Choosing the Right Workflow: A Comparative Evaluation of FIGG Genotyping Technologies for Sexual Assault Casework

Sarah Cavanaugh, Madalyn Bowers, Melinde Byrne, Robert Bever, and Jon Davoren

Green Mountain Meeting
July 2024

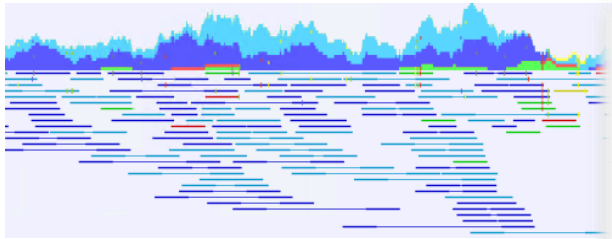
DISCLOSURE STATEMENT

- 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 presentation are those of the authors and do not necessarily reflect those of the Department of Justice.

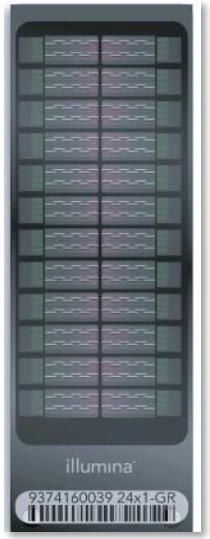
- Any commercial products or instruments described are for the purposes of complete description of experimental procedures. This does not imply a recommendation or endorsement by Bode Technology.

PROGRAM OBJECTIVE

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



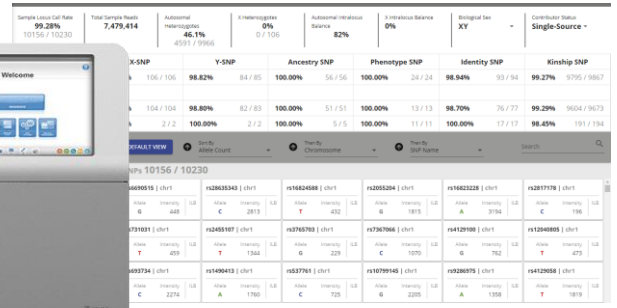
Optimal Input: >10 ng



Optimal Input: 200 ng

[GED match] PRO™

Foundation of genetic genealogical databases like GEDmatch and FamilyTreeDNA

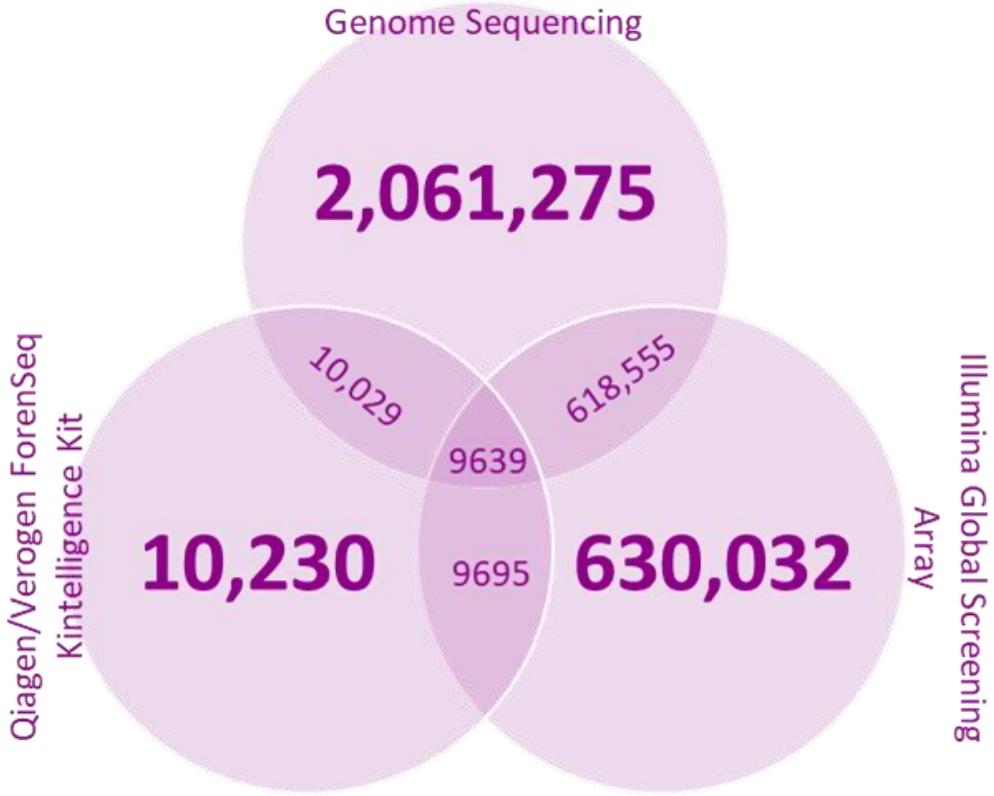


Optimal Input: 1 ng

PROGRAM OBJECTIVE – Answer 2 Main Questions

1. What effects are observed in response to **decreasing sample input** and **decreasing sample quality** for each method/technology when analyzing DNA obtained from sexual assault samples?
    - Compare results within/across technologies

2. What impacts are observed to **genealogical analyses** when compared against GEDmatch database?



Numbers of overlapping SNP loci interrogated by each method

QUESTION 1:

ASSESSMENT OF SENSITIVITY TO DECREASING SAMPLE INPUT

Bode Technology

Answers with Confidence and Accuracy

SENSITIVITY SAMPLE CONSTRUCTION

- Samples:
  - Fresh semen collections, 2 known donors
    - IRB consent for collection and genealogical matching
    - At least 1 relative available for database comparison, as distant as 2nd cousin
  - RM8393
    - Human DNA for Whole-Genome Variant Assessment (Son of Chinese Ancestry) (HG-005)
    - Extensively characterized DNA sample with high coverage sequence benchmark data
- Sample Preparation:
  - Qiagen EZ1&2 ® DNA Investigator ® Kit extraction
  - Quantifiler® Trio and Qubit™ dsDNA HS assay
  - Reference genotypes generated with GSA v2 (200ng input)

