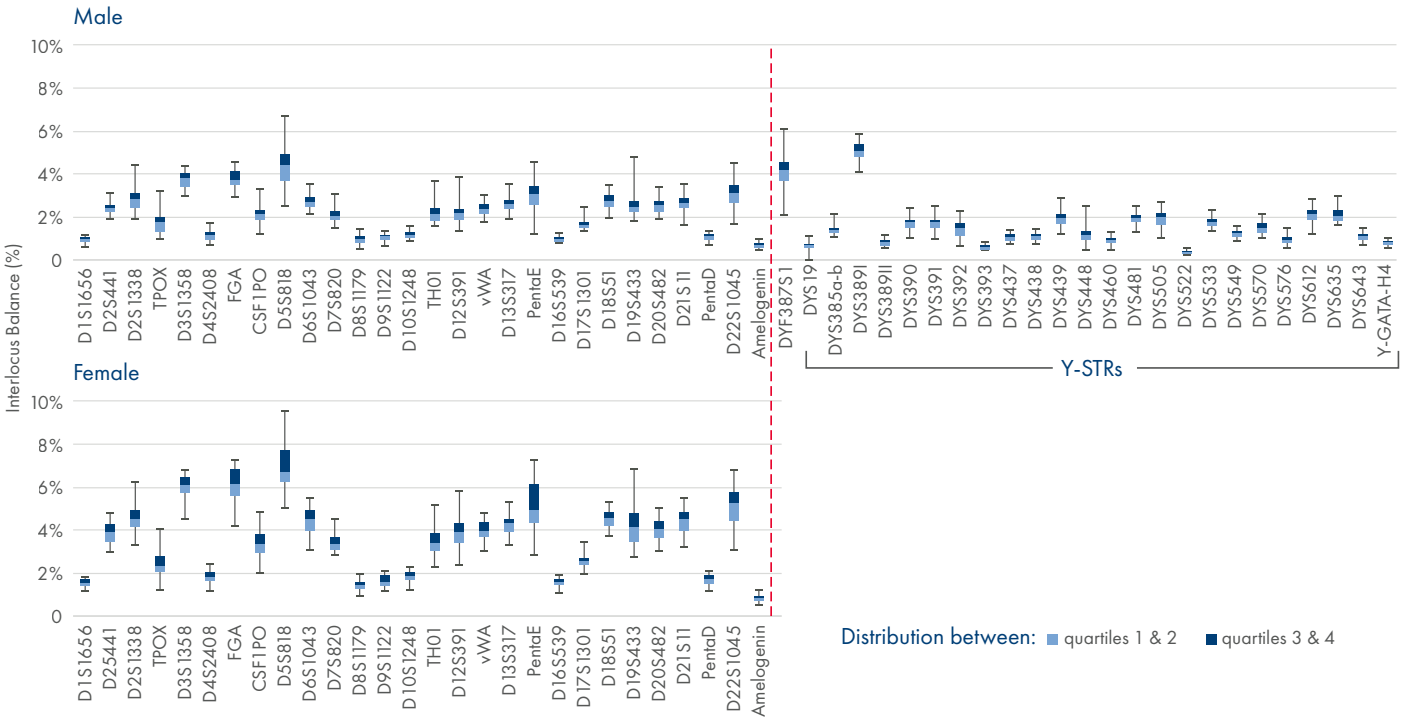

Figure 2.
Interlocus balance for genetic markers included in the ForenSeq MainstAY Kit. Autosomal and Y-STRs from male samples are represented in the top box plot. Interlocus balance for Autosomal STRs from female samples are represented in the bottom box plot.

The intralocus balance (ILB) or heterozygous balance (HB) was calculated as a ratio of the lower number of allele reads to the higher number of allele reads for a heterozygote pair. Compared to CE-based STRs, NGS shows a larger range for the ILB due to amplification biases for smaller amplicons during library preparation and sequencing. The ILB for heterozygotes and isometric genotype pairs was evaluated using the 1 ng DNA samples in this study for each sample and locus. All markers in the panel exhibited ILB values above 60% with a median ILB of 86%. The coverage and low variation of the ForenSeq MainstAY Kit indicates that amplicons are well-balanced, which reduces the likelihood of locus or allelic dropout and improves the identification of minor contributors in a mixture (Figure 3).

