Document Title:	**Comparative Evaluation of Genotyping Technologies for Investigative Genetic Genealogy in Sexual Assault Casework**
Author(s):	**Sarah Cavanaugh, Madalyn Bowers, Robert Bever, Ph.D., Melinde Byrne, Dane Plaza, Jonathan M. Davoren**
Document Number:	**310017**
Date Received:	**January 2025**
Award Number:	**15PNIJ-21-GG-04143-MUMU**

Comparative Evaluation of Genotyping Technologies for Investigative Genetic Genealogy in Sexual Assault Casework

Technical Summary
Bode Technology
June 30, 2024

Report submitted to the National Institute of Justice
Grant # 15PNIJ-21-GG-04143-MUMU
Project Period: 01/01/2022 to 06/30/2024
Award Amount: $437,383

Sarah Cavanaugh – Lead Researcher
Madalyn Bowers
Robert Bever, PhD
Melinde Byrne
Dane Plaza – Grant Administrator
Jonathan M. Davoren **– Principal Investigator**
Bode Technology
10430 Furnace Rd, Suite 107
Lorton, VA 22079
Jonathan.Davoren@bodetech.com

This project was supported by Award No. 15PNIJ-21-GG-04143-MUMU, awarded by the National Institute of Justice, Office of Justice Programs, U.S. Department of Justice. The opinions, findings, and conclusions or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect those of the Department of Justice.

Table of Contents

Project Summary

Goals and Objectives

Investigative Genetic Genealogy (IGG) offers a capability to identify investigative leads when CODIS searching is unproductive, and IGG can provide time efficient methods for removing perpetrators of serial violent crimes, such as rape and murder from the community, thereby increasing public safety. However, use of IGG has preceded establishment of best practices. The 2021 TWG operational requirements identified the need for further development, assessment, and evaluation of IGG testing procedures for use by crime labs [1]. This study supports the TWG requirements by assessing the ability of genotyping technologies to develop useful profiles from low-template and degraded sexual assault samples for genealogical searching in law enforcement accessible Direct-to-Consumer (DTC) genealogical databases and support rapid, accurate, efficient identification of the samples' source.

In Phase I, genotyping by Illumina's Infinium™ Global Screening Array (GSA) BeadChip, genome sequencing on Illumina NovaSeq 6000, and targeted sequencing with Qiagen/Verogen ForenSeq® Kintelligence Kit on the MiSeq FGx® Sequencing Sytem were compared for sensitivity to low-level DNA input concentrations and specificity for artificially degraded DNA using whole semen and nascent semen DNA samples. The high-density SNP genotype profiles were compared against databased genotypes in order to determine the maximum distance at which known or potential genealogical associations can be identified. In Phase II, the limitations were further tested by generating a mock case scenario with laboratory-created challenging samples exhibiting both low-level concentration and DNA degradation utilizing a known donor for whom verified family members of relationship distance greater than 5th degree or 2nd cousin are present in DTC databases. After genotyping mock samples with each technology, a full genealogical investigative workflow conforming to the Genealogical Proof Standard [2] was applied to demonstrate whether or not increasingly distant relatives can be identified and at what distance identification is no longer possible.

Research Questions

This study evaluated the three technologies currently available for developing high-density SNP genotypes from human DNA samples and compared their abilities to generate profiles from challenging forensic samples related to sexual assault casework across two separate phases. More specifically, this projected sought to investigate how low-template DNA (e.g., around 1-2 ng inputs) and highly degraded DNA would affect the quality, accuracy, and reproducibility of high-density SNP genotypes and ultimately affect the performance of investigative genetic genealogy (IGG) to identify potential relatives in the GEDmatch PRO database by answering the following questions:

1) What are the effects to overall call rates (total number of SNPs genotyped) when analyzing low-template and degraded semen samples?
2) What are the effects to genotype concordance compared to a high-quality known reference when analyzing low-template and degraded semen samples?
3) What are the effects to technology-specific quality metrics when analyzing low-template and degraded semen samples that would help an investigator assess genotype quality prior to upload for genealogical comparison?

4) How do the genotyping results of each technology impact the ability to support rapid, accurate, and efficient identification of the samples' source?
5) What are the impacts to matching range and accuracy within a large, multi-generational family when using genotypes of disparate numbers of target SNPs (i.e., 10,230 SNPs targeted with Kintelligence vs over 2 million SNPs obtained with genome sequencing)?

Summary of Project Design and Methods

PHASE I

Sample Preparation for Sensitivity Analysis

Each technology was assessed for sensitivity using three DNA samples: two semen samples and one NIST Reference Material (RM). Semen samples were collected from two unrelated male donors under informed consent. Each donor had at least one known relative present in the GEDmatch database. Family trees for the two semen donors are provided in Figure 1. NIST RM8393 was purchased as a control DNA sample. It is provided as extracted DNA from a cell line derived from a male individual of Chinese ancestry.

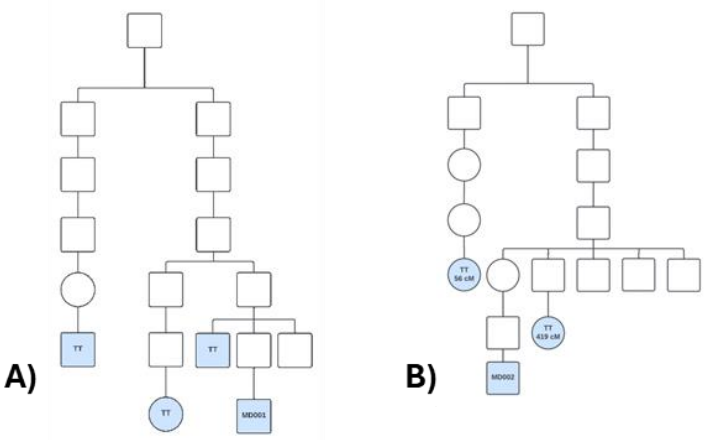

Figure 1: Family tree construction of known semen donors MD001 (A) and MD002 (B). A previously unknown distant relative was identified for each donor, represented by the highlighed box to far left on each tree. TT = test taker, designating the relationship of matching GEDmatch kits.

DNA extraction of known donor semen samples was performed on the EZ1 BioRobot instrument with the EZ1 DNA Investigator® Kit (Qiagen, Germantown, MD) and a Bode-optimized DTT lysis protocol. DNA quantification was performed in triplicate on all samples with the Quantifiler® Trio DNA Quantification Kit (Thermo Fisher Scientific, Waltham, MA) in 11 µl reaction volumes on the Applied Biosystems® 7500 Real-Time PCR System. The extracts were then split into two aliquots. One aliquot was retained at Bode for serial dilution, ForenSeq Kintelligence library preparation, library prep QC, and MiSeq FGx sequencing following manufacturer's recommended protocol [3]. The second aliquot was submitted to Gene by Gene (Houston, TX) for microarray and genome sequencing processing following their internally validated SOPs. First, a 200 ng aliquot of each donor and the control sample was processed via microarray genotyping with Illumina GSAv2 BeadChip [4] to serve as the ground truth genotype for comparisons of all other samples. Next, each DNA extract was serial diluted to generate DNA inputs ranging from 50 ng to 0.025 ng, depending on the technology. The maximum and minimum DNA concentrations tested with each of the three genotyping technologies varied based on the

optimal input range for each technology (Table 1). A total of 57 samples were processed for this task. Extraction positives, reagent blanks, and amplification positive and negative controls were processed alongside the samples, as appropriate.

Table 1: DNA concentrations that will be tested with each technology.

Technology	Optimal Input (ng)	Sensitivity Range (ng)	Replicates	Total Samples
BeadChip	200	200, 50, 10, 2, 1, 0.5, 0.25	3	21
WGS	10	50, 10, 2, 1, 0.5, 0.25	3	18
Kintelligence	1	2, 1, 0.5, 0.1, 0.05, 0.025	3	18

Data Analysis

For samples processed with the ForenSeq Kintelligence kit, raw data analysis was performed using the ForenSeq Universal Analysis Software (UAS) [5], which is part of the MiSeq FGx Forensic Genomics System. Following ForenSeq sequencing, automated processing of raw reads occurred through the UAS. During this process, raw base calls are demultiplexed, converted to sequence reads in FASTQ format, and trimmed for quality. Each read is then aligned to the human genome reference and SNP base calls are assigned. SNP calls were first assigned using Qiagen/Verogen's recommended default thresholds: 3% analytical threshold (read count representation to type an allele), 3% interpretation threshold (read count representation of an allele to contribute to a call), and 50% intra-locus read count balance at heterozygous loci. Following initial review of allele call rates and locus read coverage variability, the analysis and interpretation thresholds were reduced to 1.5%, corresponding to a minimum locus depth of coverage of 10X. All samples were reanalyzed with these thresholds. Then, a report was generated containing SNP calls, total read coverage, and quality metrics for each SNP allele. The report also provides an estimation of biogeographical ancestry and phenotype prediction (e.g., hair and eye color). Finally, a GEDmatch-formatted genotype file was exported for upload into GEDmatch PRO.

For microarray analysis, Gene by Gene executed raw data analysis in GenomeStudio® Genotyping Module v2.0 [6] using default cluster files to cluster allele signal intensity and call SNP genotypes. They then provided GEDmatch-formatted genotype files in .csv file format. Up to 630,032 SNP genotypes were interrogated using GSAv2 BeadChips.

For genome sequencing analysis, Gene by Gene executed raw data analysis using an optimized Illumina Dynamic Read Analysis for Genomics (DRAGEN) analysis pipeline [7]. To begin, sequence reads were trimmed and filtered for quality and PCR duplicates were removed. Then, DRAGEN performed alignment to the hg38 human genome reference, variant calling, quality filtering, and post-processing reporting of all sequenced libraries. A mutation report in Variant Call Format (VCF) was generated and then converted to the GEDmatch-specific genotype .csv file to upload for database searching. Gene by Gene provided genome alignments in BAM format, variant analysis VCF data files, and down selected genealogy SNP genotypes in a GEDmatch-compatible format. Up to 2,061,275 SNP genotypes were interrogated using genome sequencing.

Comparative analyses of the final genotype files within and across technologies were performed at Bode using custom Excel macros and review templates, and statistical analyses were performed in JMP Statistical Discovery Software v15.2.1.

Sample Preparation for Degradation Analysis

To evaluate the effects of environmental degradation on semen samples and assess any correlation with Quantifiler Trio DI values, three semen samples (the two samples collected for Phase I Sensitivity analysis and a third semen sample purchased from Lee BioSolutions) were

exposed to known degradative processes (Table 2). Methods detailed in Nelson [8] for oxidative damage and hydrolytic damage/depurination required optimization including increasing buffer component concentrations and length of time for exposure. Damage via UV light exposure for increasing lengths of time was also tested on both whole semen aliquots and semen DNA extracts. Damage via depurination and UV exposure was determined to only affect semen DNA extracts, not whole semen aliquots.

Once optimal reaction conditions were identified, 20 µl aliquots of whole semen from the three unrelated male donors were extracted and artificially degraded via depurination through exposure to acidic conditions in a 10X concentration sodium citrate buffer (pH=4.8) and incubated at 70 °C for 12–48 hours. For degradation via UV light exposure, 20 µl aliquots of whole semen from three unrelated male donors were extracted and incubated in a UV (245 nm) crosslinker for 120–720 seconds. For degradation via oxidative damage, 20 µl aliquots of whole semen from three unrelated male donors were exposed to the Fenton reaction[1] by incubating the semen aliquot in a $FeCl_2$-EDTA (167.5 mM-337.5 mM) solution combined with 1.8 M H_2O_2 at 37 °C for 12–52 hours. Samples were then extracted (Oxidative damage set only) and purified via DNA Fast flow Microcons, quantified with Quant Trio to obtain DI values, and STR amplified to verify that profile degradation was observed. Profile balance and FI [9] values were calculated for each sample. Based on these results, a total of 36 extracts were identified for further genotyping analysis (Table 3). Samples were selected to fit a range of DI values and levels of observed STR profile degradation.

Table 2: Degraded semen samples examined with each genotyping technology

Genotyping Technology	Degradation Process	Degradation Timepoints	Replicates	Total Samples
BeadChip	Oxidation	4	3	12
	Hydrolysis – Fenton Reaction	4	3	12
	UV	4	3	12
WGS	Oxidation	4	3	12
	Hydrolysis – Fenton Reaction	4	3	12
	UV	4	3	12
Kintelligence	Oxidation	4	3	12
	Hydrolysis – Fenton Reaction	4	3	12
	UV	4	3	12

[1] Fenton reaction: bivalent iron (Fe^{2+}) reacts with H_2O_2 ($Fe^{2+} + H_2O_2 \rightarrow Fe^{3+} + \bullet OH + OH^-$). Hydroxyl radicals create DNA lesions by converting guanine to 8-oxoguanine.