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

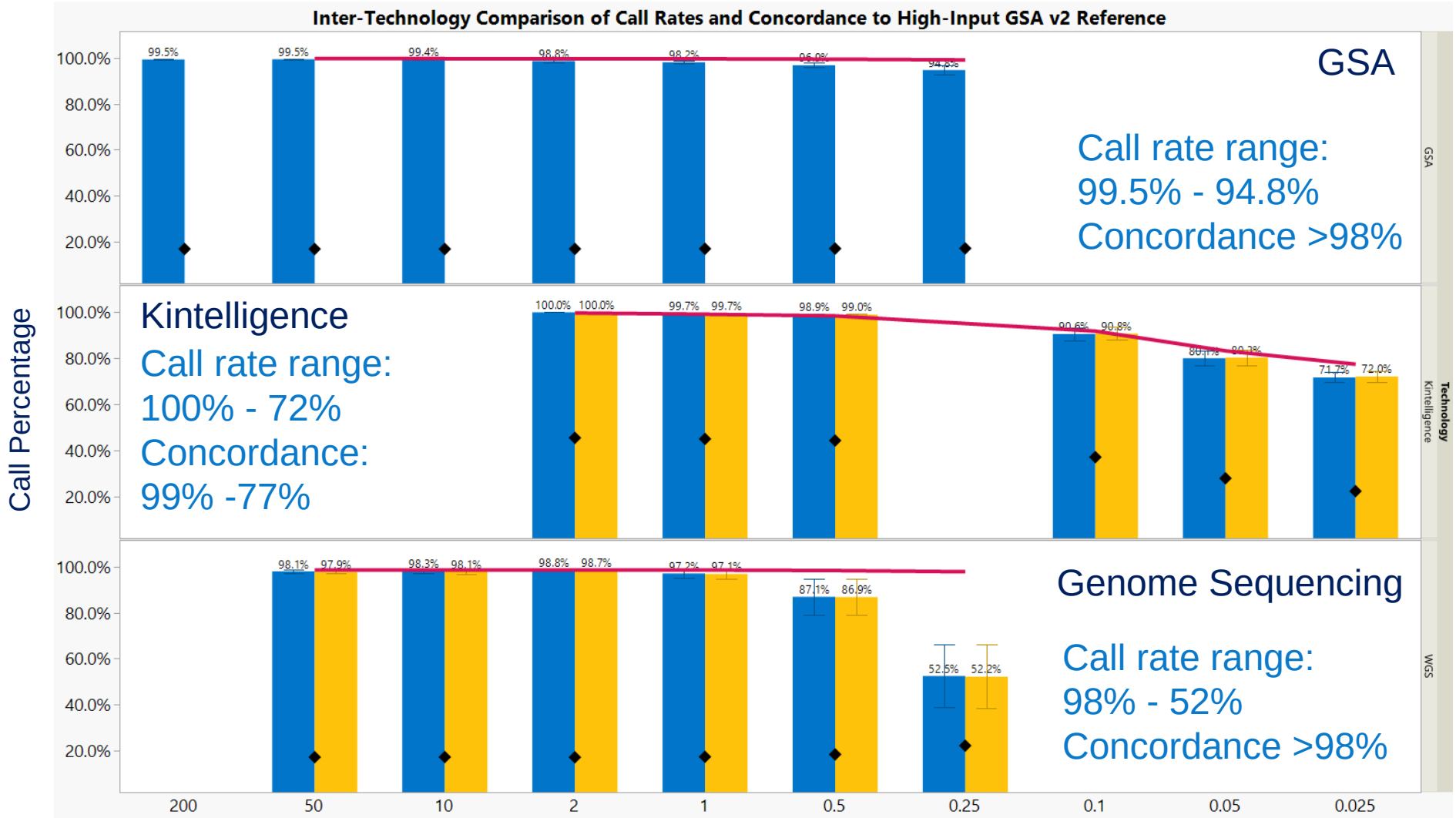
SENSITIVITY SAMPLE PROCESSING

Technology	Sample Prep	Processing Parameters	Bioinformatic Analysis
Target Sequencing	ForenSeq® Kintelligence Kit	Pooling of 3 libraries/run, MiSeq FGx Reagents w/ standard flow cell	UAS v2.5, 1.5% AT/IT (10X cov minimum), 50% Intra-locus Balance
SNP Microarray	WGA and hybridization to custom Illumina GSA v2 BeadChip	Illumina iScan	GenomeStudio® v2.0 Genotyping Module with internally optimized parameters/cluster files
Genome Sequencing	dsDNA library prep, internally optimized workflow	Illumina NovaSeq 6000, 2x150 bp reads, 30X Depth	DRAGEN Pipeline (Edico Genome, Inc) internally optimized parameters

- GEDmatch-uploadable GT data in .csv format provided from array and sequencing analyses
- Secondary analysis with SAMtools[1] and BCFtools in Galaxy (www.usegalaxy.org) and customized Excel Workbooks
- Statistical analysis in SAS JMP v15

1) https://doi.org/10.1093/gigascience/giab008

INTER-TECHNOLOGY SENSITIVITY COMPARISONS – Technologies Demonstrate Sensitivity to Forensic Level Inputs



Call Rate (% Loci Called)

GSA Consistent Call Rate (% GSA-overlapping Loci Called)

Concordance Rate to GSA-generated genotype

Calculated Heterozygosity Average

QUESTION 1:
ASSESSMENT OF SPECIFICITY TO DEGRADED DNA

Bode Technology

Answers with Confidence and Accuracy

DEGRADATION SAMPLE CONSTRUCTION

- Sample Preparation, 3 semen donors

Method	Procedure	Time Points	Sample Input
Depurination	10X Depurination Buffer/ HCl incubation @ 70 °C	12 hrs; 24 hrs; 36 hrs; 48 hrs (n =12)	Post-extraction semen aliquot, 50 µl extract
Hydrolytic/ oxidative damage	Fenton Reaction: Fe-EDTA/H_2O_2 incubation @ 37 °C	12 hrs, 24 hrs, 48 hrs, 52 hrs (n =12)	Pre-extracted whole semen aliquot, 20 µl
UV irradiation	UV Crosslinker incubation (245 nm λ)	120 sec, 360 sec, 600 sec, 720 sec (n =12)	Post-extraction semen aliquot, 50 µl extract

- Verification of Degradation
  - Agilent TapeStation Genomic ScreenTape, Quantifiler Trio, and STR profile generation with Promega PowerPlex Fusion 6C
    - DIN and DI
    - Profile recovery, balance, and Forensic Index

DEGRADATION SAMPLE PROCESSING

Technology	Sample Prep	Processing Parameters	Data Analysis
Target Sequencing	ForenSeq Kintelligence Kit, 1ng input	Pooling of 3 libraries/run, MiSeq FGx Reagents w/ standard flow cell	UAS v2.5, 1.5% AT/IT (10X cov minimum), 50% Intra-locus Balance
SNP Microarray	WGA and hybridization to custom Illumina GSA v2 BeadChip, 2 ng input	Illumina iScan	GenomeStudio® v2.0 Genotyping Module with internally optimized parameters/cluster files
Genome Sequencing	dsDNA library prep, internally optimized workflow, 2 ng input	Illumina NovaSeq 6000, 2x150 bp reads, 30X Depth	DRAGEN Pipeline (Edico Genome, Inc) internally optimized parameters

