Application Note

VEROGEN

High-Quality Outcomes from Low-Quality Samples with ForenSeq Kintelligence

Flexible and purpose-built library prep that delivers exceptional sensitivity, coverage, and call rates.

Highlights

- Superior performance to microarrays
  Sequence a spectrum of degraded samples.
- High call rates down to 50 pg
  Achieve excellent data quality from low input.
- Robust profiles with challenging samples
  Analyze inhibited and contaminated samples.

Introduction

The prevailing methods of generating DNA profiles for comparison in genealogy databases, genotyping by microarray and whole-genome sequencing (WGS), can be incompatible with the unique needs of forensic casework. The ForenSeq® Kintelligence Kit offers a fit-for-purpose solution that employs a targeted sequencing approach to optimize recovery rates for low-quality forensic samples. Designed, developed, and tested as part of an integrated next-generation sequencing (NGS) workflow, the kit works in conjunction with the Verogen MiSeq FGx® Sequencing System for a powerful combination that consistently delivers deep coverage and high call rates. The MiSeq FGx System is built on proven Illumina sequencing-by-synthesis (SBS) chemistry. This application note showcases the exceptional performance of the ForenSeq Kintelligence Kit on a variety of degraded,inhibited, and contaminated samples.[1]

Limitations of current methods

Microarray-based methods are designed for population genotyping or direct-to-consumer (DTC) tests, both of which require large quantities of high-quality DNA. Environmental exposure and source material age often result in degradation of forensic samples and microbial contamination. Cold case samples, for instance, are often degraded due to advanced age, variable extraction and storage procedures. While microarrays can generate data from challenging forensic samples, this data is usually low-quality. They may also fail with contaminated or degraded samples. Moreover, the high number of markers in arrays, which range from 650,000 to 2.5 million, contain extensive counts of medically relevant single nucleotide polymorphisms (SNPs). These SNPs are not useful for forensic genetic genealogy (FGG) and create genetic data privacy concerns.

Although WGS supports samples that are low-quality and low-quantity, and hence cannot be processed on a microarray, the workflow includes several steps that are not aligned with forensic best practices. During library prep, forensic samples are often pooled with research, clinical-research, or clinical samples that are not related to the investigation. When adequate data are unavailable, WGS uses a non-standardized bioinformatics method called imputation, which adds bases to the sequencing output. WGS also tends to be significantly more expensive than arrays or targeted sequencing. Crucially, WGS and array manufacturers did not design or validate these platforms for forensic use and do not support forensic applications (Table 1).

Table 1: Comparison of methods for FGG and degraded sample processing

Metric	Microarray[2–3]	WGS[4–6]	Targeted Sequencing
Sample types*	gDNA from blood, buccal swabs, fresh frozen, saliva, solid tumors	gDNA from fresh blood, FFPE, saliva	gDNA from degraded blood, bone, hair, teeth, semen, and buccal swabs
Sample prep method	Hybridization enrichment	Mechanical or enzymatic fragmentation and probe-based enrichment	PCR enrichment
Typing method	Hybridization capture and single-base extension	SBS	SBS
Compatible instruments	iScan System	HiSeq X System, NextSeq Systems, or NovaSeq 6000 System	MiSeq FGx System
Forensically designed platform*	No	No	Yes
Forensic application support*	No	No	Yes
DNA input*	200 ng	50 ng–50 pg	≤ 1 ng
Partially degraded DNA processing	No	Yes	Yes
Marker content	650,000–2,500,000 SNPs	3 billion data points	10,230 forensically curated SNPs
Medical relevance	Yes	Yes	No
Analysis software	GenomeStudio	Multiple custom tools	Universal Analysis Software
Bioinformatics expertise	Required	Required	Not required
Data manipulation	Not required	Imputation	Not required
Uploading to GEDmatch PRO	Manual report configuration	Manual report configuration	Preconfigured report

* Manufacturer guideline

Superior performance to microarrays

To assess the performance of targeted sequencing versus arrays, Verogen compared call rates from degraded blood samples with degradation indexes (DIs) of 1–460 using two platforms: the ForenSeq Kintelligence Kit with the MiSeq FGx System and the Infinium Global Screening Array-24 Kit with the iScan System. Due to the low recommended input volumes of the Global Screening Array (4 µl), total gDNA input amounts varied between 0.69 to 3.46 ng. The high input volume capacity of ForenSeq Kintelligence at 25 µl enabled 1 ng gDNA input for testing.