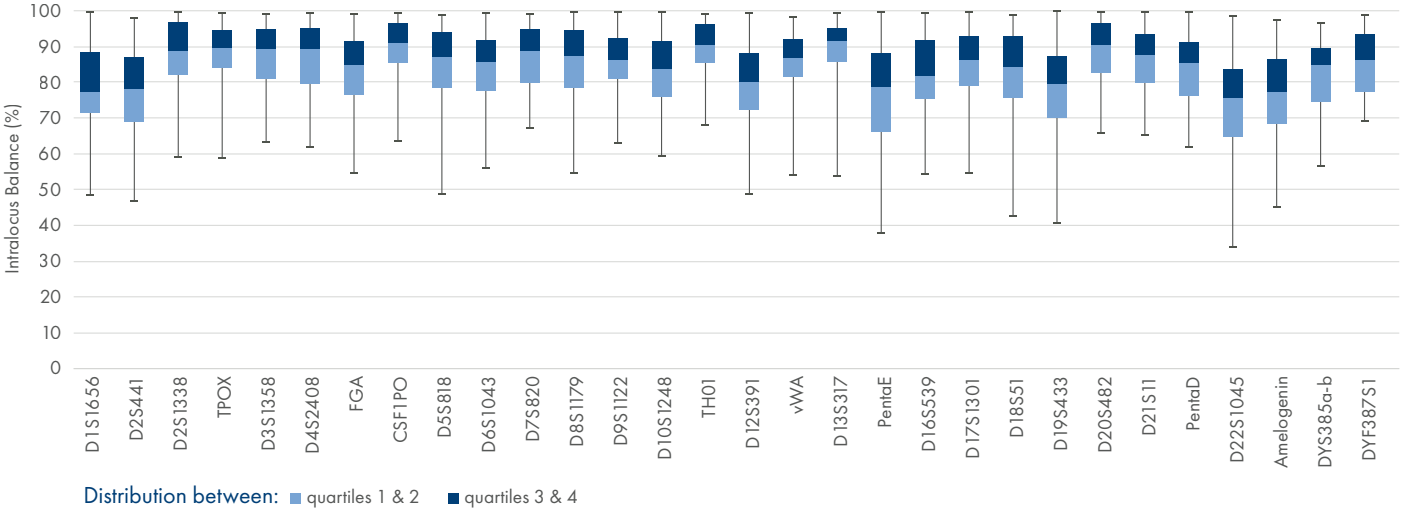

Figure 3.
Intralocus balance for genetic markers included in the ForenSeq MainstAY Kit with 1 ng DNA input. The boxes show the middle 50% or the interquartile range (IQR). Whiskers extend to the minimum and maximum ILB.

High sensitivity and multiplexing ability enable volume casework

The ability to generate genotypes and haplotypes across a range of inputs was evaluated using serially diluted gDNA in template amounts, in triplicate, of 1 ng, 500 pg, 250 pg, 125 pg, 62.5 pg, 31.25 pg, 15.625 pg and 7.82 pg. A total of 96 libraries were simultaneously sequenced using the MiSeq FGx Reagent Micro Kit. The ForenSeq MainstAY products offered a high level of sensitivity, generating full profiles (no allele loss) and 100% call rate with as little as 62.5 pg of input DNA. To determine allele call rate and comparative accuracy of the sequencing-based genotype data, the genotypes and haplotypes were generated with the MainstAY Analysis Module in the UAS using the default settings. These were then compared to orthogonal genotyping data from conventional genotyping methods (CE fragment length detection for STRs). Au-STRs demonstrated 100% concordance down to 62.5 pg, with a concordance of 95% even at 31 pg. Y-STRs demonstrated 100% concordance down to 125 pg, with 91% concordance down to 31 pg. All discordant calls were the result of stutter that exceeded the default stutter filter set for 1 ng samples. (Figure 4).