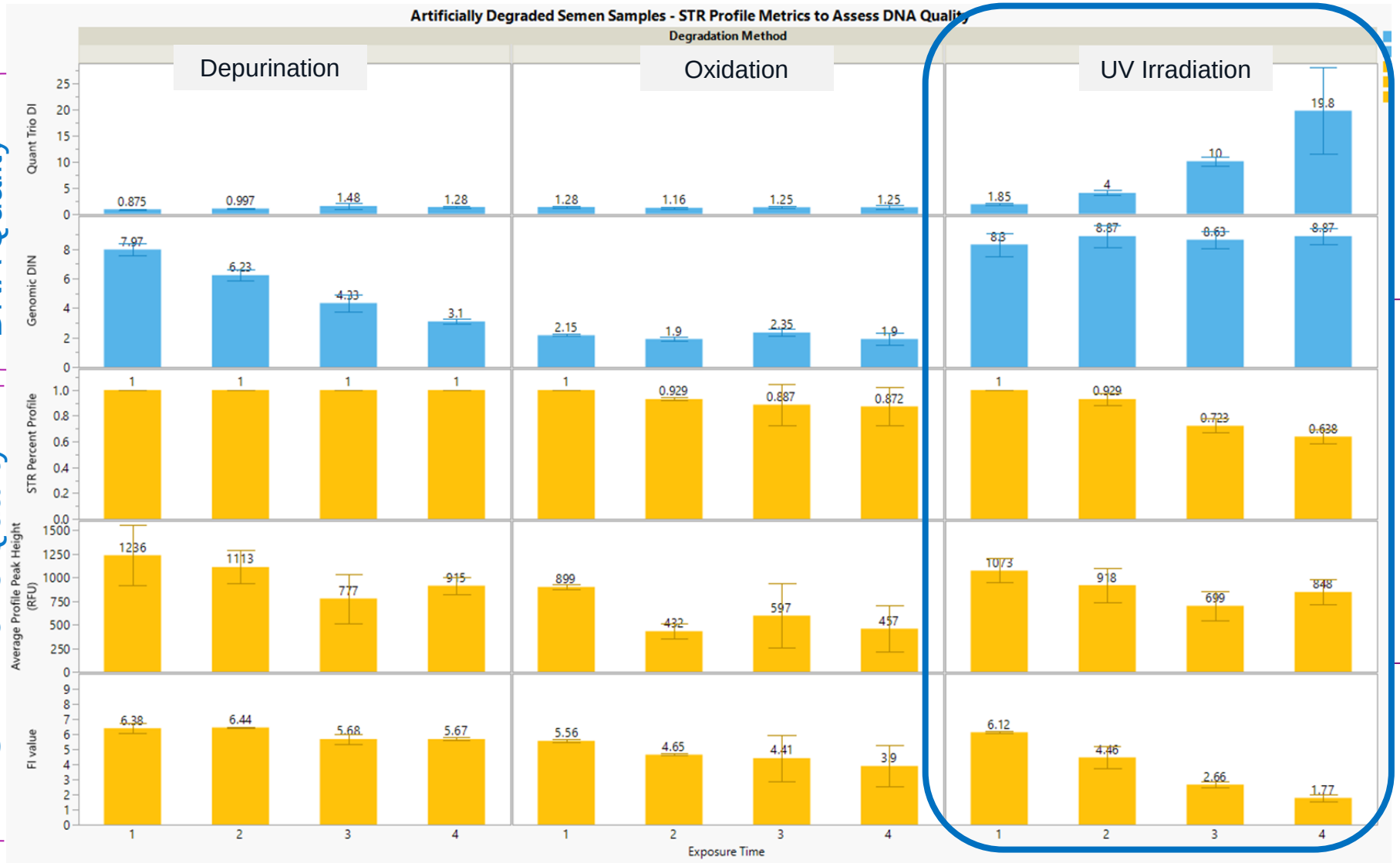
- GEDmatch-uploadable GT data in .csv format provided from array and sequencing analyses
- Secondary analysis with SAMtools[1] and BCFtools in Galaxy (www.usegalaxy.org) and customized Excel Workbooks
- Statistical analysis in SAS JMP v15

1) https://doi.org/10.1093/gigascience/giab008

DEGRADATION QUALITY CONTROL

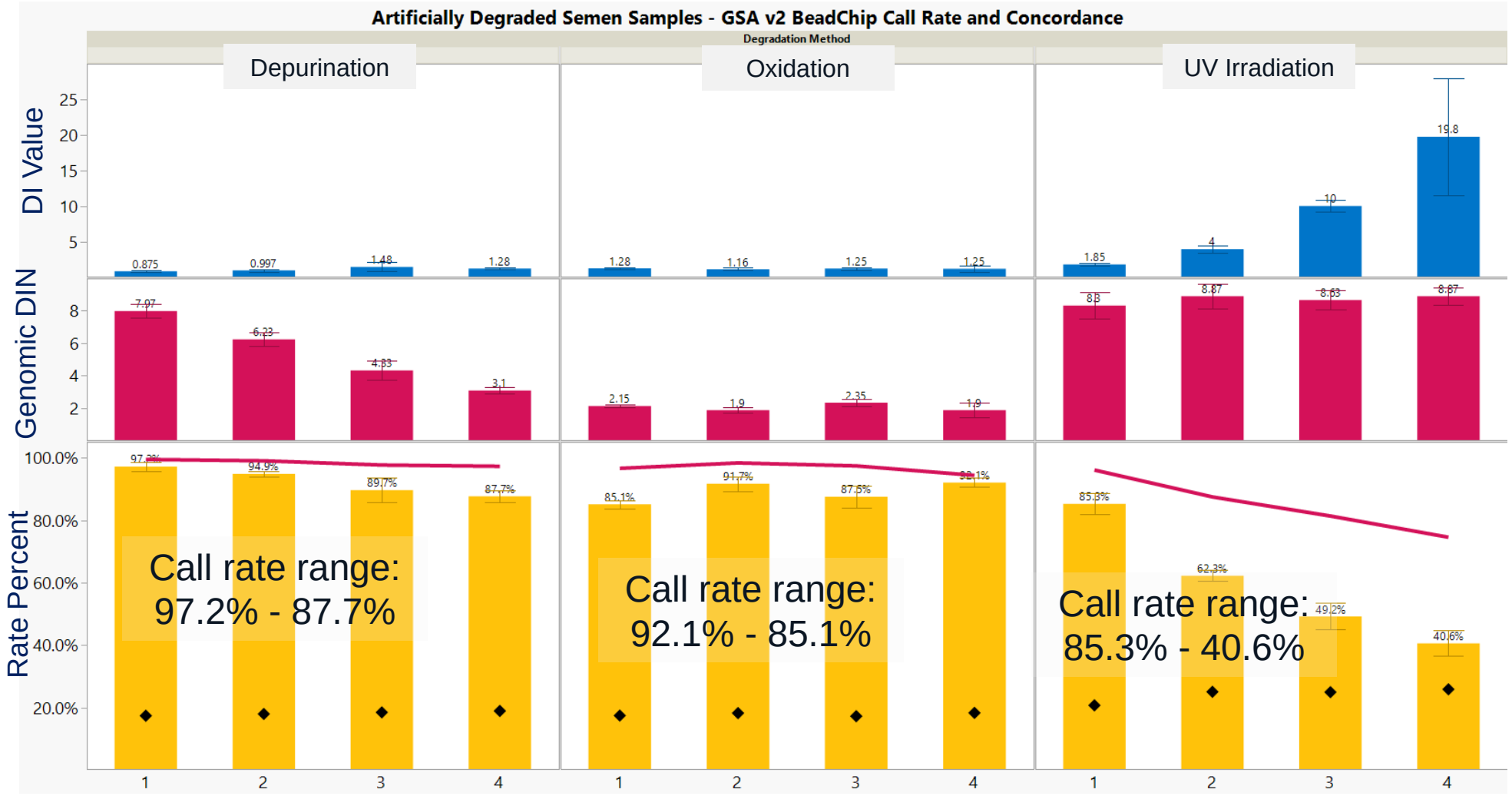
DNA Quality

STR Profile Quality



- STR results confirmed impacts to profile quality due to all reactions
- Greatest impacts due to UV irradiation

DEGRADED SAMPLE ANALYSIS WITH GSA V2 BEADCHIPS – Degradation Index Increase Correlated with Call Rate Decrease