Across all levels of degradation, quantified using the Innogenomics InnoQuant HY kit, targeted sequencing showed superior results compared to the array-based genotyping. ForenSeq Kintelligence reproducibly delivered an average locus call rate of 98.8% with > 10,000 SNPs called. Upload to GEDmatch PRO requires ≥ 7000 SNPs called, a threshold that all samples met. The blood sample with the highest DI had a locus call rate of 95.6%, a minimal SNP dropout compared to a sample with a DI of 1 and a locus call rate of 100%. In sharp contrast, even low degradation levels had a substantial impact on Global Screening Array performance, which showed a steep decline in call rate as the DI increased (Figure 1). Samples with a DI > 2 had a locus call rate below 80% and samples with a DI > 4 had a locus call rate < 60%, limiting the accuracy of downstream kinship analysis.

ForenSeq Kintelligence delivered superior performance with degraded samples when compared to Global Screening Array results. The consistent, high-confidence SNP calls generated by ForenSeq Kintelligence makes it ideal for delivering excellent kinship results without multiple extractions to meet the high input amounts required by microarrays.[7]

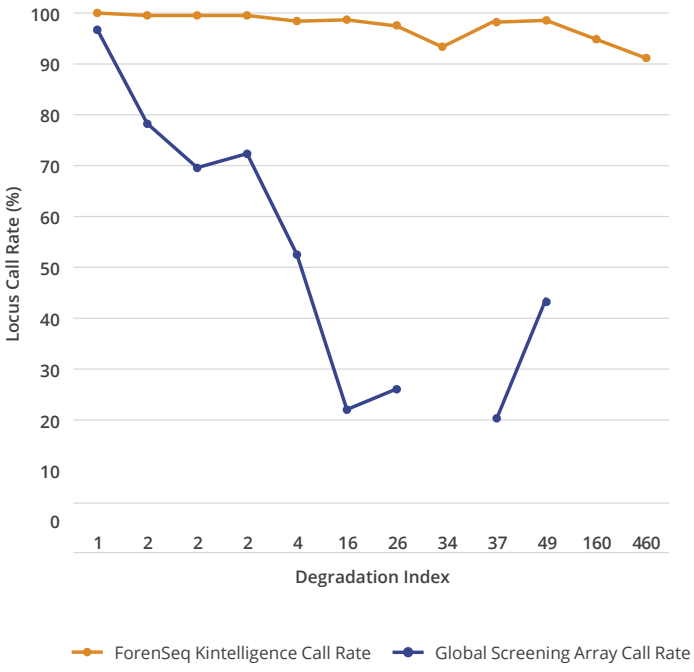

Figure 1: A comparison of locus call rates shows the superior performance of targeted sequencing, which maintained high rates across the range of DIs. Samples are 1 ng gDNA extract from blood artificially degraded by mechanical shearing. Due to low input, GSA failed on the sample with a DI of 34.

Degradation assessment with casework-type samples

To supplement the microarray comparison and characterize ForenSeq Kintelligence performance on degraded samples, Verogen conducted studies analyzing blood and bone samples. All samples exhibited similar performance to positive amplification control and generated high average call rates.

Blood samples

For the study of degraded blood, Verogen analyzed 1 ng gDNA extracted from blood and mechanically sheared by Covaris (Innogenomics), resulting in a degradation series with DIs of 1–158. DIs were calculated using the InnoGenomics InnoQuant HY kit. Calling an average of 9630 SNP loci, ForenSeq Kintelligence achieved an average locus call rate of 94.1% (Figure 2). The most degraded sample reached a locus call rate of 81% with 8263 SNPs called.

Contemporary bone samples

For the study of degraded bones, Verogen analyzed 1 ng gDNA extracted from degraded bone samples exposed to a variety of environmental insults, such as commercial cremation, partial decomposition with internment, and embalming (Table 2, Source: Sam Houston State University). The burnt bones were sourced from frozen cadavers placed in a controlled environment for two days and then burned. The post-mortem interval of these samples is 1–8 years. Commercially cremated samples were sourced from a funeral home.