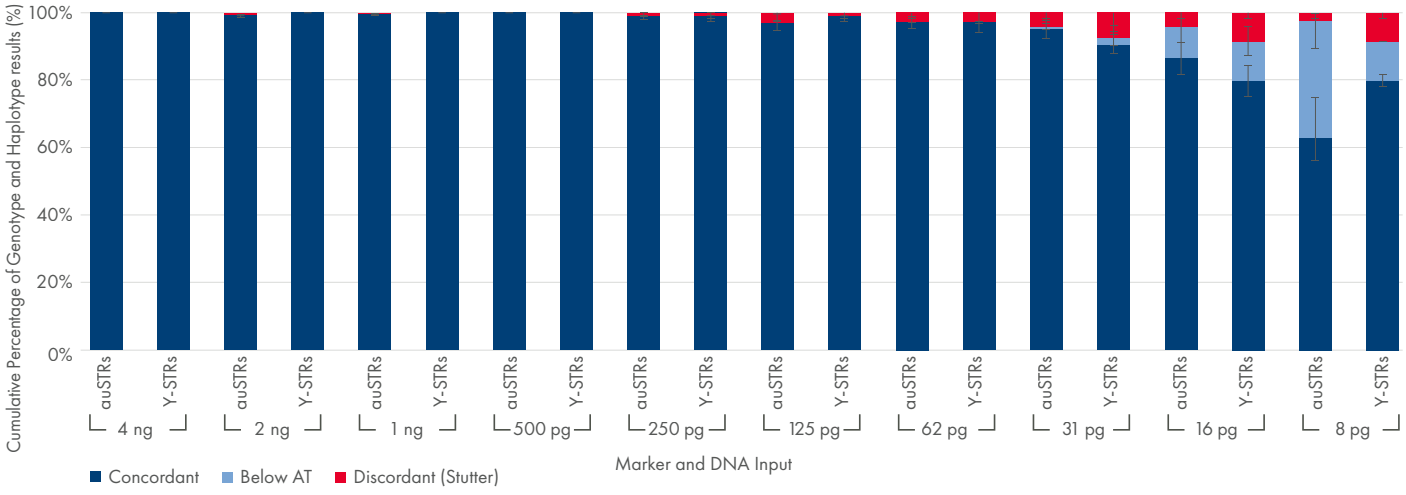

Figure 4.
Sensitivity study. Serially diluted gDNA samples amplified in triplicate. Genotype and haplotype outcomes are designated as "Concordant" (dark blue), "Below AT" (light blue), or "Stutter" (pink) based on concordance with orthogonal typing, and plotted as cumulative percentage of the total number of outcomes for each of the gDNA inputs. Shown are the average percentages of the samples.

Genotype and haplotype concordance with orthogonal technologies

To evaluate the concordance of the genotypes and haplotypes generated by the ForenSeq MainstAY Kit, 31 libraries including the NIST® SRM® 2391D standards (10), Coriell DNAs, a positive amplification control and three NTCs were prepared and sequenced in triplicate. Data were analyzed using the MainstAY analysis module in the UAS with default analysis thresholds and stutter filters. Accuracy of the genotypes and haplotype alleles was determined by comparing it to orthogonal CE data. Precision and call rates were determined using repeatability and reproducibility experiments. Libraries were prepared and sequenced from 15 control DNA samples in quadruplicate with one positive amplification control and three NTCs by three operators on three different MiSeq FGx instruments.

High accuracy was observed in Au-STRs (99.82%) and Y-STRs (100%). Six discordant alleles were detected out of 3261 Au-STR alleles. No discordant alleles were detected for the 774 Y-STRs alleles. Similarly, high precision rates of 99.87% and 99.42% were observed for Au-STRs and Y-STRs, respectively. All discrepancies were attributed to stutter alleles exceeding the default stutter filters (Table 3).

Table 3. Concordance with orthogonal technologies

	Autosomal STRs (%)	Y-STRs (%)
Accuracy	98.82	100
Precision	99.87	99.42
Call rate	99.87	99.42

Detection of minor alleles at low level contributions

To evaluate the ability of the ForenSeq MainstAY Kit to detect minor alleles in low level contributors, mixtures of control female and male samples were generated at 1 ng, with the female sample as the major contributor. The percentage of the minor contributor in the mixture ranged from 50% to 0.2%. The number of unique, unshared minor allele contributor alleles in the mixture was assessed using the default analytical and interpretation thresholds, as well as stutter filters in the MainstAY analysis module in the UAS (Figure 5).