Degradation Index

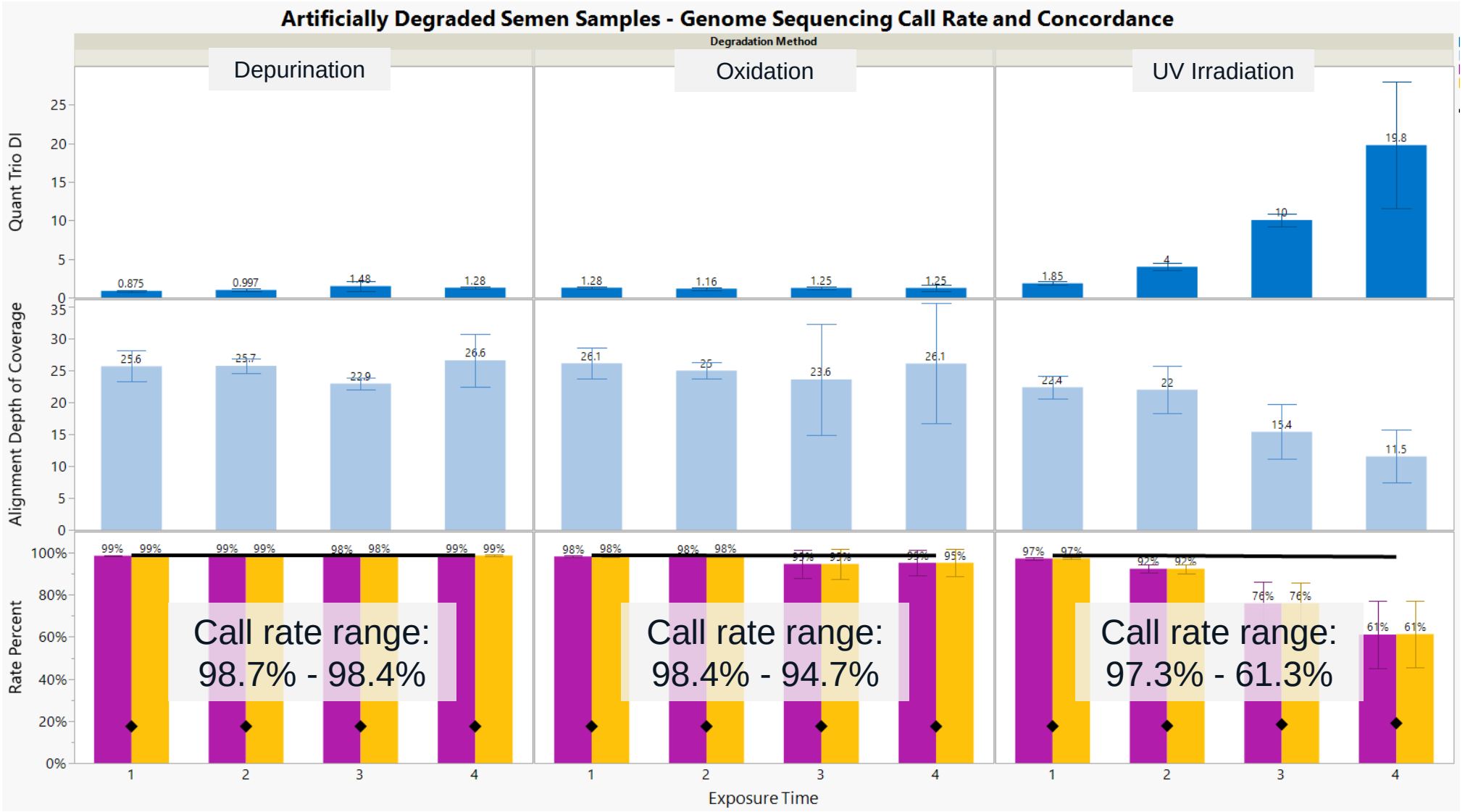
Genomic DIN

Call Rate (% Loci Called)

Concordance Rate to GSA-generated genotype

Calculated Heterozygosity Average

DEGRADED SAMPLE ANALYSIS WITH GENOME SEQUENCING – Degradation Index Increase Correlated with Call Rate Decrease



Degradation Index

Depth of Coverage

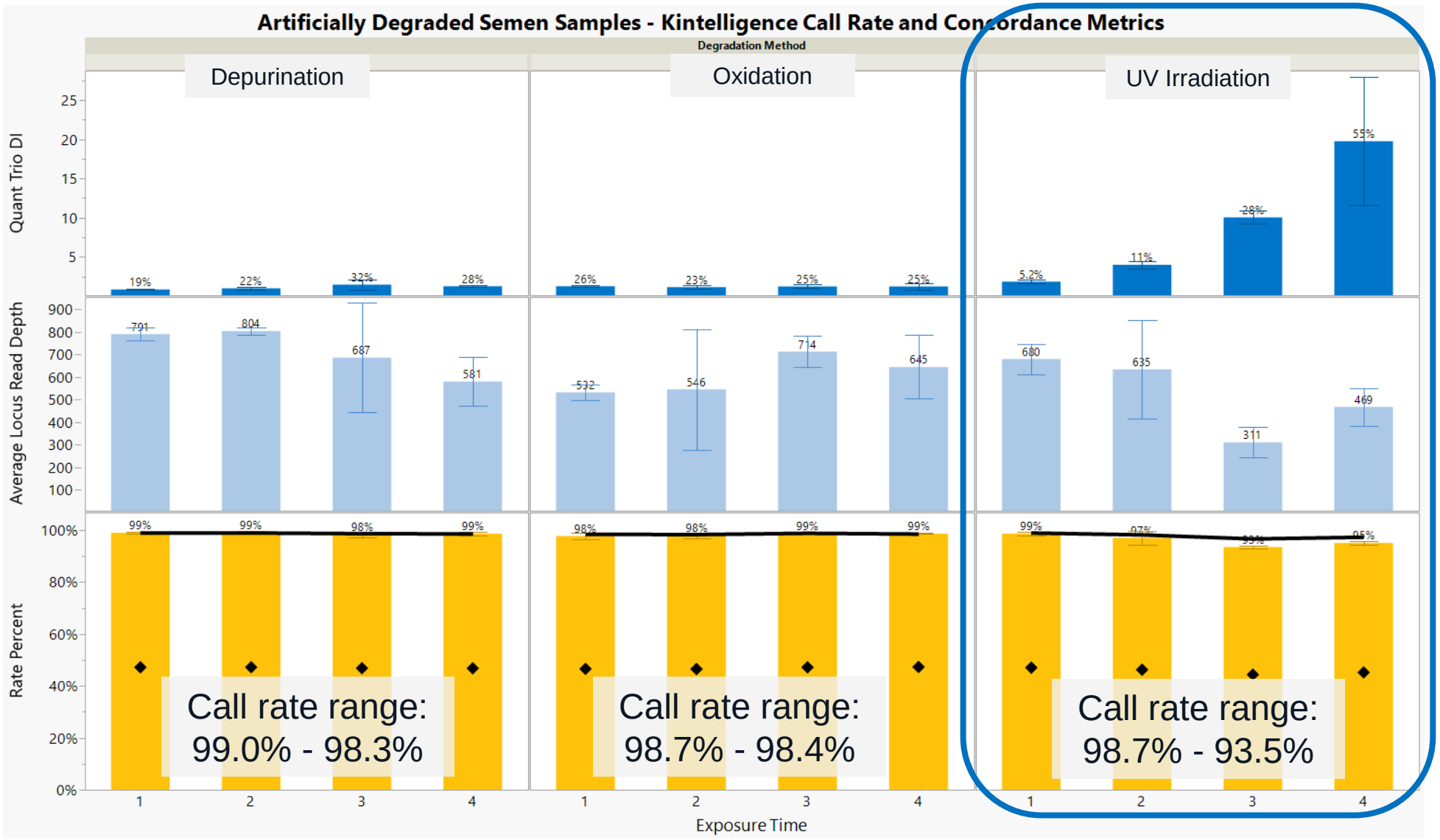
Call Rate (% Loci Called)