DNA was extracted using two methods common in forensic laboratories: total demineralization and the Applied Biosystems PrepFiler Forensic DNA Extraction Kit. DIs were calculated using the Applied Biosystems Quantifiler Trio DNA Quantification Kit. Average number of loci called with the ForenSeq Kintelligence Kit was 10,148 (call rate: 99.2%). The most degraded sample had a 99.7% loci call rate with 10,196 SNPs called. The lowest call rate for a sample was 98.2% with 10,049 SNPs called

(Figure 3). These call rates indicate robust results overall for this particularly challenging sample type, regardless of extraction method or insult.[8]

Interred bone samples

Verogen also studied a set of aged interred bones. Although extreme age and severe degradation prevented calculating DIs for all the samples, mitochondrial DNA (mtDNA) copy numbers helped ascertain a DI of 232 for one of them. This sample generated 7424 SNPs for a 71.5% locus call rate. A sample with an unknown DI achieved the highest locus call rate of 96.4% with 10,004 SNPs called. By achieving high-quality results across a spectrum of degraded samples and generating a SNP report that can be uploaded to GEDmatch PRO, ForenSeq Kintelligence reduces the need for multiple DNA extractions, bioinformatic manipulation of results, and resequencing.

Inhibited casework-type samples and Human DNA specificity studies

Verogen analyzed teeth to characterize kit performance on samples with potential inhibition. In addition, assessments on species-specificity and microbial contamination were performed. Teeth exhibited similar performance as the positive amplification controls with high average call rates. Sample evaluations on species specificity and microbial contamination produced similar results to the negative amplification controls.

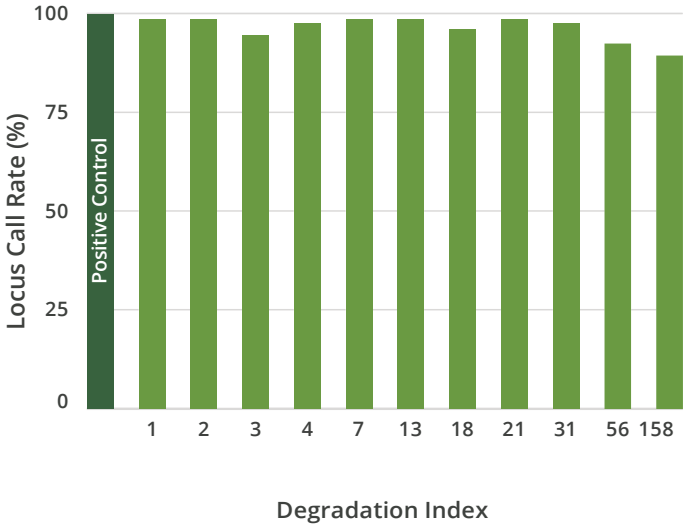

Figure 2: Degradation study. Genomic DNA extracted from blood. Dġ spanning 1 to 158. Average call rate is 96.7%

Figure 3: Degradation study. Genomic DNA extracted from contemporary bones. Dġ spanning 1 to 14. Average call rate is 99.2%

Table 2: Summary of bone samples and the associated degradation index

Sex	Ethnicity	Bone	DI	Insult	Extraction Method
Male	Caucasian	Vertebral arch	14	Cremated	PrepFiler Forensic DNA Extraction Kit
Male	Caucasian	Humerus	4	Embalmed	Total demineralization
Male	African American	Femur	6	Embalmed	PrepFiler Forensic DNA Extraction Kit
Male	Caucasian	Tibia	1	Early decomposition	PrepFiler Forensic DNA Extraction Kit
Female	Caucasian	Femur	2	Burned	Total demineralization
Female	Caucasian	Femur	2	Burned*	PrepFiler Forensic DNA Extraction Kit
Female	Caucasian	Humerus	7	Burned*	PrepFiler Forensic DNA Extraction Kit

Inhibition study

Alongside bones, teeth are frequently the only available sources of DNA for identification of degraded or fragmented human remains. Because their composition and location in the jawbone affords some protection, teeth are preferred over bones for forensic applications.

Verogen analyzed 1 ng gDNA extracted from five contemporary tooth samples. DNA was extracted using the Dental Forensic Kit (DFK). Calling an average of 10,163 SNPs, the ForenSeq Kintelligence Kit achieved an average locus call rate of > 99%, as shown in Figure 4, making all samples eligible for upload to GEDmatch PRO. Even the lowest performing sample had a locus call rate of 98.9%.