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Table 3: Sample QC Metrics Following Artificial Degradation

Donor	Incubation Time	Degradation Method	DI Value	% Complete Profile	Profile Balance	FI Value
MD001	12hrs	Depurination	0.939394	100%	4.60	6.15
MD001	24hrs	Depurination	1.022624	100%	3.39	6.48
MD001	36hrs	Depurination	1.1843	100%	6.09	5.57
MD001	48hrs	Depurination	1.302857	100%	7.52	5.74
MD002	12hrs	Depurination	0.746951	100%	2.24	6.75
MD002	24hrs	Depurination	0.815642	100%	4.09	6.44
MD002	36hrs	Depurination	1.040073	100%	5.47	6.06
MD002	48hrs	Depurination	1.161905	100%	7.98	5.55
MD003	12hrs	Depurination	0.937198	100%	2.06	6.25
MD003	24hrs	Depurination	1.151515	100%	4.94	6.39
MD003	36hrs	Depurination	2.219697	100%	4.88	5.41
MD003	48hrs	Depurination	1.382008	100%	13.58	5.72
MD001	12hrs	Fenton Reaction	1.12782	100%	8.88	5.50
MD001	24hrs	Fenton Reaction	1.190476	91%	N/A	4.57
MD001	48hrs	Fenton Reaction	1.501984	96%	N/A	5.03
MD001	52hrs	Fenton Reaction	1.730983	96%	24.17	4.61
MD002	12hrs	Fenton Reaction	1.40146	100%	9.41	5.65
MD002	24hrs	Fenton Reaction	0.95104	94%	N/A	4.69
MD002	48hrs	Fenton Reaction	1.103053	100%	13.18	5.53
MD002	52hrs	Fenton Reaction	1.150862	96%	32.68	4.78
MD003	12hrs	Fenton Reaction	1.325359	100%	9.24	5.51
MD003	24hrs	Fenton Reaction	1.333333	94%	N/A	4.67
MD003	48hrs	Fenton Reaction	1.15678	70%	N/A	2.67
MD003	52hrs	Fenton Reaction	0.863049	70%	N/A	2.30
MD001	120s	UV exposure	2.111765	100%	5.79	6.06
MD001	360s	UV exposure	4.42953	91%	44.91	4.57
MD001	600s	UV exposure	9.183673	68%	N/A	2.72
MD001	720s	UV exposure	29.22374	57%	N/A	1.58
MD002	120s	UV exposure	1.696319	100%	4.60	6.12
MD002	360s	UV exposure	3.448276	98%	23.37	5.11
MD002	600s	UV exposure	10.13453	70%	N/A	2.43
MD002	720s	UV exposure	15.18182	68%	N/A	2.04
MD003	120s	UV exposure	1.735294	100%	5.84	6.20
MD003	360s	UV exposure	4.134276	89%	N/A	3.70
MD003	600s	UV exposure	10.8168	79%	N/A	2.83
MD003	720s	UV exposure	14.87395	66%	N/A	1.69

Introduction of electrophoretic evaluation of genomic DNA fragmentation after extraction

Quantifiler Trio is used to estimate DNA quality by calculating the ratio between quantification values obtained for large (>200 bp) and small (<80 bp) human targets. High DI values indicate amplification of larger fragments is impaired due to some degree of gDNA fragmentation; however, there is limited information correlating DI value with overall DNA fragmentation and how that impacts BeadChip hybridization. BeadChip hybridization has been shown to be particularly sensitive to degradation beginning at DI values that would still generate acceptable STR profiles [10,11], likely due to reduced efficacy of the polymerase used in an initial isothermal amplification step when DNA is fragmented to segments less than 2 Kb [12]. Thus, a method to evaluate DNA fragmentation prior to genotyping was incorporated. Following extraction and qPCR of the challenged DNA samples, 1 µl of each extract was run on the Agilent 2200 TapeStation with the Genomic DNA ScreenTape gel electrophoresis assay, which provides a measurement of gDNA integrity via a DNA Integrity Number (DIN). The integrated DIN

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

algorithm determines the fragmentation of a gDNA sample by assessing the distribution of signal across the size ranges and automatically calculates the DIN, which ranges from 1 to 10 [13]. The lower the DIN, the more degraded the gDNA sample.

Data Analysis of Artificially Degraded Semen Samples

SNP genotyping and data analysis of the degraded samples followed the methods of the sensitivity samples described above. A 2 ng DNA input for both genome sequencing and GSAv2 processing was selected based on the call rates obtained in the sensitivity analysis. All samples processed with Kintelligence were amplified with the manufacturer's recommended 1 ng DNA input. A total of 108 samples (36 samples per technology) were processed for this degradation analysis.

Comparative analyses of the final genotype files within and across technologies were performed at Bode using custom Excel macros and review templates, and statistical analyses were performed in JMP Statistical Discovery Software v15.2.1.

Genealogical Comparisons

For our genealogical assessment, all matching was performed against the GEDmatch database, with uploads in both GEDmatch Classic and GEDmatch PRO. Genotypes derived from genome sequencing and GSAv2 BeadChips for the sensitivity samples were uploaded through GEDmatch Classic as "Research" samples. Due to changes in the GEDmatch Terms of Service that no longer allow the use of GEDmatch Classic for forensic research, the degraded samples were uploaded through the PRO portal. One-to-Many Segment Based matching was performed for all "standard" (i.e., GSAv2- or Genome Sequencing-generated) and match lists were evaluated for all matches exceeding total shared centimorgans of 50 cM. All Kintelligence-generated genotypes were uploaded through the GEDmatch PRO portal as "Validation" samples. One-to-Many Kinship matching [14] was performed for all Kintelligence kits and match lists were evaluated for all matches exceeding the Expanded Match List thresholds (Table 4). The kit designations sequester test samples out of the database to maintain donor privacy but still match to the known relatives in the public database as well as to other Research kits. To assess potential false positive matches, One-to-One Q matching of standard kits and One-to-One Kinship matching with Kintelligence kits [14] was performed.

Table 4: GEDmatch PRO One-to-Many Kinship Match List Thresholds

High Confidence Matches Thresholds:		
shared cM	**Longest peak**	**SNP overlap**
170	30	9000
190	30	8000
200	30	6000
Expanded Matches Thresholds:		
shared cM	**Longest peak**	**SNP overlap**
120	30	9000
140	30	8000
160	30	6000

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

The goals of database comparisons were to confirm matching to the known relatives; observing additional matches with high total shared centimorgans (cM); and monitoring trends among metrics including total shared cM, longest segment lengths, number of overlapping SNPs, and whole genome kinship coefficients from Kintelligence comparisons [14].

Phase II

Sample Preparation of Challenged Mock Sexual Assault Samples

Vaginal fluid was collected from a single donor on sterile cotton tipped swabs under informed consent following IRB guidelines. A reference buccal swab was also collected from the donor, extracted, and submitted for reference genotyping on the GSAv2 BeadChip. Semen aliquots (20 µl) from Phase I donor MD001 were spotted onto the vaginal swabs, allowed to dry, and stored at 70 °C for 15 days. These storage conditions replicate room temperature storage of 390 days [15,16]. An additional constructed swab was stored at room temperature for 15 days. Upon removal from storage, the swabs were extracted following Bode's internally optimized differential extraction method that incorporates incubation with DNase I to remove epithelial cells from the sperm fraction. The sperm fraction was then extracted with the EZ1&2 DNA Investigator Kit. To induce additional degradation, the sperm fraction extracts were subjected to 110 seconds of UV exposure in a crosslinker. A second set of samples replicating buccal/saliva swabs was constructed by collecting three buccal swabs from an unrelated donor. Two swabs were incubated at 70 °C with one swab removed from incubation after 15 days and a second swab removed after 21 days, replicating room temperature storage for 390 and 547 days, respectively. The third swab was stored at room temperature for 21 days. The buccal swabs were extracted with the EZ1&2 DNA Investigator Kit, 500 µL lysis and eluted into 50 µL TE buffer. All extracts were then quantified with Quantifiler Trio and concentrated with Microcon DNA FastFlow devices. Extraction quality and degradation was confirmed with PowerPlex Fusion 6C™ (Promega Madison, WI) STR typing in 25 µl reactions prior to preparation for genotyping.

Data Analysis

SNP genotyping and data analysis of the degraded samples followed the methods described above for Phase I samples. The extracts were split to generate replicate samples for genotyping with BeadChip analysis, genome sequencing on NovaSeq 6000, and targeted resequencing using the Kintelligence kit. Table 5 describes the sample inputs used to generate genotyping results. Data

analyses to establish call rates and concordance with reference genotypes and GEDmatch comparisons were performed following previously described methods.

Table 5: Mock Sample Descriptions and DNA Inputs per Genotyping Technology

Sample Name	Sample Type	Degradation	DI Value	GSAv2 Genotyping DNA Input	Genome Sequencing Genotyping DNA Input	KINT Genotyping Total DNA Input
SAS_20_RT_SF	Sperm Fraction	None - Room Temp Control	0.7203	2 ng	2 ng	1 ng
SAS_20_15_SF	Sperm Fraction	High Heat 15 days - UV	6.211538	1.3 ng 0.5 ng	1.3 ng 0.5 ng	1 ng 0.5 ng
SAS_10_15_SF	Sperm Fraction	High Heat 15 days - UV	5.638298	0.4 ng	0.4 ng	0.4 ng
MD006_1	Buccal/saliva	None - Room Temp Control	2.439716	2 ng 0.5 ng	2 ng 0.5 ng	1 ng 0.5 ng
MD006_0626	Buccal/saliva	High Heat - 15 days	14.8913	2 ng 0.5 ng	2 ng 0.5 ng	1 ng 0.5 ng
MD006_0703	Buccal/saliva	High Heat - 21 days	61.35957	2 ng 0.5 ng	2 ng 0.5 ng	1 ng 0.5 ng

Genealogical Comparisons

Genealogical comparisons with Phase II mock samples were performed as described above. As the goal of this phase was to assess the range of accuracy in relationship matching, the buccal swab donor (MD006) was selected from the members of the autosomal DNA 5G Proof Project database (Figure 2), which is comprised of more than 100 family members, properly documented to five generations and genetically cross-checked, who have provided explicit written consent for genealogical studies using DNA evidence. Buccal swabs were collected from each member between 2017-2019 to generate FamilyTree DNA genotype kits. This database is privately held and administered to protect true identities, but genotype data are present and opted-in to law enforcement matching in the GEDmatch database. Donor MD006 sits in the third generation to allow assessment of both up generations and down generations. Kit matching was evaluated using One-to-Many Segment Based match lists and Kintelligence One-to-Many Kinship High Confidence, Expanded Matches, and All Matches match lists.

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

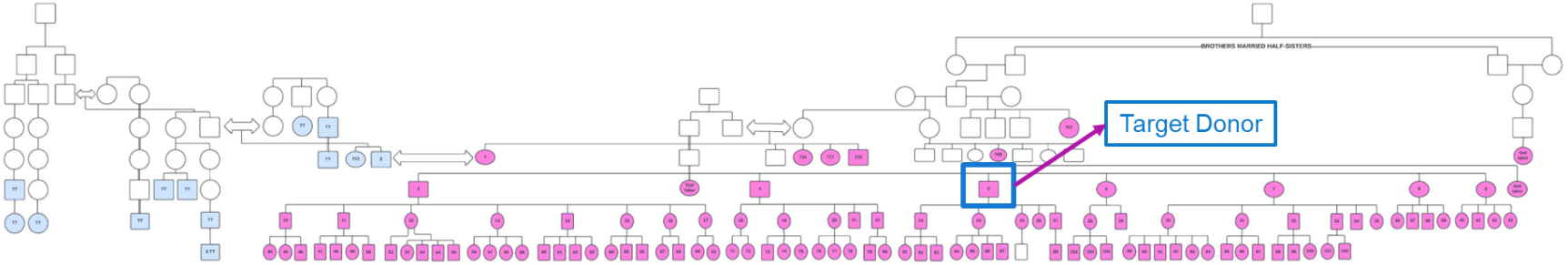

Figure 2: 5G Proof Project Database Family Tree. MD006 (no. 5, designated by arrow) and matching test takers with charted relationships; created 16 June 2024. Pink shapes are test takers from MD006's mother's side. Blue shapes are from MD006's father's side. On the maternal side brothers married half-sisters in the second of eight generations charted here. This was sufficiently distant that no significant endogamy was observed.

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Summary of Results

Phase I - Sensitivity Analysis

All three genotyping technologies were evaluated for the effects of decreasing DNA input based on three main performance metrics: SNP call rate, observed heterozygosity, and genotype concordance rate compared to the profile developed from a 200 ng sample using the Illumina GSA BeadChip. The sensitivity sample set was evaluated to establish input limitations where not only loss of call rate occurs but where genotype call accuracy is impacted.

Figure 3 compares sensitivity call rate and concordance metrics for all three genotyping technologies. Illumina GSAv2 BeadChip processing demonstrates genotyping call rates above 95% with DNA inputs as low as 250 pg with concordance rates >99% for all input amounts tested. Heterozygosity was calculated as a sample quality measure. Russell *et al* determined the average heterozygosity of GSAv2 chips to be ~17% [10]. No variation in heterozygosity was observed as DNA input decreased for the sensitivity samples processed on GSAv2 BeadChips. Thus, both the call rates and heterozygosity rates are in line with previous reports [10,17].

Kintelligence processing demonstrates genotyping call rates above 90% with DNA inputs as low as 100 pg while call rates drop to ~70% with the lowest input of 25 pg. Heterozygosity was again calculated as a sample quality measure. The expected heterozygosity for a sample of optimal input is ~45% [18]. Heterozygosity begins to drop below expected with inputs around 100 pg; at inputs of 25 pg, heterozygosity is < 20%.