GSA Consistent Call Rate (% GSA-overlapping Loci Called)

Concordance Rate to GSA-generated genotype

Calculated Heterozygosity Average

DEGRADED SAMPLE ANALYSIS WITH KINTELLIGENCE – Robust to increasing degradation with optimal input



Degradation Index

Locus Read Coverage

Call Rate (% Loci Called)

Concordance Rate to GSA-generated genotype

Calculated Heterozygosity Average

QUESTION 2: GENEALOGICAL MATCHING ASSESSMENT

GENEALOGICAL ASSESSMENT

- GEDmatch/GEDmatch PRO® comparisons
  - Genome Sequencing and GSA-generated sensitivity GT uploaded through GEDmatch Classic as “Research” samples
    - Degraded GT uploaded through GEDmatch PRO
  - Kintelligence-generated GT uploaded through GEDmatch PRO portal as “Validation” samples
    - Designations allows comparison against the database and known relatives without making the samples searchable to outside users

- Are we matching to the known relatives in the database?
- What effects observed on:
  - Number of usable SNPs
  - Total shared cM
  - Length of longest shared segment
- Assess application of Kintelligence data to genealogical workflows

GENEALOGICAL ASSESSMENT – MATCHING ALGORITHMS

GSAv2 and Genome Sequencing (standard kits)

One-to-Many Segment Based

Total shared cM >50 cM

Kintelligence

One-to-Many Kinship[1]

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

1. Snedecor et al. (2022) FSI:Genetics 61:102769

GEDmatch GENEALOGICAL COMPARISONS – GSA and Genome Sequencing Sensitivity Samples

One-to-Many Match List – GSAv2 Genotypes

MD002 - GEDmatch PRO One-to-Many Segment Based Match Lists Generated with Decreasing DNA Input - GSAv2

	200ng	50ng	10ng	2ng	1ng	0.5ng	0.25ng
Known Relative 1C1R	Alias 1 421.2/82.5 2.55 (77894 SNPs)	Alias 1 421.2/82.5 2.55 (77910 SNPs)	Alias 1 421.2/82.5 2.55 (77904 SNPs)	Alias 1 421.2/82.5 2.55 (77852 SNPs)	Alias 1 421.2/82.5 2.55 (77729 SNPs)	Alias 1 421.2/82.5 2.55 (77363 SNPs)	Alias 1 421.2/82.5 2.55 (76906 SNPs)
Alias 2 59/28.5 3.96 (87915 SNPs)	Alias 2 59/28.5 3.96 (87952 SNPs)	Alias 2 59/28.5 3.96 (87935 SNPs)	Alias 2 59/28.5 3.96 (87750 SNPs)	Alias 2 59/28.5 3.96 (87462 SNPs)	Alias 2 59/28.5 3.96 (86791 SNPS)	Alias 2 59/28.5 3.96 (85904 SNPs)
Alias 3 58.5/11.7 3.97 (281725 SNPs)	Alias 4 57/47.4 3.99 (79261 SNPs)	Alias 4 57/47.4 3.99 (79245 SNPs)	Alias 4 57/47.4 3.99 (79150 SNPs)	Alias 4 57/47.4 3.99 (78979 SNPs)	Alias 4 57/47.4 3.99 (78469 SNPs)	Alias 4 57/47.4 3.99 (77869 SNPs)
Alias 4 57/47.4 3.99 (79232 SNPs)	Alias 5 54.8/22.6 4.02 (279321 SNPs)	Alias 5 54.8/22.6 4.02 (279152 SNPs)	Alias 5 54.8/22.6 4.02 (278440 SNPs)	Alias 5 54.8/22.6 4.02 (277368 SNPs)	Alias 5 54.8/22.6 4.02 (274727 SNPs)	Alias 5 54.8/22.6 4.02 (271181 SNPs)
Alias 5 54.8/22.6 4.02 (279191 SNPs)	Alias 6 53.3/14.4 4.04 (281848 SNPs)	Alias 6 53.3/14.4 4.04 (281673 SNPs)	Alias 6 53.3/14.4 4.04 (280930 SNPs)	Alias 6 53.3/14.4 4.04 (279853 SNPs)	Alias 6 53.3/14.4 4.04 (277122 SNPs)	Alias 7 52.8/11.1 4.04 (274287 SNPs)
Alias 6 53.3/14.4 4.04 (281734 SNPs)	Alias 7 52.8/11.1 4.04 (282593 SNPs)	Alias 7 52.8/11.1 4.04 (282440 SNPs)	Alias 7 52.8/11.1 4.04 (281690 SNPs)	Alias 7 52.8/11.1 4.04 (280590 SNPs)	Alias 7 52.8/11.1 4.04 (277859 SNPs)	Alias 8 51.3/51.3 4.06 (88761 SNPs)
Alias 7 52.8/11.1 4.04 (282466 SNPs)	Alias 8 51.3/51.3 4.06 (90989 SNPs)	Alias 8 51.3/51.3 4.06 (90964 SNPs)	Alias 8 51.3/51.3 4.06 (90774 SNPs)	Alias 8 51.3/51.3 4.06 (90474 SNPs)	Alias 8 51.3/51.3 4.06 (89692 SNPs)	Alias 3 50.7/11.7 4.07 (273520 SNPs)
Alias 8 51.3/51.3 4.06 (90957 SNPs)	Alias 3 50.7/11.7 4.07 (281857 SNPs)	Alias 3 50.7/11.7 4.07 (281676 SNPs)	Alias 3 50.7/11.7 4.07 (280937 SNPs)	Alias 3 50.7/11.7 4.07 (279881 SNPS)	Alias 3 50.7/11.7 4.07 (277147 SNPs)	Alias 9 50.7/10.5 4.07 (273544 SNPs)
Alias 9 50.7/10.5 4.07 (281510 SNPs)	Alias 9 50.7/10.5 4.07 (281643 SNPs)	Alias 9 50.7/10.5 4.07 (281483 SNPs)	Alias 9 50.7/10.5 4.07 (280760 SNPs)	Alias 9 50.7/10.5 4.07 (279683 SNPs)	Alias 9 50.7/10.5 4.07 (277012 SNPs)	Alias 10 50.5/32.1 4.08 (86573 SNPs)
Alias 10 50.5/32.1 4.08 (88582 SNPs)	Alias 10 50.5/32.1 4.08 (88629 SNPs)	Alias 10 50.5/32.1 4.08 (88601 SNPs)	Alias 10 50.5/32.1 4.08 (88406 SNPs)	Alias 10 50.5/32.1 4.08 (88154 SNPs)	Alias 10 50.5/32.1 4.08 (87467 SNPs)	Alias 11 50.5/32.1 4.08 (86209 SNPs)
Alias 11 50.5/32.1 4.08 (88205 SNPs)	Alias 11 50.5/32.1 4.08 (88236 SNPs)	Alias 11 50.5/32.1 4.08 (88214 SNPs)	Alias 11 50.5/32.1 4.08 (88031 SNPs)	Alias 11 50.5/32.1 4.08 (87775 SNPs)	Alias 11 50.5/32.1 4.08 (87069 SNPs)	Alias 12 50.3/11.5 4.08 (272469 SNPs)