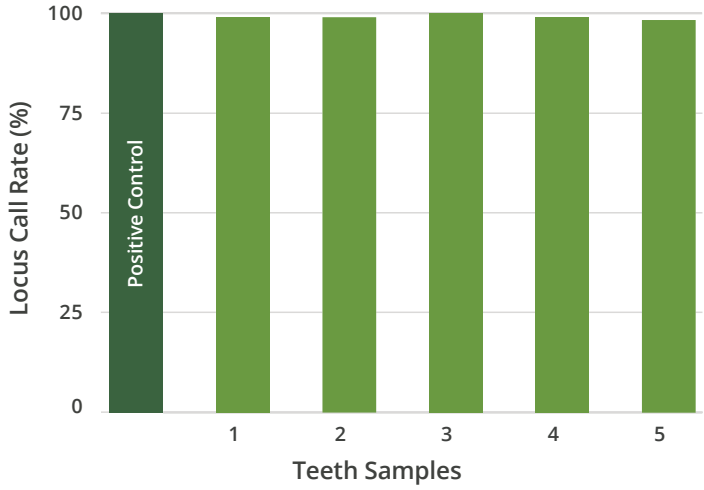

Figure 4: Genomic DNA extracted from five teeth samples generated an average call rate of 99 %

Human DNA specificity

Forensic samples, especially those recovered from unidentified human remains, are often contaminated with non-human DNA, which limits effective recovery of human DNA for forensic evaluation. The ForenSeq Kintelligence Kit features a primer set designed for human DNA. To test ForenSeqKintelligence for species specificity, Verogen prepared libraries in triplicate from 1 ng gDNA extracted from nonhuman samples spanning procaryotes and eucaryotes. The samples included the following species plus a positive control and a negative control:

- One E. coli bacterial sample[9]
- One avian species, a domesticated rooster[10]
- One Old World monkey, a rhesus macaque[11]
- One male cynomolgus monkey[12]
- Seven non-primate mammals: Cat[10], Male bovine[10], Male dog[10], Male horse[10], Male mouse[13], Male porcine[10], Male rat[14]

The ForenSeq Kintelligence Analysis Module in Universal Analysis Software (UAS) analyzed the data using the default settings for analytical threshold (AT), interpretation threshold (IT), and intralocus balance. The negative control had 36 SNPs called (locus call rate: 0.35%), with comparatively low rates across all samples (Table 3, Figure 5). Samples from the two monkey species unsurprisingly produced locus call rates of 10.2% and 10.8% respectively. Overall, the amplifications of nonhuman species with the ForenSeq Kintelligence Kit are comparable to current forensic DNA typing conducted with capillary electrophoresis (CE).[15,16]

Table 3: Call rates demonstrating human specificity

Species	Avg Call Rate (%)	Number of SNPs Called
E. Coli	1.52	156
Avian	1.15	118
Porcine	4.35	445
Non-primate mammals except porcine	0.69	71
Rhesus macaque	10.2	1105
Cynomolgus monkey	10.8	1043