Processing with genome sequencing also demonstrates genotyping call rates, of the more than 2 million SNPs interrogated, above 85% with DNA inputs as low as 500 pg. With inputs around 250 pg, the call rate drops below 60% but still produces more than 1 million SNP genotypes. The heterozygosity of sequenced samples was calculated using the autosomal SNPs that overlap the GSAv2 BeadChip for comparisons. A slight increase in average heterozygosity, up to 20%, was observed for the lowest DNA inputs tested.

Inter-technology comparison demonstrates high concordance rates across all inputs (Figure 3). Kintelligence and GSAv2 capture 9695 SNPs in common. When Kintelligence inputs were greater than 100 pg, >90% concordance to the 200 ng GSAv2 Reference genotypes was observed. With the lowest input of 25 pg, Kintelligence captured an average of 6985 common SNPs, corresponding to a concordance rate of 77.4%. Genome sequencing and GSAv2 capture 618,555 common SNPs. Call rates of the common SNPs mirror the call rates of the genome sequencing SNPs overall, >80% with inputs of at least 500 pg. Concordance rates at common SNPs are >98% between genome sequencing genotype calls and GSAv2 genotype calls.

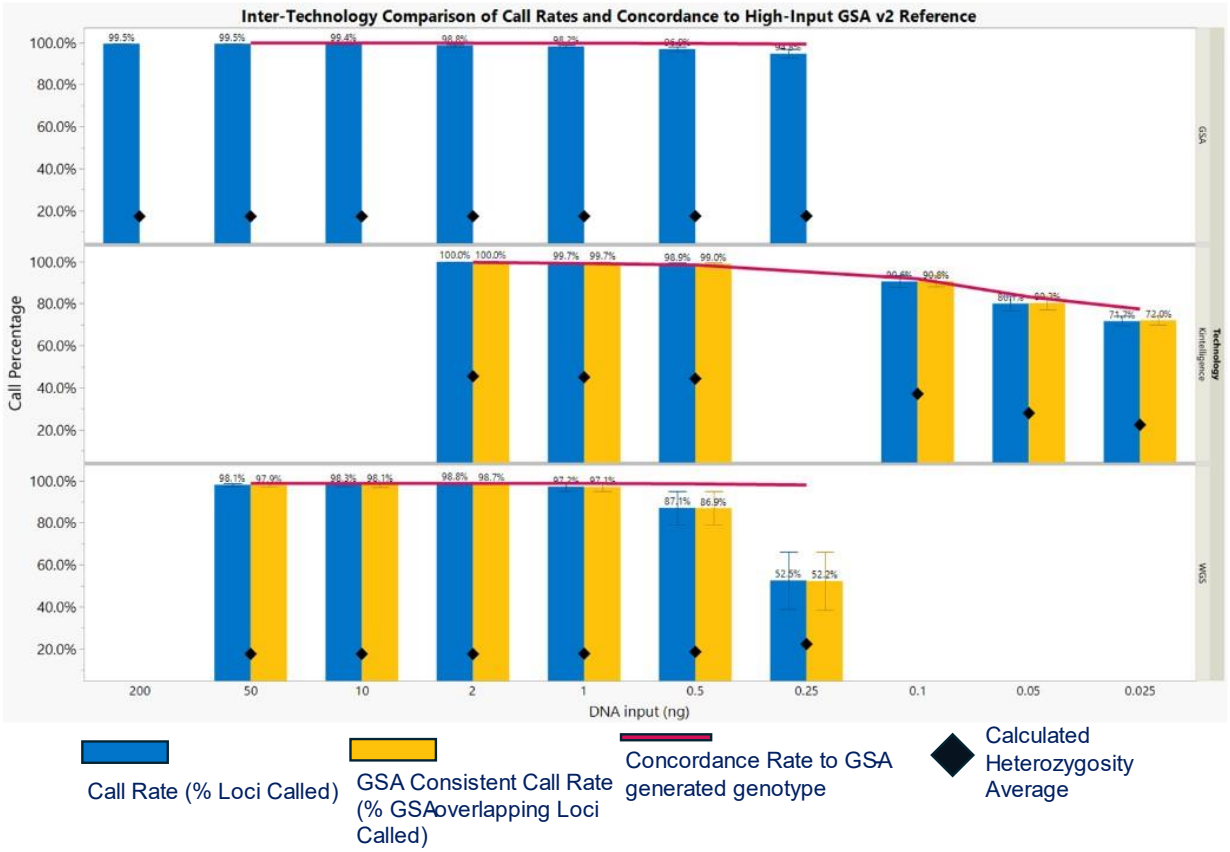

Figure 3: Comparion of Call rate metrics and concordance for sensitivity samples processed with three genotyping technologies. Top panel – Call rates, concordance, and heterozygosity with GSAv2 processing. Second panel – Call rates, concordance, and heterozygosity with Kintelligence processing. Bottom panel – Call rates, concordance, and heterozygosity with genome sequencing processing.

Discordant genotype calls relative to the 200 ng GSAv2 reference genotype for each donor were evaluated to determine the source of discordance in each technology (Figure 4). The largest percentages of GSAv2-generated discordant genotypes at DNA inputs less than 2 ng are due to false heterozygosity, or "allele drop-in" relative to the high DNA input reference profile. Genome sequencing discordance was largely due to false genotypes, which includes both opposite strand reporting between GSAv2 and genome sequencing (e.g., GA vs CT) and opposite homozygous calls (e.g., AA vs GG), until DNA inputs drop below 500 pg when an increase in false heterozygous calls was observed. The main cause of genotype discordance in Kintelligence is due to loss of heterozygosity: either the sister allele is not amplified or does not exceed the interpretation threshold read coverage.

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

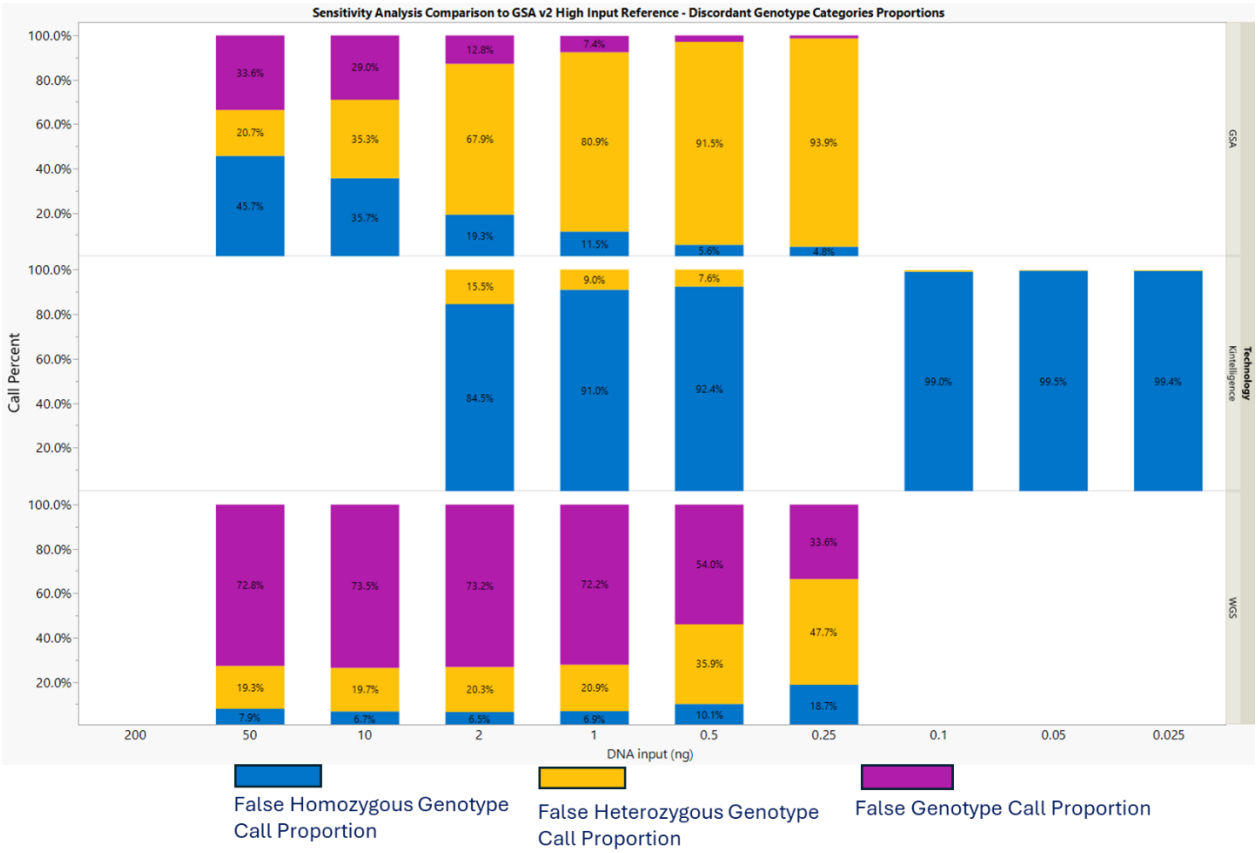

Figure 4: Comparion of discordant genotype category proportions for sensitivity samples processed with three genotyping technologies. Overall discordant rates were less than 2% of call for genome sequencing and GSAv2. Overall discordant rates ranged from <1% to 22% of call for Kintelligence. Top panel – GSAv2 processing. False genotype calls correspond to opposite homozygote calls. Second panel – Kintelligence processing. Bottom panel – genome sequencing processing. False genotype calls correspond to opposite homozygote calls and opposite strand reporting.

In addition to call rates and concordance, overall sequencing metrics were captured for both Kintelligence processing and genome sequencing. These metrics provide an overview of the quality of the data generation, which was generally high-quality. However, reductions in average locus read depth (Kintelligence processing) and average autosomal read depth (genome sequencing processing) were observed with decreasing DNA input, which correspond to the observed reductions in call rates. These metrics have been compiled into a data file that has been made publicly available at the completion of this project [19].

Summary Results of Degradation Analysis

Quantifiler Trio DI values were compared to quantitative metrics obtained for the genotype data to assess if a DI value cut-off level can be determined where each genotyping technology results in unusable SNP genotypes. DI values were only impacted by UV irradiation, which demonstrated no effect on genomic fragmentation. Figure 5 compares the average Quantifiler Trio DI values and the average DIN values observed for each degradation exposure time point.

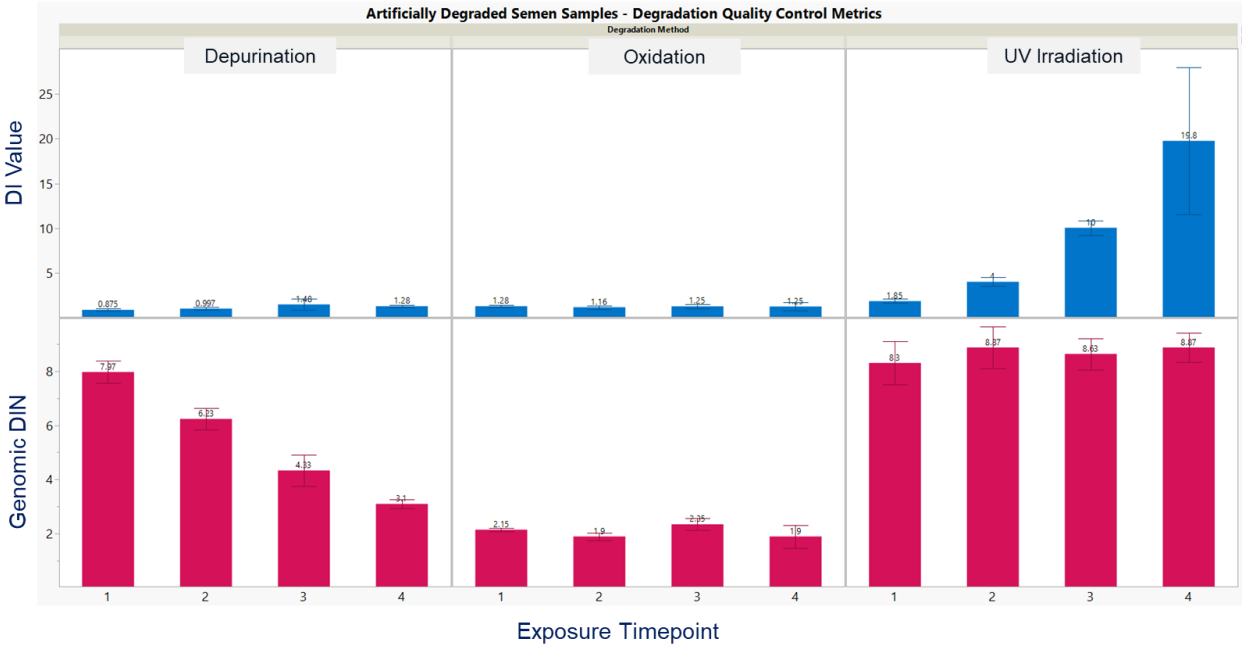

Figure 5: Comparison of semen DNA extract quality following arificial degradation using Quantifiler Trio Degradation Index (DI) (top panels) vs. Agilent TapeStation Genomic ScreenTape DIN values (bottom panels). Depurination time points 12, 24, 36, 48 hours. Oxidation via Fenton Reaction time points 12, 24, 48, 52 hours. UV incubation time points 120, 360, 600, 720 seconds. Oxidation DINs are determined using only 2 of 3 donors tested due to sample concentrations falling below the minium 10ng/ul required to visualize with the Genomic ScreenTape. Error bars are one standard deviation from the mean.