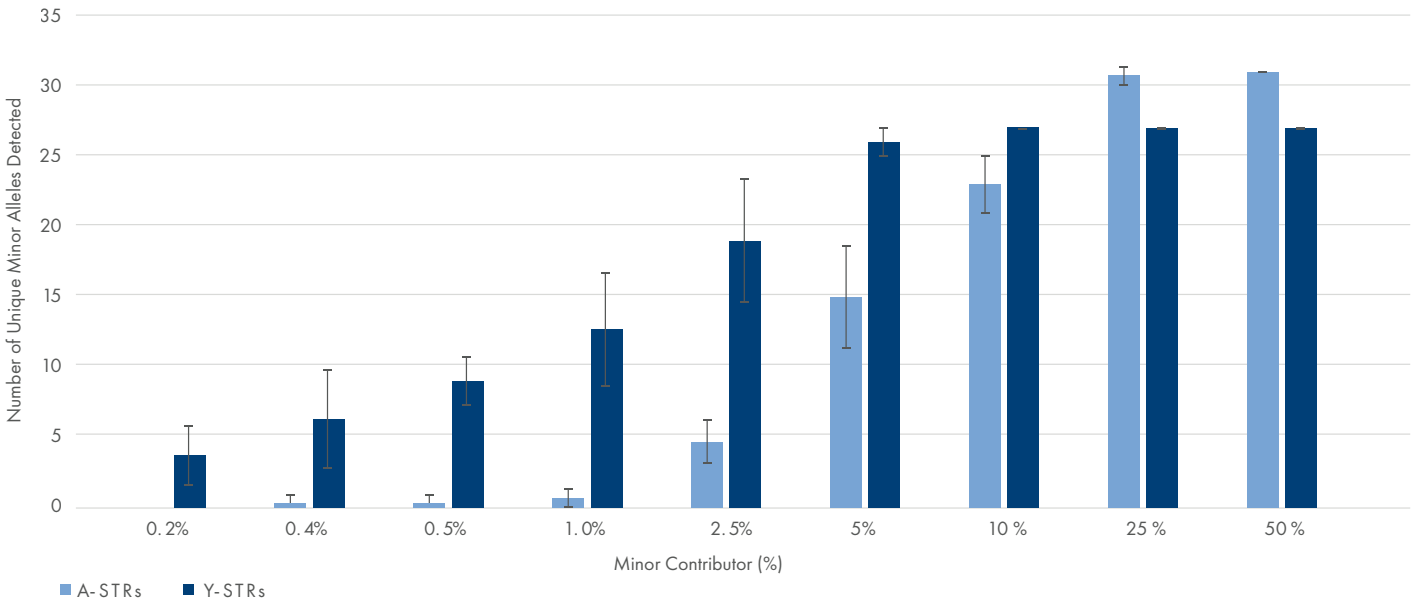

Figure 5.
Mixtures study. Genomic DNA mixtures of female:male control samples at serially diluted levels of the minor contributor. The number of unique autosomal alleles (light blue) and Y-STR alleles (dark blue) are shown.

At 50% minor contributor, 31 unique minor autosomal alleles and 27 Y alleles were detected. Approximately 15 unique minor autosomal alleles (50% unique Au-STR alleles) and 26 Y alleles were detectable at 5% minor contribution. At 1% minor contribution, over 10 Y-STR loci were detectable. ForenSeq MainstAY enables efficient identification of minor contributors in a mixture even at low level contributions.

Robust allele detection with degraded samples

The ability to type genotypes and haplotypes of degraded DNA samples with the ForenSeq MainstAY workflow was assessed using partially degraded gDNA that mimics forensic samples exposed to environmental and chemical stresses. Libraries from 1 ng of degraded blood from 30 samples, one positive amplification control and NTC, in triplicate, were prepared and 96 samples were sequenced using the Miseq FGx Reagent Micro Kit. Degradation index (DI) was used to assess the sample quality of the libraries. Samples with a DI of 1–4 were considered not degraded, samples with a DI of 4–10 were considered moderately degraded; those with a DI >10 were considered severely degraded. The samples tested with MainstAY had a DI that ranged from 1 to 56, spanning low, moderate and highly degraded samples. Results showed that 100% of STRs were accurately typed in samples with no degradation. Samples with moderate degradation showed >85% call rate of accurately typed STRs. Samples that were severely degraded had a call rate >71%, showing minimal loss of amplifiable STRs. The sample with the highest DI of 56 showed 60% recovery of alleles (40 of 74). ForenSeq MainstAY enables the detection of a high number of alleles even from severely degraded samples (Figure 6).