Key:	Kit Alias Shared cM/Longest Segment Generation Estimate (SNP Overlap)

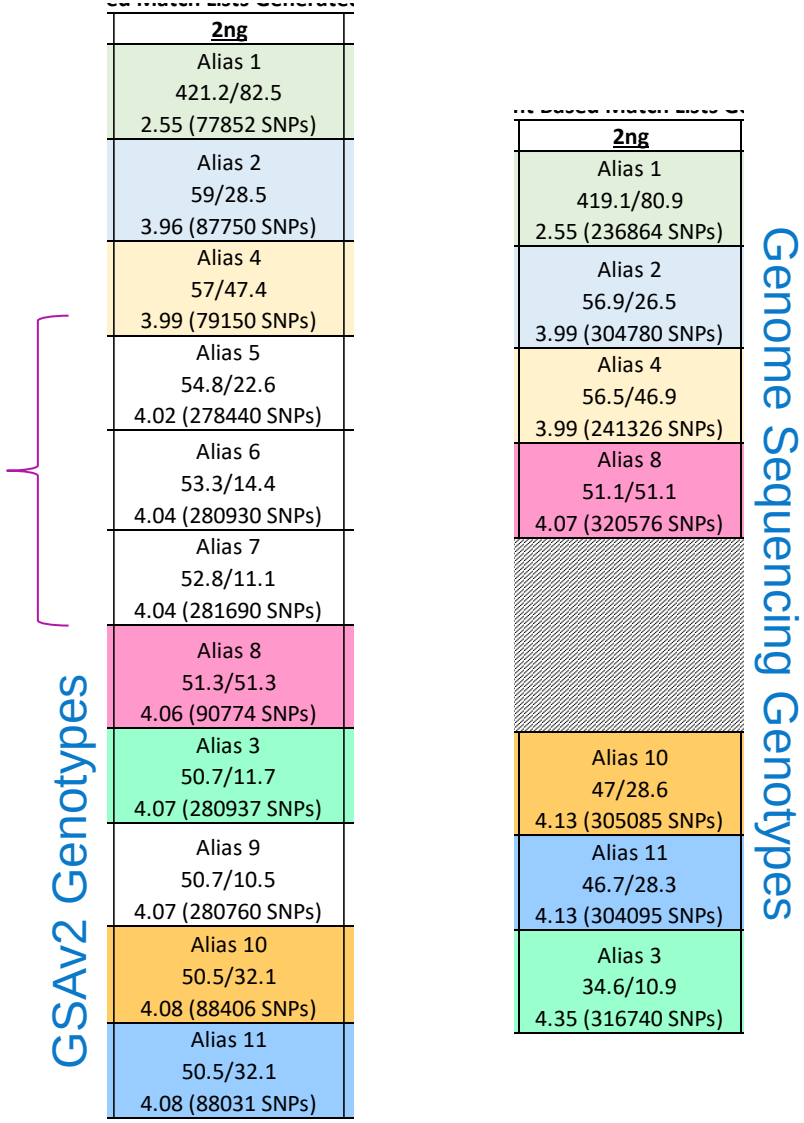
One-to-Many Match List – Genome Sequencing Genotypes

MD002 - GEDmatch PRO One-to-Many Segment Based Match Lists Generated with Decreasing DNA Input - Genome Sequencing

	50ng	10ng	2ng	1ng	0.5ng	0.25ng
Known Relative 1C1R	Alias 1 419.4/80.9 2.55 (235236 SNPs)	Alias 1 420.7/82.5 2.55 (236781 SNPs)	Alias 1 419.1/80.9 2.55 (236864 SNPs)	Alias 1 419.1/80.9 2.55 (236894 SNPs)	Alias 1 407.2/49.8 2.57 (218529 SNPs)	Alias 1 375.3/41.7 2.63 (156040 SNPs)
Alias 2 56.7/26.5 3.99 (302743 SNPs)	Alias 2 56.9/26.5 3.99 (304751 SNPs)	Alias 2 56.9/26.5 3.99 (304780 SNPs)	Alias 2 56.9/26.5 3.99 (304785 SNPs)	Alias 2 56.9/26.5 3.99 (280626 SNPs)	Alias 4 54.6/45.1 4.02 (158309 SNPs)
Alias 4 56.5/46.9 3.99 (239714 SNPs)	Alias 4 56.5/46.9 3.99 (241254 SNPs)	Alias 4 56.5/46.9 3.99 (241326 SNPs)	Alias 4 56.5/46.9 3.99 (241305 SNPs)	Alias 8 51.1/51.1 4.07 (295017 SNPs)	Alias 2 54.3/23.9 4.02 (197923 SNPs)
Alias 8 51.1/51.1 4.07 (318398 SNPs)	Alias 8 51.1/51.1 4.07 (320505 SNPs)	Alias 8 51.1/51.1 4.07 (320576 SNPs)	Alias 8 51.1/51.1 4.07 (320601 SNPs)		Alias 13 54.2/47.2 4.11 (201238 SNPs)
					Alias 8 51.1/51.1 4.07 (208060 SNPs)
Alias 10 47/28.6 4.13 (303008 SNPs)	Alias 10 47/28.6 4.13 (305043 SNPs)	Alias 10 47/28.6 4.13 (305085 SNPs)	Alias 10 47/28.6 4.13 (305074 SNPs)	Alias 10 46.7/28.3 4.13 (280769 SNPs)	Alias 10 31.3/18.3 4.42 (197958 SNPs)
Alias 11 46.7/28.3 4.13 (302003 SNPs)	Alias 11 46.7/28.3 4.13 (304021 SNPs)	Alias 11 46.7/28.3 4.13 (304095 SNPs)	Alias 11 46.7/28.3 4.13 (304072 SNPs)	Alias 11 46.7/28.3 4.13 (279857 SNPs)	Alias 11 44/18.3 4.17 (197295 SNPs)
Alias 3 34.5/10.9 4.35 (313836 SNPs)	Alias 3 34.5/10.9 4.35 (316423 SNPs)	Alias 3 34.6/10.9 4.35 (316740 SNPs)	Alias 3 34.5/10.9 4.35 (316338 SNPs)	Alias 3 34.5/10.9 4.35 (288551 SNPs)	Alias 3 42.5/10.9 4.2 (203616 SNPs)
				Alias 4 35.6/26 4.33 (222597 SNPs)	

- No impact to known relative matching, out to 2nd cousin, was observed with decreasing DNA input and GSA sample processing.