Artificially degraded samples processed with all three genotyping technologies were again evaluated for three main performance metrics: SNP call rate, observed heterozygosity, and genotype concordance rate compared to the profile developed from a 200 ng donor reference sample using the Illumina GSAv2 BeadChip.

Variation in genotyping metrics was observed when degraded samples were processed with the GSAv2 BeadChips (Figure 6). Over time, significant impacts to the call rates were observed for all three degradation methods. Depurination ($p=0.0040$) and oxidation induced by the Fenton reaction ($p=0.0159$) produced statistically significant decreases in call rate with increasing exposure time that correspond to the decreases in observed DIN values; however, the call rates remained high at >85% and concordance exceeded 95% across timepoints. These degradation methods demonstrated increasing sample fragmentation but no impact to Quant Trio DI values. UV irradiation demonstrated the greatest reduction in call rates with increasing time ($p<0.0010$) with call rate reductions corresponding to the increasing DI values. The impact to concordance to reference mirrors the impacts to call rates. The most degraded samples demonstrated average call rates down to 40.6% with approximately 70% concordance. Heterozygosity also demonstrated a marked increase with exposure time for UV irradiated samples.

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Figure 6: Call rate and concordance metrics for artificially degraded semen samples processed with GSAv2 BeadChips correspond to DI values. Top panels – Semen extract Quantifiler Trio Degradation Index values. Error bars are one standard deviation from the mean. Bottom panels – Genotyping metrics.

Figure 7 details the observed causes of discordant genotypes for each degradation method across exposure timepoints. The majority of discordance, regardless of degradation method, was due to false heterozygous calls, specifically GG to AG and CC to TC calls. Slight increases in false homozygous calls were observed in the oxidation degraded samples with the lowest call rates. UV irradiation demonstrated increases in both false homozygous calls and false genotype calls (mostly characterized by opposite homozygous transition calls, e.g., CC to TT) over time, corresponding to increasing DI value.

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

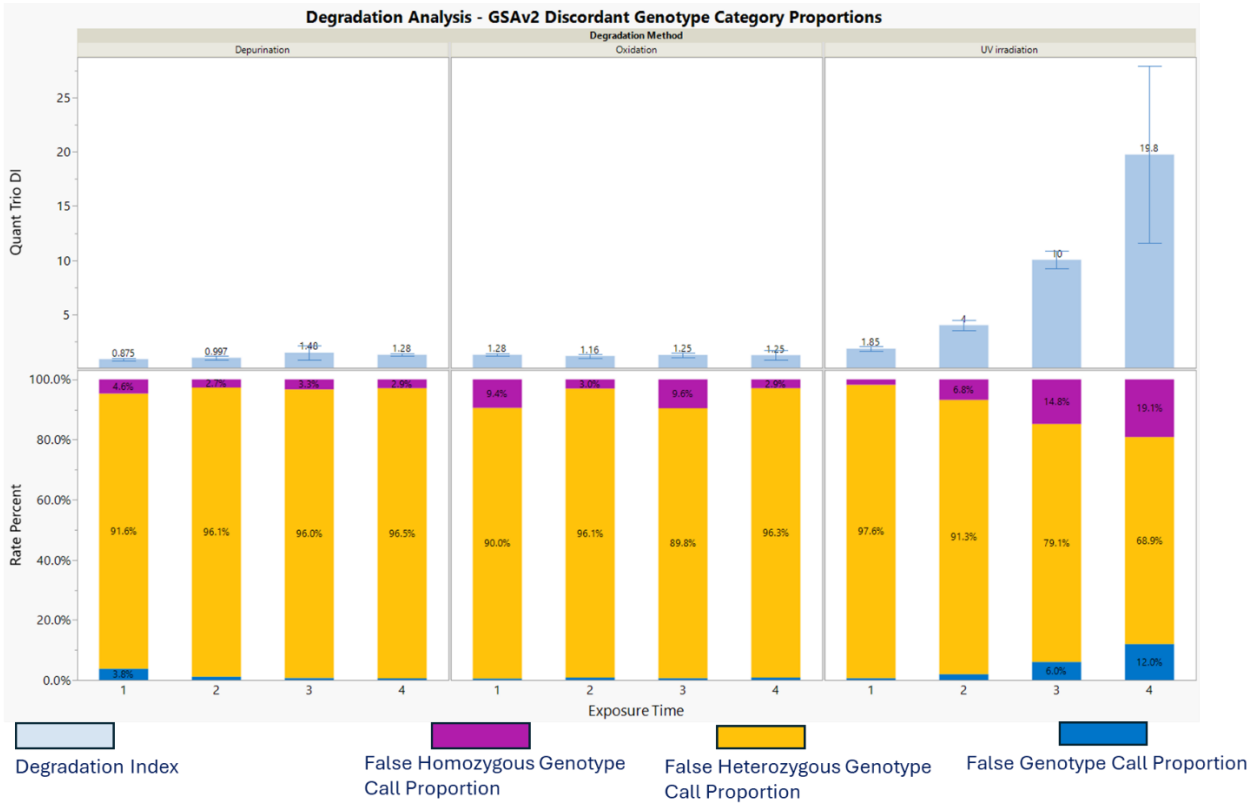

Figure 7: Discordant genotype category proportions for artificially degraded semen samples processed with GSAv2 BeadChips correspond to DI values. Overall discordant rates were up to 30% of total calls. Top panels – Semen extract Quantifiler Trio Degradation Index values. Error bars are one standard deviation from the mean. Bottom panels – Discordant category proportions. Discordant genotypes determined by comparison to donor reference genotypes derived from 200 ng DNA inputs on GSAv2 BeadChips.

When processed with genome sequencing, most samples generated call rates >92% (Figure 8) except those degraded with UV irradiation, which resulted in a statistically significant ($p=0.0054$) decrease in call rate relative to undegraded controls. However, these degradation levels still genotyped more than 1 million SNPs with concordance >99% to a 50 ng non-degraded WGS genotype. The call rates and concordance rates for called SNPs common to the GSAv2 BeadChip were also greater than 92% and 98% respectively when compared to a 200 ng GSAv2 reference genotype except those degraded with UV irradiation for longer than 600 seconds. Additional sequencing quality metrics were reviewed to assess the overall quality of the sequencing data including Q-scores, duplication rates, average autosomal read depth, and heterozygosity of GSAv2 overlapping SNPs (the latter two metrics are displayed in Figure 8, other metrics are compiled in available dataset [19]). Average Q-scores were greater than 30, indicating high quality sequence data, but other quality metrics demonstrated variation consistent with the reduction in call rates relative to increased DI degradation (e.g., duplication rates increase while depth of coverage decreases relative to poorer quality DNA input).

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Figure 8: Call rate and concordance metrics for artificially degraded semen samples processed with genome sequencing correspond to DI values. Top panels – Semen extract Quantifiler Trio Degradation Index values. Error bars are one standard deviation from the mean. Middle panels – Average autosomal depth of coverage across treatment methods. Expected Depth of Coverage was ~30X. Bottom panels – Genotyping metrics.

Figure 9 details the observed causes of discordant genotypes in genome sequencing results relative to the GSAv2 200ng Reference genotypes for each degradation method across exposure timepoints. The majority of discordance, regardless of degradation method, was loci that were not called with GSAv2 processing. The second largest proportion of discordant loci was driven by "false genotype" calls, or homozygous transitions. For example, a homozygous AA was called with BeadChips, but a homozygous GG was called with genome sequencing. An increase in false homozygous calls was observed with increasing exposure in the UV irradiated samples, again corresponding to increasing DI value.

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

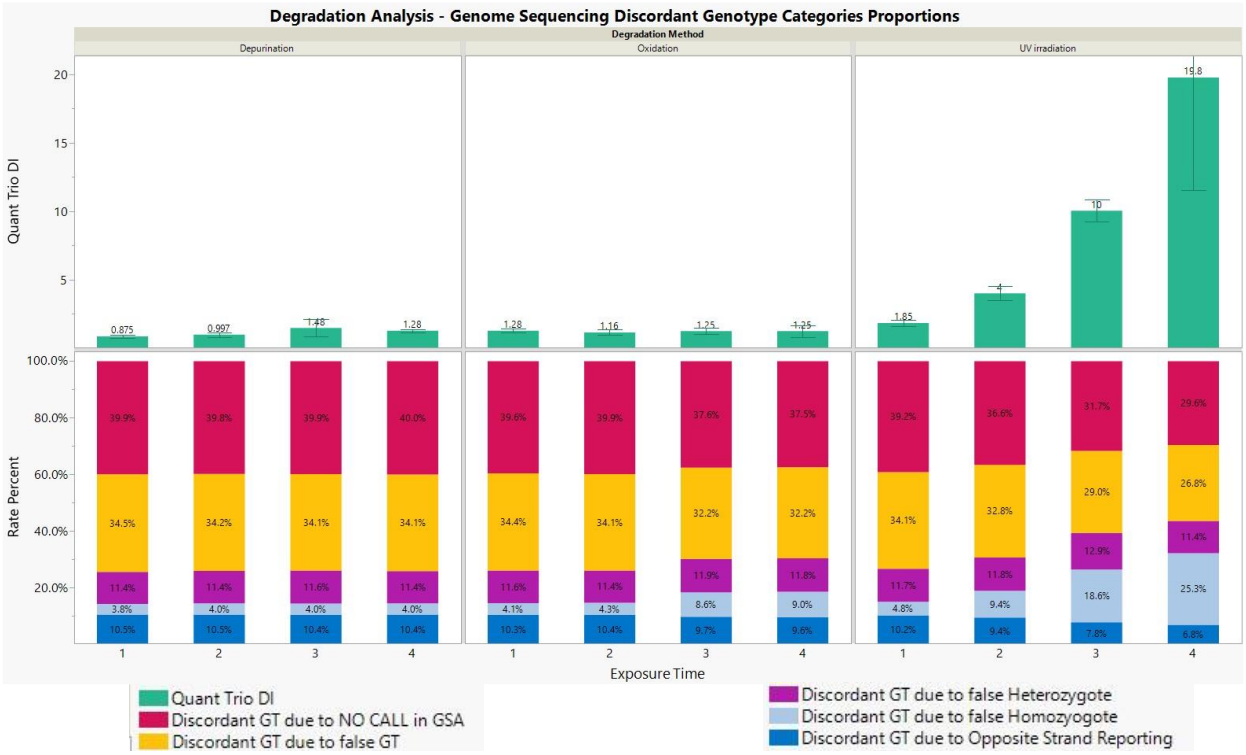

Figure 9: Discordant genotype category proportions for artificially degraded semen samples processed with genome sequencing correspond to DI values. Overall discordant rates were less than 2% of calls. Top panels – Semen extract Quantifiler Trio Degradation Index values. Error bars are one standard deviation from the mean. Bottom panels – Discordant category proportions. Discordant genotypes determined by comparison to donor reference genotypes derived from 200 ng DNA inputs on GSAv2 BeadChips.

When processed with Kintelligence (Figure 10), only UV irradiation indicated a statistically significant decrease (p=0.003) in call rate relative the undegraded control in response to increased DI and reduced average locus read depth; however, the minimum call rate regardless of degradation treatment was 93%. There was no significant difference between heterozygosity rates observed after *in vitro* degradation and the undegraded 1 ng control. The concordance rates for called SNPs were also greater than 96%. As with the sensitivity series, the majority of discordant genotypes were due to false homozygous genotypes as a result of allele drop out.

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

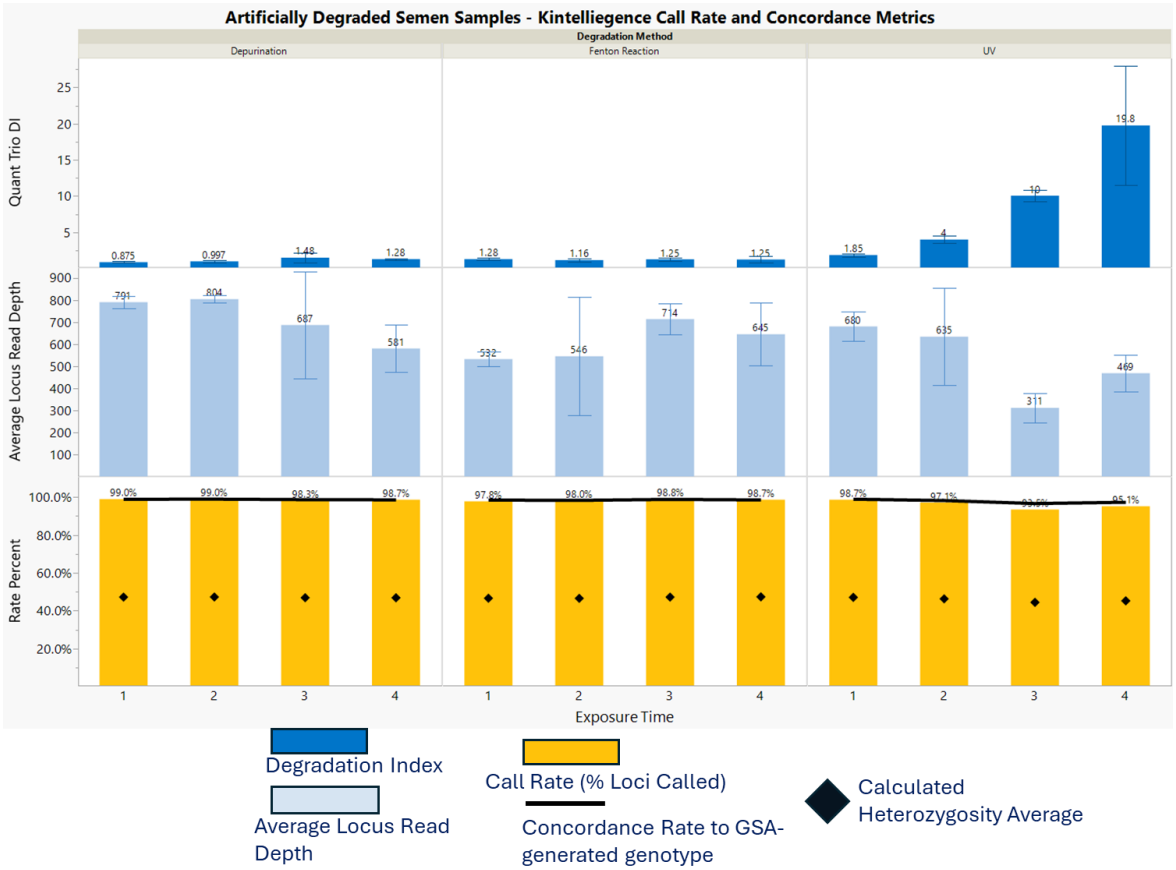

Figure 10: Call rate and concordance metrics for artificially degraded semen samples processed with Kintelligence compared to DI values. Top panels – Semen extract Quantifiler Trio Degradation Index values. Error bars are one standard deviation from the mean. Middle panels – Average locus read depth across treatment methods. Error bars are one standard deviation from the mean. Bottom panels – Genotyping metrics.