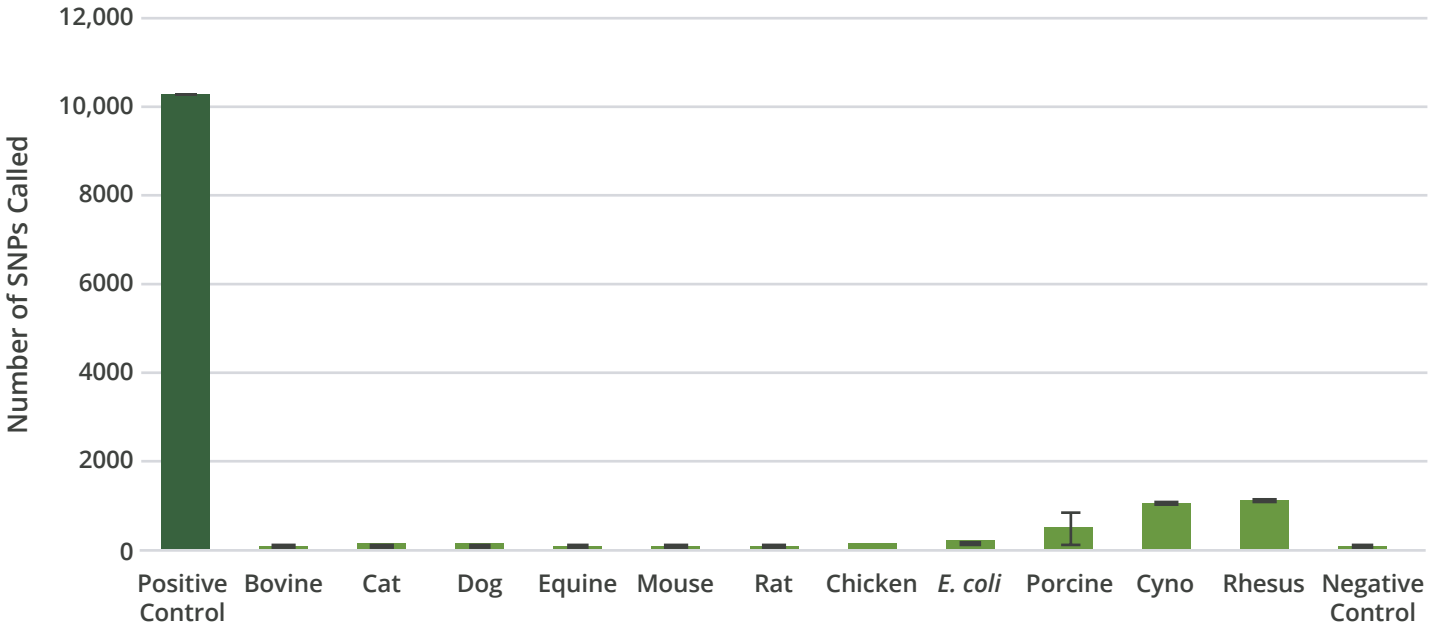

Figure 5: Bacterial, avian, and nonhuman mammal samples demonstrated low numbers of SNPs called, indicating high human specificity.

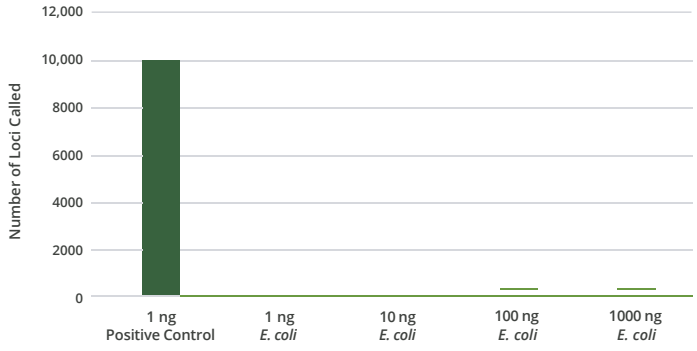

Figure 6: Input titration of *E. coli* samples demonstrated low numbers of SNPs called, indicating that the assay has high specificity for human DNA.

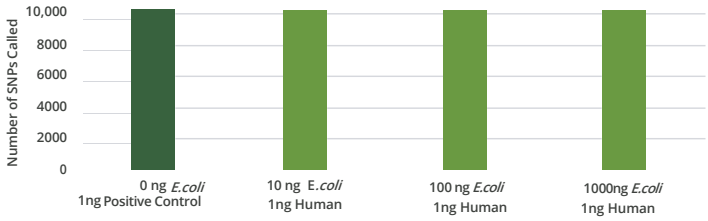

Figure 7: 1 ng human DNA samples spiked with increasing amounts of *E. coli*, consistently demonstrating high numbers of SNPs called.