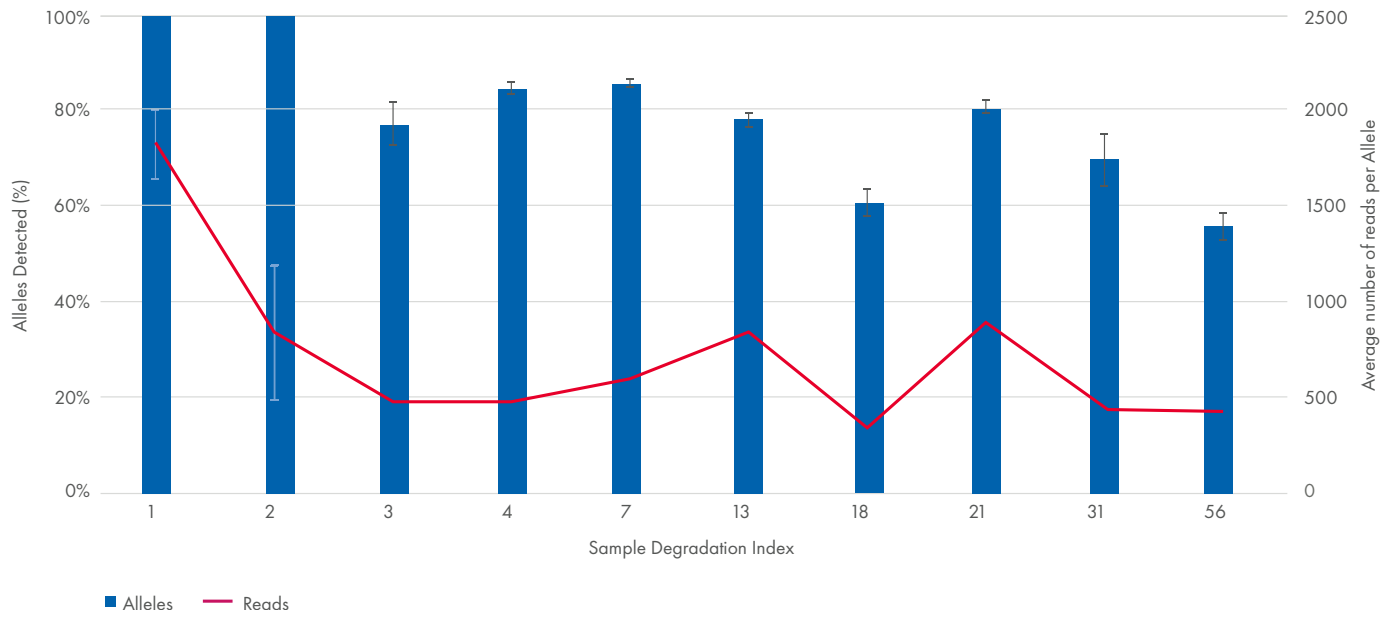

Figure 6.
Degradation study. Genomic DNA extracted from degraded blood samples. Samples were amplified and sequenced in triplicate. DIs spanning 1 to 56.

Fast, integrated end-to-end workflow

The ForenSeq MainstAY product line enables an end-to-end workflow in under 30 hours when used in conjunction with the MiSeq FGx Sequencing System, the MiSeq FGx Reagent Micro Kit and the MainstAY and MainstAY SE analysis modules in the UAS. The UAS enables the evaluation of sequencing data using familiar quality scores, threshold and guidance. Preconfigured analysis settings are provided within the solution that can be modified by the laboratory as needed. In addition, this module has been designed to increase the operational efficiency of high-volume forensic labs.

Streamlined project and sample overviews make it easy to review markers with QC indicators. New sorting and filtering capabilities allow users to categorize and evaluate a subset of STRs using a variety of QC indicators. Sequence analysis of STR alleles from populations of mixed alleles, such as isometric heterozygotes, is made possible by the ability to easily navigate between both allele calls and sequence level information.