GEDmatch PRO Searching for Genealogical Comparisons – Sensitivity Samples

One-to-Many Segment Based match lists were reviewed for standard kits (GSAv2- and genome sequencing-derived) and all matches with >50 cM total shared content were compiled across DNA inputs. All but one kit for MD001 (genome sequencing derived kit with an initial input of 250 pg, deemed "too matchy") successfully matched against the GEDmatch database. The following trends were observed:

- When known relatives were available for matching, the known relative kits were consistently the top matches, regardless of DNA input, in both GSAv2- and genome sequencing-derived test sample kits. There was minimal to no variation in total shared cM for relationships out to known 2nd cousins (Figure 11 and Figure 12).
- GSAv2-derived kits returned more matches with >50 cM total shared content than genome sequencing-derived kits. Many of the same kit IDs were returned in both match lists; however, the genome sequencing matches generally showed lower total shared cM values, likely due to more specific comparisons afforded by increased numbers of overlapping SNPs. Kit IDs that no longer appeared in genome sequencing-derived kit match lists often had longest segments of <20cM when matching to GSAv2-derived kits. Anonymized match list comparisons are provided as part of the project data package [19].

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

- No false positive matches (Kit ID matching >100 cM shared content with no known relationship to the donor) were observed with either technology as DNA input decreased.
- The donor of non-European ancestry, RM8393, generated highly disparate match lists between GSAv2- and genome sequencing-derived kits. In GSAv2-derived kits, no more than 7 matches were identified sharing >50 cM. The highest match was between 90 and 114 cM with low SNP overlap (<45,000 SNPs). Genome sequencing-derived kits matched with a maximum total shared cM of 37 cM for inputs between 50 ng and 500 pg with SNP overlaps around 51,000 to 53,000. None of the highest matching kits in GSAv2-derived kit match lists were retained in the genome sequencing-derived kit match lists.
- GSAv2-derived and genome sequencing-derived kits for donors M001 and M002 were analyzed by Eurogenes K13 and RM8393 was analyzed by HarappaWorld database to predict the admixture of each donor. There was no effect on admixture prediction between technologies or with decreasing DNA input; the same percentage of the contributing population group was predicted for each test kit from a given donor.

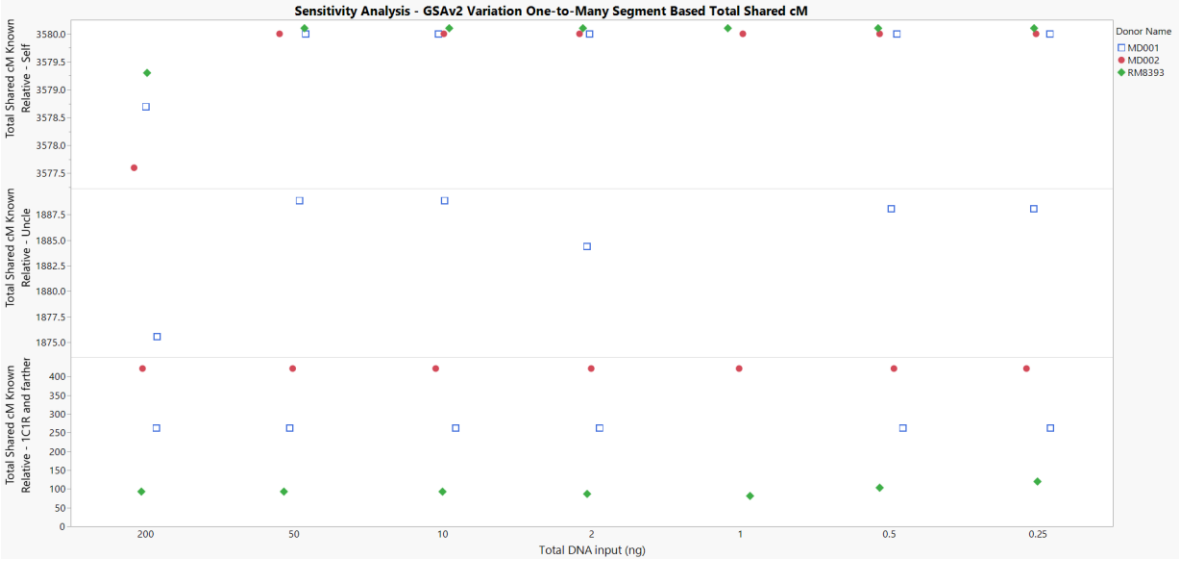

Figure 11: Variation observed in total shared cM for each relationship of each sensitivity donor, GSAv2 derived kits. Top panel = Shared cM when matching to self-reference using 200ng GSAv2 Reference kit. Note self-matching for 200 ng samples performed to 50 ng input genome sequencing-derived kits. Minimal to no variation in shared content with self-reference kits. Middle panel = Shared cM when matching to known relative database kit at 1st degree relationship. Only MD001 has relative in this category. Minimal variation present but does not alter relationship probability. Bottom panel = Shared cM when matching to known relative database kits at further cousin relationships. MD001 and MD002 have a relative at this distance. RM8393 top match shown here but unknown to be true relative.

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

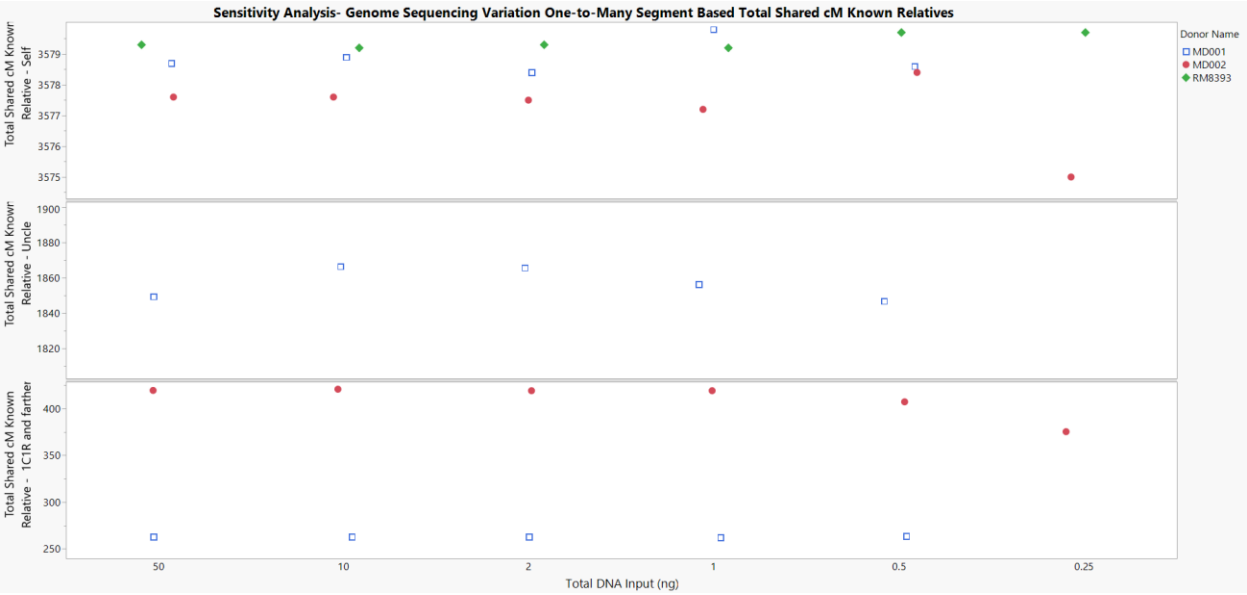

Figure 12: Variation observed in total shared cM for each relationship of each sensitivity donor using genome sequencing-derived kits. Top panels = Shared cM when matching to self-reference using 200ng GSAv2 Reference kit. Middle panels = Shared cM when matching to known relative database kit at 1st degree relationship. Only MD001 has relative in this category. Slight variation present, but does not alter relationship probability. Bottom panels = Shared cM when matching to known relative database kits at further cousin relationships. MD001 and MD002 have a relative at this distance. RM8393 not included because top match unknown to be true relative and matched with total shared cM <40 cM. Again, a slight drop in shared content observed for MD002 but does not alter relationship probability.

One-to-Many Kinship match lists were reviewed for Kintelligence kits and all matches exceeding Expanded Match list thresholds were compiled across DNA inputs. All SNP genotype profiles produced with Kintelligence were successfully compared against the GEDmatch database. The following trends were observed:

- When genotypes were generated with DNA input of at least 100 pg, true matches out to the 5th Degree (2nd cousin) relationship (>200 shared cM) were identified.
- Known 4th Degree (1C1R) relatives could still be observed in match lists when inputs were at or below 50 pg; however, the degree of relationship was determined as more distant and the match exhibits less statistical confidence (Figure 13).

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

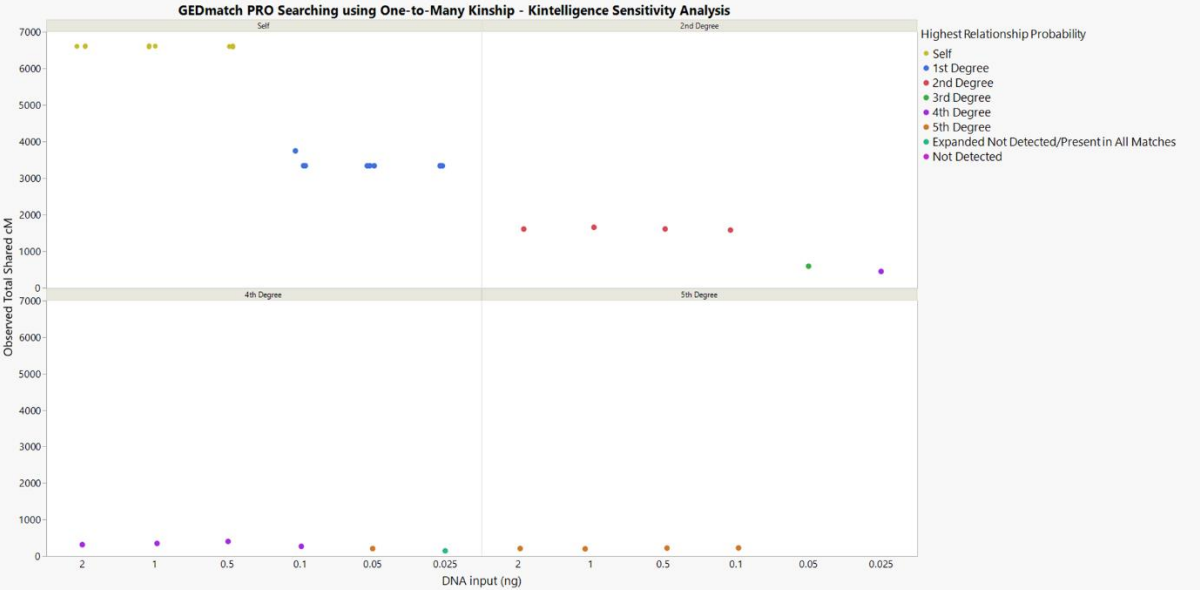

Figure 13: Total shared cM values generated One-to-Many Kinship matching comparing each Sensitivity test sample Kintelligence kit to the known relative databased kits. Expected relationships are: Self (top left), 2nd Degree or 1st Cousin (top right), 4th Degree or 1C1R (bottom left), and 5th degree or 2nd Cousin (bottom right). Data points are colored according to the GEDmatch PRO generation chart's likely relationship based on observed total shared cM. No match observed to the expected 5th degree relationship with inputs of 50 pg and 25 pg.