Microbial sensitivity and contamination studies

Microarrays may fail in the presence of microbial contamination, particularly with samples from challenging sources such as bones and teeth. A study using a 1 ng positive amplification control and libraries generated from 1 ng, 10 ng, 100 ng, and 1000 ng input DNA extracted from *E. coli* demonstrated that even in the presence of high concentrations of *E. coli,* ForenSeq Kintelligence amplifies human DNA with a high degree of specificity. When compared to the positive control, which called all 10,230 SNPs, the 1 ng *E. coli* sample recovered 156 SNPs for a 2% locus call rate. The highest concentration of *E.coli* , at 1000 ng called 7 SNPs (locus call rate: 0.07%) (Figure 6)

To further test kit performance in the presence of microbial contamination, Verogen spiked 1 ng human DNA samples with increasing inputs of *E. coli* at ratios of 1:10, 1:100, and 1:1000. Compared to the 100% call rate of the positive control, the *E. coli* sample with the highest contamination of 1000 ng had the lowest locus call rate of 99.93%, with 10,168 SNPs called (Figure 7). Overall, the ForenSeq Kintelligence Kit is robust in the presence of microbial contamination.

Sensitivity study with control samples

Verogen optimized ForenSeq Kintelligence to reproducibly generate SNP calls across a range of input amounts and replicates.

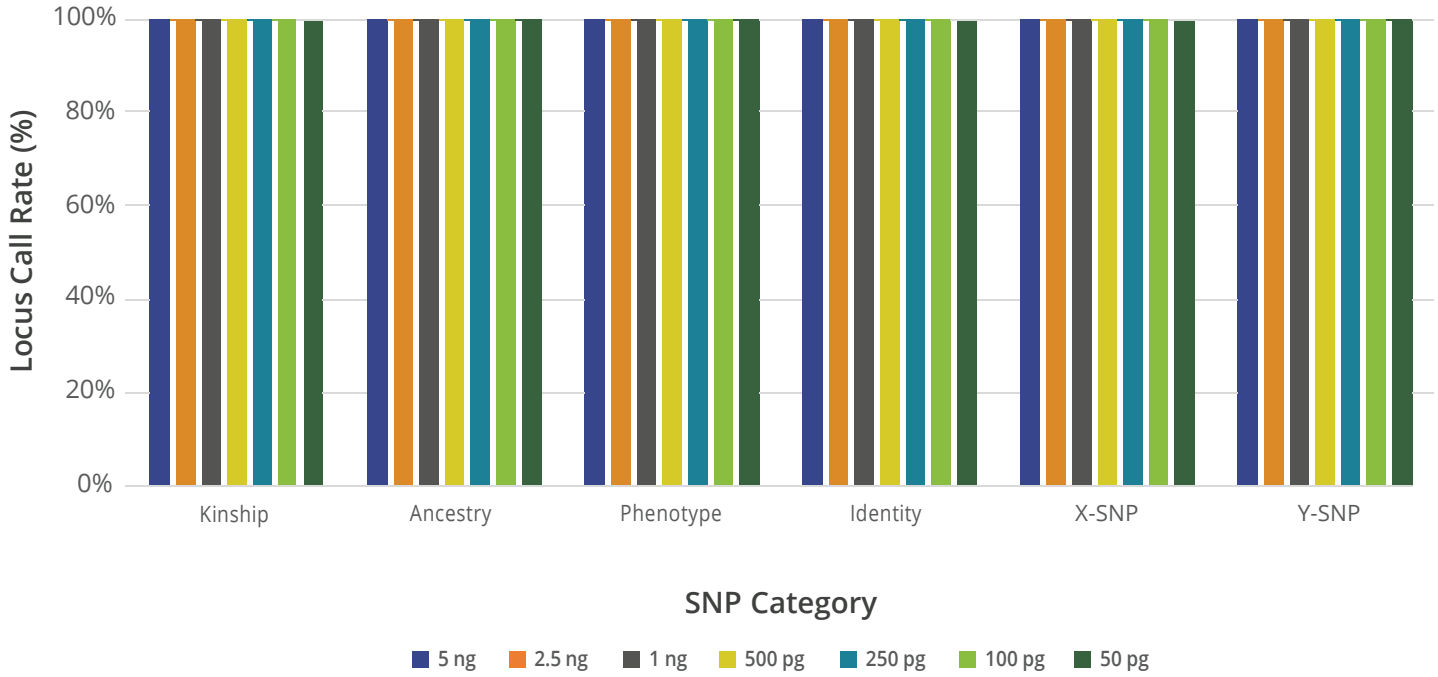

Figure 8: A sensitivity titration of control DNA at a range of inputs shows SNP call rates across all six SNP categories included in ForenSeq Kintelligence. Call rates are 100% for 5 ng to 250 pg, 99.8% at 100 pg, and 99.6% at 50 pg.