Conclusion

The ForenSeq MainstAY product line offers an easy transition point to include NGS into your operational workflow. The ForenSeq MainstAY Kit contains reagents to amplify 27 autosomal loci (28 loci with the ForenSeq MainstAY SE Kit), 25 Y-STR loci and Amelogenin in a single reaction as small amplicons to generate sequencing data on the MiSeq FGx Sequencing System using the MiSeq FGx Reagent Micro Kit. With this workflow, both sequence and allele length polymorphism in the Au-STRs and Y- STRs can be identified in a single amplification. This eliminates the need to run multiple workflows and maximizes the informative value of a forensic sample, thereby increasing the statistical power of inclusion. NGS also eliminates the problem of overlapping alleles, allowing users to interrogate more alleles in the multiplex.

The combination of these STR loci and Amelogenin makes the ForenSeq MainstAY product line an effective tool for human identification while maintaining compatibility with existing databases worldwide. Low input recommendations of 1 ng gDNA deliver a reproducibly high call rate. The seamless integration of the MainstAY Analysis Module allows laboratories to easily analyze sequence data in a familiar and easy-to-use interface while generating reports in less than 1.5 days.

Ordering Information

Product	Contents	Cat. no.
ForenSeq MainstAY Kit (96)	Includes all the required reagents for 96 reactions to prepare sequencing libraries generating data for mainstream casework and forensic genetic genealogy confirmatory testing	V16000142
ForenSeq MainstAY Kit (384)	Includes all the required reagents for 384 reactions to prepare sequencing libraries generating data for mainstream casework and forensic genetic genealogy confirmatory testing	V16000128
ForenSeq MainstAY SE Kit (96)	Includes all the required reagents for 96 reactions to prepare sequencing libraries generating data for mainstream casework and forensic genetic genealogy confirmatory testing including the same markers as in ForenSeq MainstAY Kit + SE33	V16000183
MiSeq FGx Sequencing System	Desktop instrument with two run modes for a range of forensic genomics applications within a validated NGS workflow	15048976
ForenSeq Universal Analysis Software (UAS)	Software pre-installed as a dedicated server specific for forensic genomics for run setup, sample management, analysis and report generation. This product includes server, mouse, keyboard and monitor.	9003364
MiSeq FGx Reagent Micro kit	Supports up to 5 million paired-end reads for small batch sizes and faster turnaround times	20021681

Learn more about NGS for HID in your lab. Visit **qiagen.com/MainstAY**

References:

1 Edwards A, et al. DNA typing with trimeric and tetrameric tandem repeats: Polymorphic loci, detection systems, and population genetics. In: The Second International Symposium on Human Identification 1991; Promega Corporation, 31–52. *Am J Hum Genet.* 1991; 49: 746-56.

2 Edwards A, et al. Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups. *Genomics* 1992; 12: 241–53. doi: 10.1016/0888-7543(92)90371-x.

3 Bruijns B, et al. Massively parallel sequencing techniques for forensics: A review. *Electrophoresis.* 2018; 39(21): 2642-2654. doi: 10.1002/elps.201800082.

4 Yang Y, et al. Application of next-generation sequencing technology in forensic science. *Genomics Proteomics Bioinformatics* 2014; 12: 190–197. doi: 10.1016/j.gpb.2014.09.001.

5 Gettings KB, et al. STR allele sequence variation: Current knowledge and future issues. *Forensic Sci Int Genet.* 2015; 18: 118–130. doi: 10.1016/j.fsigen.2015.06.005.

6 Hares DR. Selection and implementation of expanded CODIS core loci in the United States. *Forensic Sci Int Genet.* 2015; 17: 33–34. doi: 10.1016/j.fsigen.2015.03.006

7 The Y-Chromosome Haplotype Reference Database (YHRD). https://yhrd.org/pages/resources/composition

8 National Institute of Standards and Technology (NIST), Core STR Loci Used in Human Identity Testing https://strbase.nist.gov/

9 DNA samples from the National Institute of General Medical Sciences (NIGMS) Human Genetic Cell Repository at the Coriell Institute for Medical Research. https://www.coriell.org/

10 NIST Material Details: SRM 2391d - PCR-Based DNA Profiling Standard https://www-s.nist.gov/srmors/view_detail.cfm?srm=2391d