- At 100 pg inputs, additional high confidence false positive matches (>200 shared cM) were identified.
- False positive matches (>120 cM) were identified in the Expanded match lists for all donors with all DNA inputs (Figure 14). Relationship likelihoods placed the matches as 2nd cousin or farther. Whole genome kinship coefficients were calculated for all additional matches using One-to-One Kinship Matching, and any match with a coefficient >0.01 (the threshold observed for true 2nd Cousin matching) was compared to the donors' GSAv2 Reference Kit ID in One-to-One Q matching. Out of 69 kit comparisons, only nine matches demonstrated shared content with the standard kits; however, no more than one matching segment with a max shared content of 9.61 cM was identified.

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

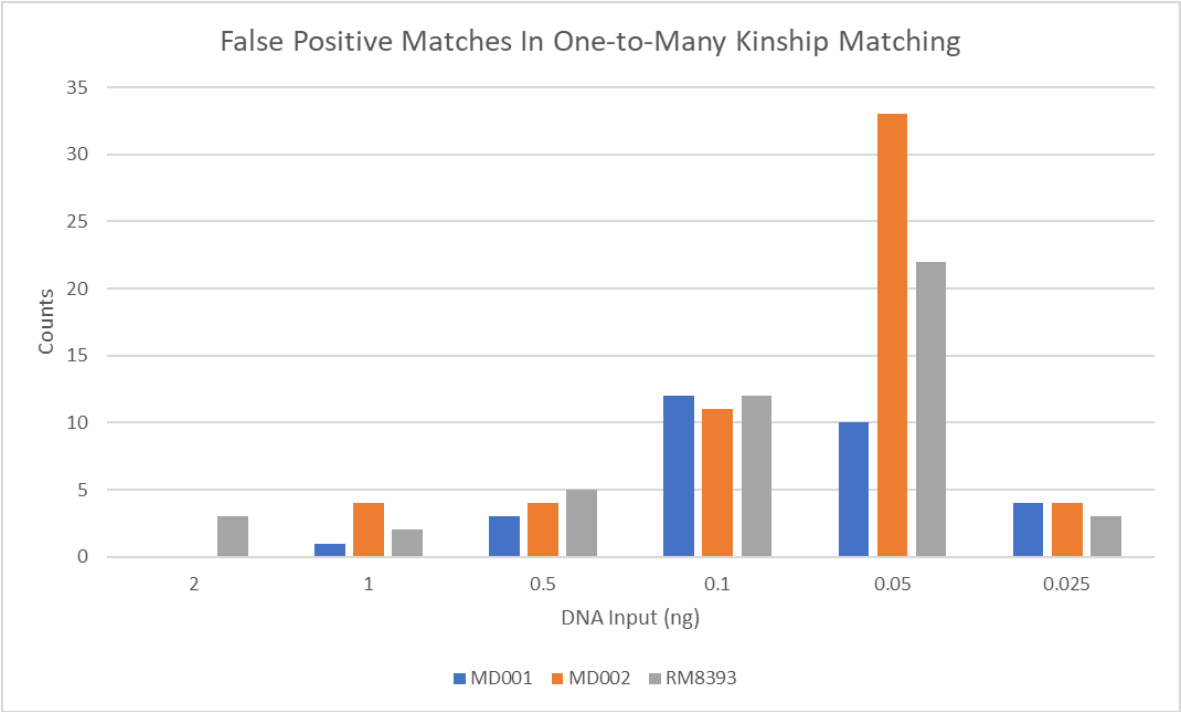

Figure 14: Total counts of "false positive" matches to Kintelligence sensitivity data. These matches were observed to exceed the 6000 SNP-180 cM/ 9000 SNP-12cM with 30 cM longest stretch threshold in One-to-Many Kinship windowed matching.

- None of the matches exceeding the Expanded match list thresholds were observed in the standard kit match lists for sensitivity samples.
- Results are consistent with recommendations by Radecke *et al* [20], that disclose matches in the 4th and 5th degree range in the expanded match list are unreliable.
- High confidence matches were identified for RM8393 donor at inputs of 500 pg, 100 pg, and 50 pg at likely relationships of Half GG-Aunt / Uncle, 2C, Half 1C1R, 1C2R, or Half GG-Niece / Nephew. One-to-One Q matching between RM8393's standard reference kit and the high confidence matching kits indicate no real relationship.

GEDmatch Searching for Kinship Matching – Degradation Samples

One-to-Many Segment Based match lists were reviewed for standard kits (GSAv2- and genome sequencing-derived) and all matches with >50 cM total shared content were compiled across degradation treatments. All but one kit for MD002 (GSAv2-derived kit exposed to 360 seconds of UV irradiation, deemed "too matchy") successfully matched against the GEDmatch database. The following trends were observed:

- All degraded samples processed with genome sequencing consistently matched to known relatives out to 2nd cousins, and known relatives were consistently the top matches. The minimal loss in total shared cM observed with oxidized and UV irradiated treatments had no effect on the relationship estimates.
- Depurinated and oxidized test samples processed with GSAv2 consistently matched to known relatives out to 2nd cousins, and known relatives were consistently the top matches (Figure 15). The UV irradiated test samples lost the ability to match to 2nd cousins when

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

the DI >4. Also, total shared cM values decreased for the closest relationships to the point where generation estimates increased, suggesting more distant relationships (Figure 15).

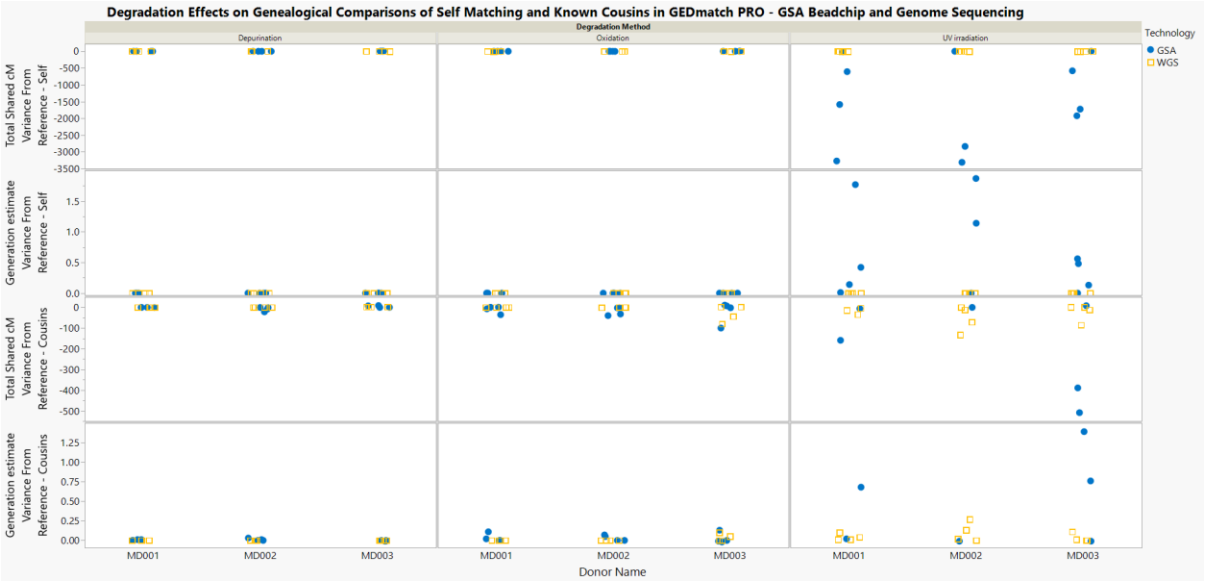

Figure 15: Degradation effects on genealogical matching in GEDmatch PRO. Observed variance in One-to-Many Segment Based total shared cM from expected shared cM for the undegraded donor reference matching for each sample for "Self" matching (Top panels). Observed variance in Generation estimate from expected generation for the undegraded donor reference matching for each sample for "Self" matching (Second panels). Observed variance in One-to-Many Segment Based total shared cM from expected shared cM for the undegraded donor reference matching for each sample for cousin matching (Third panels). Observed variance in Generation estimate from expected generation for the undegraded donor reference matching for each sample for cousin matching (Bottom panels).

- More matches in general were obtained with >50 cM total shared when comparing GSA-derived kits of Degraded test samples than with GSAv2-derived kits from the Sensitivity test samples. Matching metrics were consistent for reoccurring Kit IDs (Kit IDs observed matching to more than one test sample), but kits sporadically matched with more shared cM.
- No false positive matches (Kit ID matching at >100 cM shared content with no known relationship to the donor) were observed with either technology across degradation treatments.

One-to-Many Kinship match lists were reviewed for Kintelligence kits and all matches exceeding Expanded Match list thresholds were compiled across degradation treatments. All SNP genotype profiles produced with Kintelligence were successfully compared against the GEDmatch database. The following trends were observed:

- Regardless of degradation method, all three donors matched to their known relatives with high confidence. The shared cM values at a given relationship were consistent for nearly all comparisons across degradation methods and time points. Only two comparisons to a 4th degree relative of UV degraded samples resulted in a drop in shared cM values that would indicate a more distant relationship (Figure 16).

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

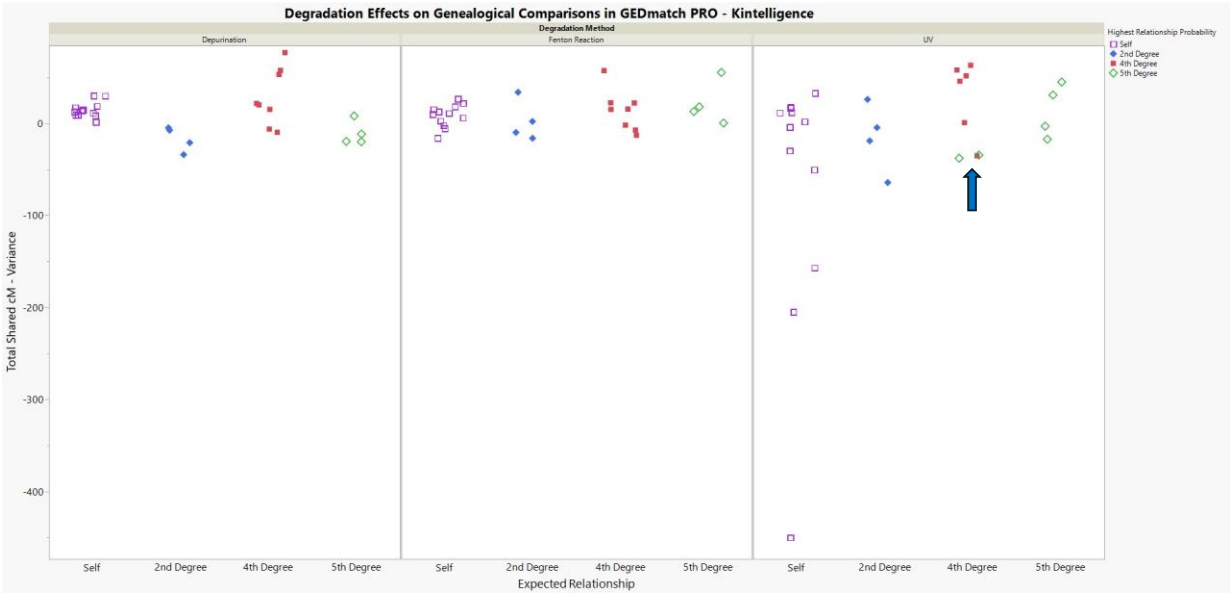

Figure 16: Observed variance in One-to-Many Kinship total shared cM from expected shared cM for the undegraded donor reference matching for each sample, color-coded by the determined relationship degree. Arrow indicates the two matches to a 4th degree (1C1R) relative with total shared cM that correspond a more distant relationship.

- For the purchased semen sample donor, a second possible relative was matched with high confidence at a 5th degree or 2C level in 2 out of 12 comparisons, but this same kit ID was present in the expanded match lists for all other comparisons, with shared cM values in the range of 120-150 cM. Additionally, this kit was observed as a top match (170-180 cM total shared cM) in the match lists generated using GSAv2 and genome sequencing genotype files. A third match in the range of 180-190 cM with GSAv2- and genome sequencing-derived kit comparisons was NOT observed in the One-to-Many Kinship match lists for this donor.
- An increase in the number of non-known relative matches exceeding Expanded Match thresholds was observed as sample degradation increased (Figure 17). Across 36 test samples uploaded for comparison to GEDmatch PRO, 157 additional matches with less than 200 cM (5th degree relationships and farther) were returned. To assess the likelihood that these additional matches were true positive matches, comparisons in One-to-One Q matching were performed against the donor reference kits generated with GSAv2 genotyping. Only three matches demonstrated shared content with the respective donor reference. Two matches to MD001 shared 11.43 cM and 7.22 cM with 1 segment each. With One-to-Many Kinship, these matches share total 168.3 cM (7571 SNPs overlap) and total 127.5 cM (9526 SNPs overlap), respectively. One additional kit matched to two MD003 test samples shared 13.91 cM in a single segment with the MD003 GSA reference kit. With One-to-Many Kinship, this match shares max 122.6 cM (9529 max SNPs overlap). One-to-One Q matching comparisons indicates that these additional matches past 5th degree relationship are likely false positive matches to the Kintelligence generated genotype kits.