A sensitivity study assessed total DNA inputs of 5 ng, 2.5 ng, 1 ng, 500 pg, 250 pg, 100 pg, and 50 pg. The recommended input of 1 ng achieved a 100% locus call rate while the sample with the lowest input, 50 pg, had a 98.9% locus call rate. From at least 250 pg, a full SNP genotype was generated across all SNP categories.

The high average locus call rate is reproducible across all six SNP categories. Ancestry, phenotype, and Y-SNPs achieved a 100% locus call rate across inputs from 5 ng to 50 pg. Identity and X-SNPs had a slightly lower rate of 99% at 50 pg. In both cases, each category suffered a dropout of only one SNP. Kinship SNPs reached a 100% locus call rate down to 250 pg—well below the recommended input of 1 ng. The 100 pg level accurately generated a 99.9% kinship SNP locus profile with four dropouts. At 50 pg, the locus call rate was 99.6% with a dropout of 32 SNPs out of 9867 total kinship SNPs, enabling the generation of a SNP report that can be uploaded to GEDmatch PRO across the entire sensitivity study. With a large multiplex at very low input levels, dropouts are expected stochastic effects. Ultimately, ForenSeq Kintelligence is a highly sensitive kit that generates profiles with > 98.9% locus call rates at input amounts ranging from 5 ng to 50 pg (Figure 8).

Conclusion

FGG provides critical information to analysts and investigators. Due to high input requirements and oversized SNP sets, FGG has historically been challenging to apply to forensic samples, particularly those that are degraded. However, by combining curated content design with a low input requirement and targeted sequencing, forensic laboratories can obtain an accurate DNA profile from degraded and inhibited samples for FGG purposes. The ForenSeq Kintelligence Kit reproducibly delivers high-quality results across a broad rangeof low-quantity samples. Furthermore, a forensically validated system, intuitive software, and dedicated genealogy database provide efficient and accurate comparisons for long-range kinship outcomes whilst also minimizing data privacy concerns. This workflow empowers all forensic laboratories to generate investigative leads for cases of violent crimes, missing persons, and unidentified human remains.

Learn more at **verogen.com/products/forenseq-kintelligence-kit**.

References

1. Tillman, Andreas, Peter Sjölund, Bo Lundqvist, Therese Klippmark, Catja Älgenäs, and Henrik Green, "Whole-genome sequencing of human remains to enable genealogy DNA database searches – A case report," *Forensic Science International: Genetics* 46 (January 2020): 102233. https://doi.org/10.1016/j.fsigen.2020.102233.
2. Illumina, *Infinium Global Screening Array-24 v3.0 BeadChip Data Sheet*, 2020, pub. no. 370-2016-016-G.
3. Illumina, *Infinium Omni2.5-8 v1.5 BeadChip Data Sheet*, 2019, pub. no. 370-2011-008-C.
4. Illumina, *Nextera XT DNA Library Preparation Kit Data Sheet*, November 2014, pub. no. 770-2012-011.
5. Illumina, *TruSeq DNA PCR-Free Data Sheet*, 2017, pub. no. 770-2013-001-D.
6. Illumina, *TruSeq DNA Nano Data Sheet*, 2020, pub. no. 770-2013-012-F.
7. InnoGenomics InnoQuant HY (https://innogenomics.com/products/innoquant-hy/)
8. Applied Biosystems PrepFiler Forensic DNA Extraction Kit (https://www.thermofisher.com/order/catalog/product/4463351?us&en#/4463351?us&en)
9. Affymetrix, lot 4345173
10. Zyagen, lot 771619G010
11. Zyagen, lot 662519GM140R
12. Zyagen, lot 060818GM140C
13. Zyagen, lot 1102120GMW150M
14. Zyagen, lot 1102120GRL180M
15. "DNA Profiling," ScienceDirect, accessed March 23, 2021, https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/dna-profiling.
16. "Capillary Electrophoresis," ScienceDirect, accessed March 23, 2021, https://www.sciencedirect.com/topics/medicine-and-dentistry/capillary-electrophoresis.

Verogen • 1.833.837.6436 toll-free (North America) • +1.858.285.4101 tel • +44 (0) 208 054 8706 (United Kingdom) info@verogen.com • www.verogen.com