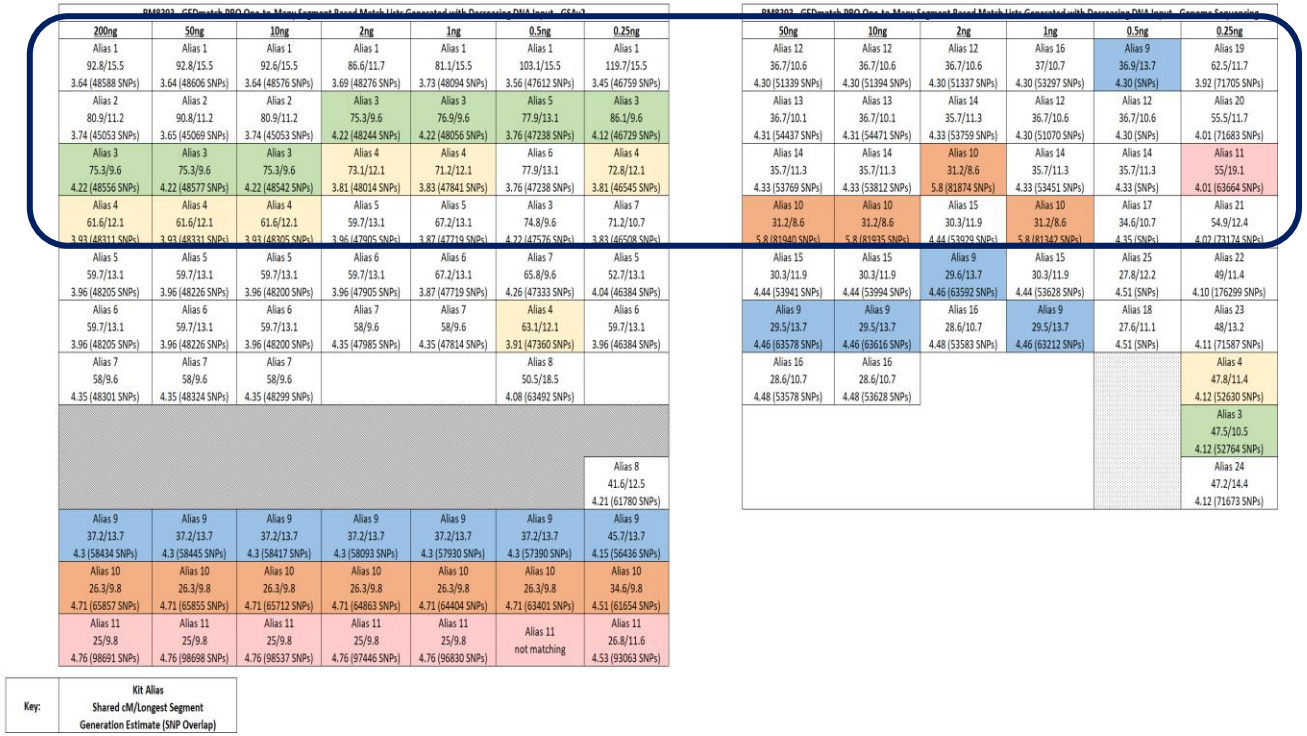
GEDmatch GENEALOGICAL COMPARISONS – GSA and Genome Sequencing Sensitivity Samples



- Observed trends between One-to-Many Segment Based Matching with GSAv2 vs Genome Sequencing genotypes:
  - Reductions in shared cM for genome sequencing results
    - Within a 10% expected variation among kit types
  - Kits removed from genome sequencing match lists
    - Longest segment <20 cM
    - Increasing number of DNA genotypes may improve match calculations
  - Genealogical perspective – No functional differences when working with either kit type

GEDmatch GENEALOGICAL COMPARISONS – GSA and Genome Sequencing Sensitivity Samples

RM8393 Matching Kits – GSA | Genome Sequencing ~2 million SNP



200ng	50ng	10ng	2ng	1ng	0.5ng	0.25ng
Alias 1 92.8/15.5 3.64 (48588 SNPs)	Alias 1 92.8/15.5 3.64 (48606 SNPs)	Alias 1 92.6/15.5 3.64 (48576 SNPs)	Alias 1 86.6/11.7 3.69 (48276 SNPs)	Alias 1 81.1/15.5 3.73 (48094 SNPs)	Alias 1 103.1/15.5 3.56 (47612 SNPs)	Alias 1 119.7/15.5 3.45 (46759 SNPs)
Alias 2 80.9/11.2 3.74 (45053 SNPs)	Alias 2 90.8/11.2 3.65 (45069 SNPs)	Alias 2 80.9/11.2 3.74 (45053 SNPs)	Alias 3 75.3/9.6 4.22 (48244 SNPs)	Alias 3 76.9/9.6 4.22 (48056 SNPs)	Alias 5 77.9/13.1 3.76 (47238 SNPs)	Alias 3 86.1/9.6 4.12 (46729 SNPs)
Alias 3 75.3/9.6 4.22 (48556 SNPs)	Alias 3 75.3/9.6 4.22 (48577 SNPs)	Alias 3 75.3/9.6 4.22 (48542 SNPs)	Alias 4 73.1/12.1 3.81 (48014 SNPs)	Alias 4 71.2/12.1 3.83 (47841 SNPs)	Alias 6 77.9/13.1 3.76 (47238 SNPs)	Alias 4 72.8/12.1 3.81 (46545 SNPs)
Alias 4 61.6/12.1 3.93 (48311 SNPs)	Alias 4 61.6/12.1 3.93 (48331 SNPs)	Alias 4 61.6/12.1 3.93 (48305 SNPs)	Alias 5 59.7/13.1 3.96 (47905 SNPs)	Alias 5 67.2/13.1 3.87 (47719 SNPs)	Alias 3 74.8/9.6 4.22 (47576 SNPs)	Alias 7 71.2/10.7 3.83 (46508 SNPs)
Alias 5 59.7/13.1 3.96 (48205 SNPs)	Alias 5 59.7/13.1 3.96 (48226 SNPs)	Alias 5 59.7/13.1 3.96 (48200 SNPs)	Alias 6 59.7/13.1 3.96 (47905 SNPs)	Alias 6 67.2/13.1 3.87 (47719 SNPs)	Alias 7 65.8/9.6 4.26 (47333 SNPs)	Alias 5 52.7/13.1 4.04 (46384 SNPs)
Alias 6 59.7/13.1 3.96 (48205 SNPs)	Alias 6 59.7/13.1 3.96 (48226 SNPs)	Alias 6 59.7/13.1 3.96 (48200 SNPs)	Alias 7 58/9.6 4.35 (47985 SNPs)	Alias 7 58/9.6 4.35 (47814 SNPs)	Alias 4 63.1/12.1 3.91 (47360 SNPs)	Alias 6 59.7/13.1 3.96 (46384 SNPs)
Alias 7 58/9.6 4.35 (48301 SNPs)	Alias 7 58/9.6 4.35 (48324 SNPs)	Alias 7 58/9.6 4.35 (48299 SNPs)			Alias 8 50.5/18.5 4.08 (63492 SNPs)	
						Alias 8 41.6/12.5 4.21 (61780 SNPs)
Alias 9 37.2/13.7 4.3 (58434 SNPs)	Alias 9 37.2/13.7 4.3 (58445 SNPs)	Alias 9 37.2/13.7 4.3 (58417 SNPs)	Alias 9 37.2/13.7 4.3 (58093 SNPs)	Alias 9 37.2/13.7 4.3 (57930 SNPs)	Alias 9 37.2/13.7 4.3 (57390 SNPs)	Alias 9 45.7/13.7 4.15 (56436 SNPs)
Alias 10 26.3/9.8 4.71 (65857 SNPs)	Alias 10 26.3/9.8 4.71 (65855 SNPs)	Alias 10 26.3/9.8 4.71 (65712 SNPs)	Alias 10 26.3/9.8 4.71 (64863 SNPs)	Alias 10 26.3/9.8 4.71 (64404 SNPs)	Alias 10 26.3/9.8 4.71 (63401 SNPs)	Alias 10 34.6/9.8 4.51 (61654 SNPs)
Alias 11 25/9.8 4.76 (98691 SNPs)	Alias 11 25/9.8 4.76 (98698 SNPs)	Alias 11 25/9.8 4.76 (98537 SNPs)	Alias 11 25/9.8 4.76 (97446 SNPs)	Alias 11 25/9.8 4.76 (96830 SNPs)	Alias 11 not matching	Alias 11 26.8/11.6 4.53 (93063 SNPs)