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

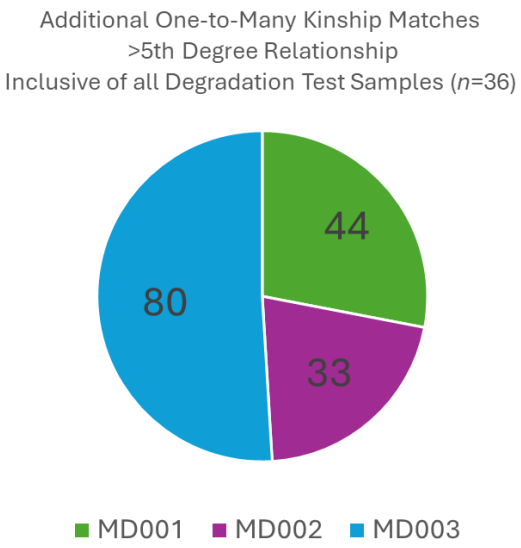

Figure 17: Counts of matches beyond known relative per donor observed in One-to-Many Kinship Matching, exceeding Expanded Match List thresholds, for degraded test samples processed with Kintelligence.

PHASE II

Two mock sample sets were constructed for Phase II to more closely replicate sexual assault evidence. First, a set of vaginal cell/semen mixture swabs was subjected to high temperature incubation and UV irradiation to induce degradation to the sperm fraction. The second set consisted of buccal/saliva swabs collected from a known donor and subjected to accelerated aging through high temperature incubation. These samples, now exhibiting varying levels of degradation, were then processed through all three genotyping technologies with DNA inputs ranging from 2 ng to 400 pg.

Genotyping Results – Mock Vaginal/Semen Swab Sperm Fraction

Sperm fraction DI values are shown in Table 5. Quantifiler Trio results also indicated male fractions <100%, and subsequent STR typing confirmed the presence of low-level female contribution to the sperm fraction for two of the extracts. Extract SAS-20-RT-SF indicated an ~14% contribution from the female vaginal cell donor and extract SAS-20-15-SF indicated an ~9% minor female contribution. The third extract showed no evidence of the female contributor.

As in Phase I, mock samples were evaluated for three main performance metrics: SNP call rate, observed heterozygosity, and genotype concordance rate compared to the profile developed from a 200 ng donor reference sample using the Illumina GSAv2 BeadChip (Figure 18). For the mock samples, all metrics were impacted relative to undegraded, high-quality samples of similar input. With GSAv2 processing (Figure 18A), the test samples with DI >5 produced call rates around 30% with concordance <70% and elevated heterozygosity. Only the SAS-20-15-RT sample with low DI produced a call rate >80%; however, heterozygosity was still elevated about 20% due to the presence of the minor contributor. Genome sequencing of the sperm fractions also resulted in reduced call rates (Figure 18B); however, as more than 2 million SNPs are interrogated, these call rates range from >493,000 SNPs (23% call rate, SAS-20-15-SF) to >1.95 million SNPs (94.8% call rate, SAS-20-15-RT). Heterozygosity was slightly elevated to ~20% across samples, but

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

concordance to the sperm fraction donor reference exceeded 95% for all samples. Call rates with Kintelligence (Figure 18C) exceeded 85% but the presence of the female contributor in SAS-20-15-RT and both SAS-20-15-SF inputs resulted in elevated heterozygosity and concordance reductions to less than 70%. Sample SAS-10-15-SF, despite the low DNA input of 400 pg, produced a call rate of 88% with ~90% concordance and heterozygosity in the expected range at 42%.

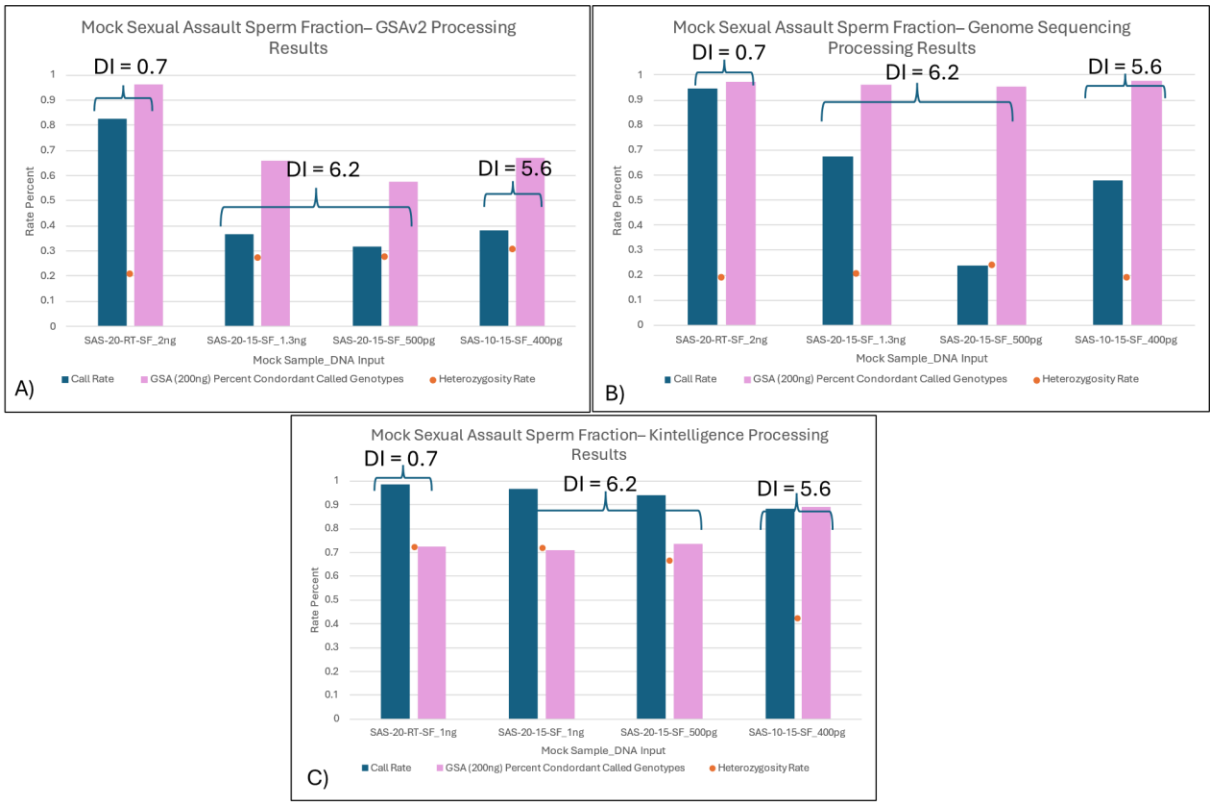

Figure 18: Comparion of Call rate metrics and concordance for mock sexual assault sperm fractions processed with three genotyping technologies. A) – Call rates, concordance, and heterozygosity with GSAv2 processing. B) – Call rates, concordance, and heterozygosity with genome sequencing processing. C) – Call rates, concordance, and heterozygosity with Kintelligence processing.

Genealogical Searching Results – Mock Vaginal/Semen Swab Sperm Fraction

The impacts to GEDmatch searching are shown in Figure 19, which compares the observed total shared cM for each of the expected donor relatives, self-matching to the 200 ng GSAv2 donor reference kit and matching to the minor female contributor's donor reference kit. The GSAv2-derived kit comparisons were unusable (Figure 19A). The kit for SAS-20-15-RT-SF was uploaded but was deemed "too matchy". For the remaining three sample kits, no known relatives were identified, and self-matching was inaccurate. Three of the four genome sequencing-derived kits successfully matched to the GEDmatch database (Figure 19B); SAS-20-15-RT-SF with an input of 500 pg was deemed "too matchy". For all three kits, accurate self-matching was observed and the top matches in the One-to-Many Segment Based match list were the donor's two known relatives. Total shared cM for each match were slightly reduced relative to the expected values, but the reduced shared content did not affect the relationship estimations. Additionally, the minor female contributor did not match to any kit.

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

One-to-Many Kinship matching of Kintelligence kits produced inaccurate results (Figure 19C). Self-matching of all four kits indicated a more distant relationship. Only the SAS-10-15-SF sample, which did not indicate a mixture genotype, matched both known relatives with total shared cM values near expected. The remaining three samples matched to the known 1st cousin relative, but with lower-than-expected total shared cM and also matched to the minor female contributor's reference kit at a 5th degree (2nd cousin) relationship.

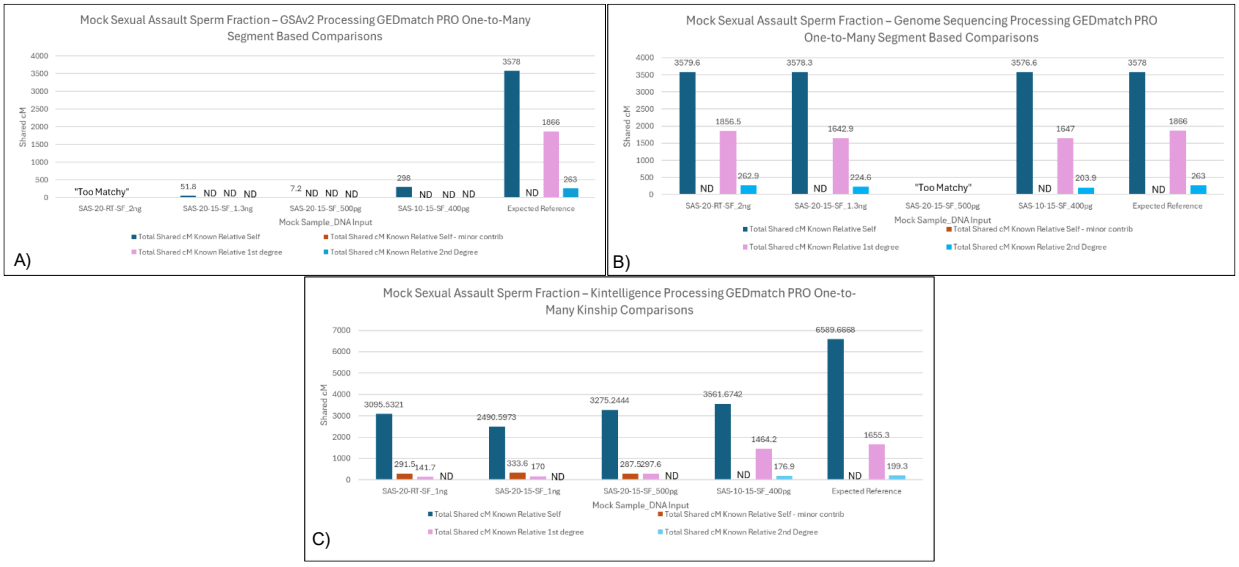

Figure 19: Variation observed in total shared cM for expected relationship matching of each sperm fraction test sample. A) One-to-Many Segment Based Matching results for uploaded GSAv2-derived kits. B) One-to-Many Segment Based Matching results for uploaded genome sequencing-derived kits. C) One-to-Many Kinship Matching results for uploaded Kintelligence-derived kits. ND = Not detected, no match identified to kit donor. The expected total shared cM values are included on each graph for self-matching and the two expected realtive kit matches.

Genotyping Results – Buccal/Saliva Swabs

SNP call rates, observed heterozygosity, and genotype concordance rate compared to the profile developed from a 200 ng donor reference sample using the Illumina GSAv2 BeadChip are shown for the mock buccal/saliva swabs in Table 6. GSAv2 call rates rapidly declined as DI value increased and the total DNA input decreased. No call rate exceeded ~75% and all samples demonstrated poor concordance rates. However, similarly to mock sperm fraction extracts, genome sequencing and Kintelligence results were of better quality. The most degraded extract (DI=61.4) with only 500 pg inputs produced ~70 % call rate (>1 million SNP genotypes) with 99%

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

concordance when processed with genome sequencing and a nearly 90% call rate with ~90% concordance when processed with Kintelligence.

Table 6: Genotyping metrics obtained for mock saliva swabs

		GSAv2*		Genome Sequencing**		Kintelligence		
Sample	Quant Trio DI	DNA input (ng)	Call Rate	Percent Condordance	Call Rate	Percent Condordance	DNA input (ng)	Call Rate	Percent Condordance
MD006_1	**2.4**	2	74.84%	92.73%	97.17%	98.30%	1	99.30%	98.85%
	2.4	0.5	62.72%	84.87%	93.28%	98.34%	0.5	95.99%	97.24%
MD006_0626	**14.9**	2	46.74%	72.75%	84.20%	98.30%	1	97.16%	97.30%
	14.9	0.5	33.72%	57.97%	78.57%	98.12%	0.5	93.34%	95.26%
MD006_0703	**61.4**	2	41.49%	52.82%	84.95%	98.21%	1	92.77%	95.13%
	61.4	0.5	67.31%	36.78%	68.31%	97.84%	0.5	88.44%	91.20%

*Call rate determined from 630,000 total SNPs interrogated **Call rate determined from 2,061,275 total SNPs interrogated

Note: Red highlighted boxes indicate call rates below 60% of total SNPs interrogated by the given technology.

Genealogical Searching Results – Buccal/Saliva Swabs

For all test samples, the GEDmatch match lists were reviewed to confirm matching of all expected Proof Project members and additional known relatives present outside the Proof Project dataset Total shared cM were compiled for a known half grand-aunt and known 2nd cousin match (Table 7) to demonstrate effects of decreasing sample input and sample quality on match ability. Using GSAv2 with DNA exhibiting minimal degradation (DI=2.4) at an input of 2 ng demonstrated close to a 100% match to commercial FamiyTreeDNA match results. GSAv2 kits associated with higher degrees of degradation (DI of 14.9 and 61.4) failed to produce any matching results including unsuccessful self-matching. All six levels and sizes of degraded DNA used to develop WGS high density SNP genotypes matched to all expected documented kits of the Proof Project with minimal decrease in shared cM. Genome sequencing match lists also included more distant, documented relatives to the 39 cM level confirmed to share 8th Great-Grandparents with donor MD006. The Kintelligence derived kits matched all expected Proof Project test takers; however, the variation in shared content of more distant relationships resulted in several relatives only being identified in the Expanded Match list as DI increased and DNA input decreased. Also,

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

increasing numbers of false positive matches were identified in the Expanded Match and All Matches lists as DI increased and DNA input decreased.

Table 7: Total shared cM for mock saliva samples obtainded from One-to-Many Segment Based or Kinship Matching

		GSAv2			Genome Sequencing			Kintelligence		
Sample	Quant Trio DI	Total Family Kits Matched	Half Great-Aunt shared cM	2nd Cousin shared cM	Total Family Kits Matched	Half Great-Aunt shared cM	2nd Cousin shared cM	Total Family Kits Matched (High Confidence)	Half Great-Aunt shared cM	2nd Cousin shared cM
Expected Match		**105**	**465**	**245**	**105**	**465**	**245**	**105**	**465**	**245**
MD006_1	**2.4**	**105**	468.4	244.8	105	469	245	105	408	210
	2.4	**105**	290.5	172.2	105	469	243	104	395	ND
MD006_0626	**14.9**	**1**	ND	ND	105	442	220	104	387	ND
	14.9	**0**	ND	ND	105	406	203	105	316	175
MD006_0703	**61.4**	**0**	ND	ND	105	412	211	104	308	ND
	61.4	**0**	ND	ND	105	414	203	104	294	ND

ND = Not Detected in One-to-Many Match Lists

Applicability to Criminal Justice

The 2021 TWG operational requirements identified the need for further development, assessment, and evaluation of IGG testing procedures for use by crime labs [1]. IGG offers a capability to identify investigative leads when CODIS searching is unproductive, and IGG can provide time efficient methods for removing perpetrators of serial violent crimes, such as rape and murder from the community, thereby increasing public safety. However, use of IGG has preceded establishment of best practices. Development of best practices must start with a systematic evaluation of the laboratory technologies currently used to generate high-density SNP genotypes. The experiments performed provide the community with much needed systematic analyses and direct comparisons of available technologies and allow practitioners to make more informed decisions when working with limited resources. Results will assist in developing lab-specific criteria for processing irreplaceable DNA evidence samples with IGG. Additionally, this systematic evaluation of technological limitations provides backing for development of new genealogical workflows leading to a robust framework supporting effective use of less-than-optimal results in genealogical sample identifications. The ultimate goal of this study is to expand the current body of IGG literature with basic information regarding effects of DNA quality and quantity on the accuracy and call rate of high-density SNP genotype profiles to hopefully expand the use of IGG. Publication of the study's results contributes to the foundational literature required for potential Frye and Daubert hearings.

Conclusions

This study set out to evaluate the impact of sample quality and quantity on SNP genotype development and genealogical comparisons and to provide the forensic community direct comparisons of relevant methodologies to inform decision makers. Bode performed a systematic evaluation comparing data generation from the Illumina GSAv2 BeadChip, Genome Sequencing on the Illumina NovaSeq 6000, and targeted sequencing with ForenSeq Kintelligence on the MiSeq FGx. While all three technologies are sensitive to decreasing DNA input of high-quality extracts, impacts to call rate, heterozygosity, and concordance were observed as DI value increased and sample input decreased. Degradation (DI >4) most profoundly impacted the accuracy of

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

GSAv2 genotyping. Genome sequencing resulted in call rate reductions from >95% to ~60% at low inputs and high DI, but concordance remained high for the >1 million SNPs generated from even the lowest call rates. The Kintelligence chemistry is robust to degradation at optimal inputs, but suffers from allele loss with inputs < 100 pg.

From a genealogist perspective to identify unknown forensic samples, GSAv2 works well with high template samples with minimal degradation but quickly becomes invalid as sample template drops and degradation increases past DI=4. Genome sequencing and Kintelligence, while both applicable to samples of advanced degradation, the methods demonstrate different strengths for genealogical comparisons. Kintelligence produces reliable High Confidence results up through first cousins once removed or equivalent for low input, degraded samples prepared in this study. When such a kit is not returned in High Confidence with northern European test takers, it is because no such individual is in the available database. It was also observed that the Kintelligence shared cM displayed were between 10% and 24% lower than the expected totals from DTC kits in the greater GEDmatch database. For successful genealogical outcomes, Kintelligence should only be used when first cousin once removed (1C1R) or closer relatives are expected in the matching database. If they are expected but are not returned, kits not appearing in High Confidence should not be accepted as reliable when reported only in Expanded or All unless there is support from supplemental data.

Genome sequencing produces results most similar to standard commercial kits, returning higher, more accurate shared cM counts than Kintelligence-derived kits with all but the most degraded or smallest samples Variation in match lists were observed between those of GSAv2-derived kits and genome sequencing-derived kits for the same samples, but the shear increase in number of available genotyped loci may serve to improve matching calculations and eliminate less likely matches, saving relationship review time.

In conclusion, all three IGG genotyping approaches have their applicability in forensic casework and practitioners should carefully review their case/sample metrics prior to selecting a genotyping technology to ensure the best outcomes in both genotype quality and impacts to genealogical workflows.

Products

Scholarly Products

- Manuscript in progress
- Study-level information and a link to the Dryad dataset for NIJ award 15PNIJ-21-GG-04143-MUMU have been submitted to the National Archive of Criminal Justice Data (NACJD).
- Cavanaugh SC, Bowers M, Bever RA, Byrne M, Davoren JM. Comparative Evaluation of Genotyping Technologies for Investigative Genetic Genealogy in Sexual Assault Casework (2024) Dryad, Dataset. https://doi.org/10.5061/dryad.g1jwstr04

Dissemination Activities

1) Cavanaugh SC, Bowers M, Bever RA, Davoren JM. Comparative Evaluation of Illumina Global Screening Array BeadChip, Whole Genome Sequencing, and Verogen ForenSeq Kintelligence to Analyze Degraded and Low Template DNA Extracted from Semen

Samples. Oral presentation at the 76th Annual AAFS Scientific Conference. Denver, CO. February 22, 2024.

2) Cavanaugh SC, Bowers M, Bever RA, Davoren JM. Choosing the Right Workflow: A Comparative Evaluation of FIGG Genotyping Technologies for Sexual Assault Casework. Oral presentation at NIST Symposium on Forensic DNA Analysis. Rockville, MD. April 24, 2024.

3) Cavanaugh SC, Bowers M, Bever RA, Davoren JM. Choosing the Right Workflow: A Comparative Evaluation of FIGG Genotyping Technologies for Sexual Assault Casework. Oral presentation to National Technology Validation and Implementation Collaborative (NTVIC) FIGG-TVWG. Virtual. May 10, 2024.

4) Cavanaugh SC, Bowers M, Bever RA, Byrne M, Davoren JM. Choosing the Right Workflow: An Evaluation of Genotyping Technologies for Forensic Investigative Genetic Genealogy casework. Oral presentation at Bode Technology's 23rd annual Forensic DNA Conference. Atlanta, GA. June 27, 2024.

5) Cavanaugh SC, Bowers M, Bever RA, Byrne M, Davoren JM. Choosing the Right Workflow: An Evaluation of Genotyping Technologies for Forensic Investigative Genetic Genealogy casework. Oral presentation at Green Mountain annual DNA Conference. Burlington, VT. July 23, 2024

6) Cariola M and Bathrick AS. Enhancing Forensic Accuracy: Comparative Evaluation of Genotyping Technologies in FIGG. The ISHI Report. May 2024. https://promega.foleon.com/theishireport/the-ishi-report-may-2024/enhancing-forensic-accuracy-comparative-evaluation-of-genotyping-technologies-in-figg

7) Cavanaugh SC, Bowers M, Bever RA, Byrne M, Davoren JM. FIGG Genotyping Workflows for Sexual Assault Casework: A systematic Evaluation of Laboratory Processing and Genealogical Impacts. Abstract submitted for Oral Presentation at International Symposium on Human Identification. San Antonio, TX. September 23-26,2024

References

1. Forensic Science Technology Working Group. Operational Requirements. April 2021.
2. Board for Certification of Genealogists, Genealogical Standards, 50th Anniversary Edition (Nashville & New York: Ancestry Imprint, Turner Publishing, 2014), 1-3, and Thomas W. Jones, Mastering Genealogical Proof (Arlington, Va.: National Genealogical Society, 2013).
3. Verogen. ForenSeq™ Kintelligence Kit Reference Guide, VD2020053, Rev. C. Accessed 30 October 2023.
4. Illumina, Inc. Infinium™ Global Screening Array-24 v2.0 BeadChip. Pub. No. 370-2016-016-E QB 6363. 2018. Available: https://countingchromosomes.com/blogfiles/2019-08-17/Illumina-GSA-data-sheet-370-2016-016.pdf

5. Verogen. Universal Analysis Software v2.0 Reference Guide, VD2019002 Rev. E. Accessed 25 October 2022.
6. Illumina, Inc. GenomeStudio® Genotyping Module v2.0 Software Guide. Document #11319113 v01. Sept 2016.
7. Illumina, Inc. DRAGEN™ secondary analysis. Data Sheet, M-GL-00680 v9.0. 2024. Accessed June 27, 2024.
8. Nelson J. (2009). Repair of damaged DNA for forensic analysis. National Institute of Justice. Available: https://www.ojp.gov/pdffiles1/nij/grants/227498.pdf
9. Hedman J, Ansell R, Nordgaard A. A ranking index for quality assessment of forensic DNA profiles. BMC Res Notes. 2010;3:290.
10. Russell DA, Gorden EM, Peck MA, Neal CM, Heaton MC, Bouchet JL, Koeppel AF, Ciuzio E, Turner SD, Reedy CR. Developmental Validation of the Illumina Infinium Assay Using the Global Screening Array on the iScan System for Use in Forensic Laboratories. Forensic Genomics 2023;3:15-24. DOI: 10.1089/forensic.2022.0013
11. Bever R, Davoren J, Sweeney E, Vreeland T, Cariola M. Assessing the reliability of SNP microarray data for forensic genealogy. Presentation at: American Academy of Forensic Sciences Annual Meeting, 2020: Anaheim, CA.
12. Jorde LB. Genome-wide forensic DNA analysis. NCJ Document number: 251934 [Grant Final Report]. National Institutes of Justice award number 2012-DN-BX-K037. Received Aug 2018. Available: https://www.ojp.gov/ncjrs/virtual-library/abstracts/genome-wide-forensic-dna-analysis
13. Gassmann M, McHoull B. DNA Integrity Number (DIN) with the Agilent 2200 TapeStation System and the Agilent Genomic DNA ScreenTape assay. Agilent Technologies, Inc. Publication number 599105258EN. 1 Dec 2015. Available: https://www.agilent.com/cs/library/applications/5991-5258EN.pdf
14. Snedecor J, Fennell T, Stadick S, Homer N, Antunes J, Stephens K, Holt C. Fast and accurate kinship estimation using sparse SNPs in relatively large database searches. Foren Sci Int Genet 2022;61:102769. DOI: 10.1016/j.fsigen.2022.102769
15. Hemmerich KJ. General aging theory and simplified protocol for accelerated aging of medical devices. Med. Plast. Biomater. 1998;5:16–23.
16. Hukins DWL, Mahomed A, Kukureka SN. Accelerated aging for testing polymeric biomaterials and medical devices. Med. Eng. Phys. 2008;30(10):1270–1274.
17. de Vries JH, Kling D, Vidaki A, Arp P, Kalamara V, Verbiest MMPJ, Piniewska-Róg D, Parsons TJ, Uitterlinden A, Kayser M. Impact of SNP microarray analysis of compromised DNA on kinship classification success in the context of investigative genetic genealogy. Forensic Sci Int Genet 2022;56:102625. DOI: 10.1016/j.fsigen.2021.102625
18. Antunes J, Walichiewicz P, Forouzmand E, Barta R, Didier M, Han Y, Perez JC, Snedecor J, et al. Developmental Validation of the ForenSeq® Kintelligence kit, MiSeq FGx® Sequencing System and ForenSeq Universal Analysis Software. Forensic Sci Int Genet 2024;71:103055. DOI: 10.1016/j.fsigen.2024.103055

19. Cavanaugh SC, Bowers M, Bever RA, Byrne M, Davoren JM. Comparative Evaluation of Genotyping Technologies for Investigative Genetic Genealogy in Sexual Assault Casework (2024) Dryad, Dataset. https://doi.org/10.5061/dryad.g1jwstr04
20. Radecke SM, Antunes J, Snedecor J, Zlatkov C, Devesse L, Padmabandu G, Stephens KM. Evaluation of a High-Throughput Dense Single-Nucleotide Polymorphism PCR Multiplex Next Generation Sequencing Application for Human Remains Identification. Forensic Genomics 2023;3(3):75-93. DOI: 10.1089/forensic.2023.0005