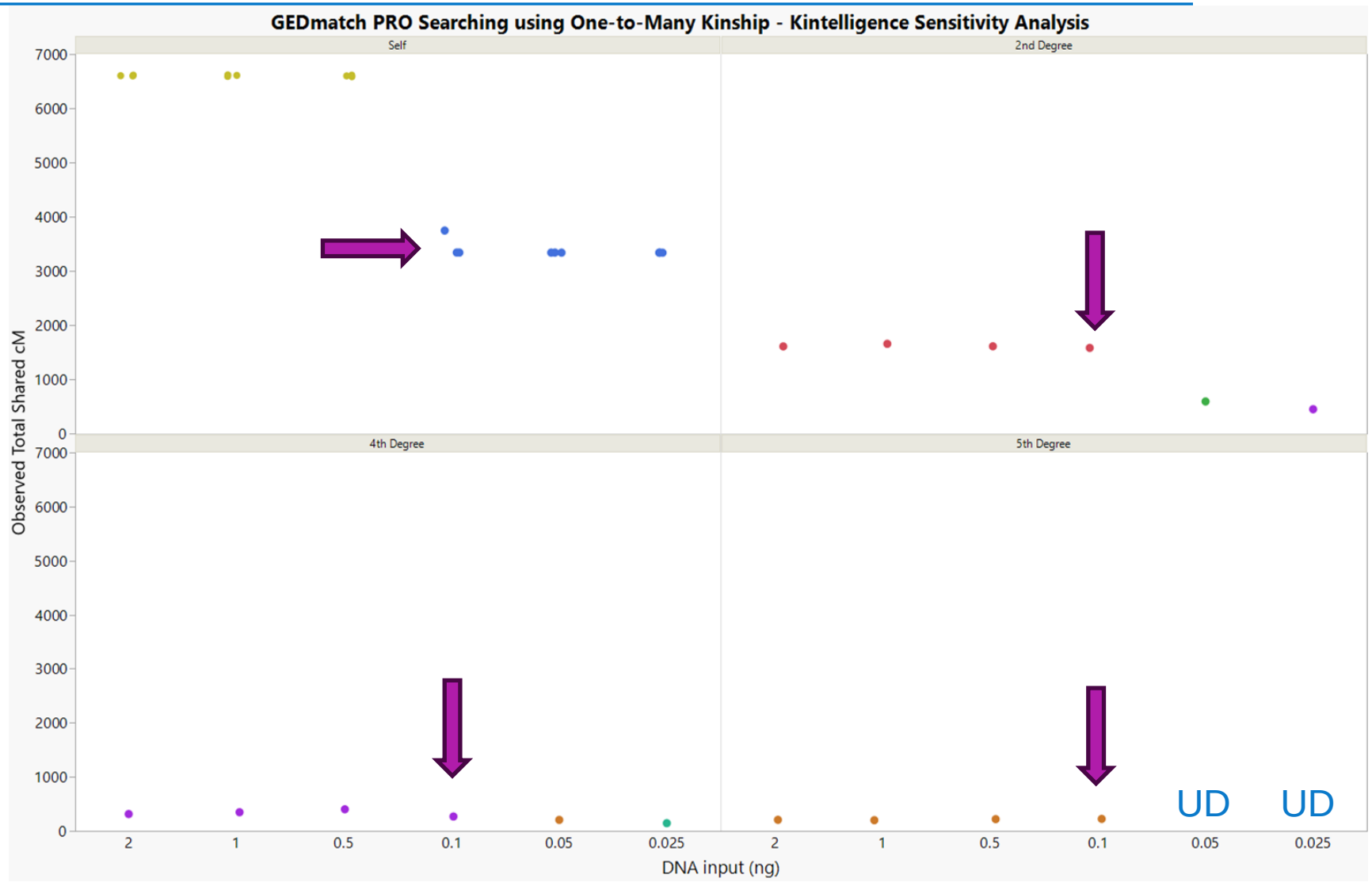
50ng	10ng	2ng	1ng	0.5ng	0.25ng
Alias 12 36.7/10.6 4.30 (51339 SNPs)	Alias 12 36.7/10.6 4.30 (51394 SNPs)	Alias 12 36.7/10.6 4.30 (51337 SNPs)	Alias 16 37/10.7 4.30 (53297 SNPs)	Alias 9 36.9/13.7 4.30 (SNPs)	Alias 19 62.5/11.7 3.92 (71705 SNPs)
Alias 13 36.7/10.1 4.31 (54437 SNPs)	Alias 13 36.7/10.1 4.31 (54471 SNPs)	Alias 14 35.7/11.3 4.33 (53759 SNPs)	Alias 12 36.7/10.6 4.30 (51070 SNPs)	Alias 12 36.7/10.6 4.30 (SNPs)	Alias 20 55.5/11.7 4.01 (71683 SNPs)
Alias 14 35.7/11.3 4.33 (53769 SNPs)	Alias 14 35.7/11.3 4.33 (53812 SNPs)	Alias 10 31.2/8.6 5.8 (81874 SNPs)	Alias 14 35.7/11.3 4.33 (53451 SNPs)	Alias 14 35.7/11.3 4.33 (SNPs)	Alias 11 55/19.1 4.01 (63664 SNPs)
Alias 10 31.2/8.6 5.8 (81940 SNPs)	Alias 10 31.2/8.6 5.8 (81935 SNPs)	Alias 15 30.3/11.9 4.44 (53929 SNPs)	Alias 10 31.2/8.6 5.8 (81342 SNPs)	Alias 17 34.6/10.7 4.35 (SNPs)	Alias 21 54.9/12.4 4.02 (73174 SNPs)
Alias 15 30.3/11.9 4.44 (53941 SNPs)	Alias 15 30.3/11.9 4.44 (53994 SNPs)	Alias 9 29.6/13.7 4.46 (63592 SNPs)	Alias 15 30.3/11.9 4.44 (53628 SNPs)	Alias 25 27.8/12.2 4.51 (SNPs)	Alias 22 49/11.4 4.10 (176299 SNPs)
Alias 9 29.5/13.7 4.46 (63578 SNPs)	Alias 9 29.5/13.7 4.46 (63616 SNPs)	Alias 16 28.6/10.7 4.48 (53583 SNPs)	Alias 9 29.5/13.7 4.46 (63212 SNPs)	Alias 18 27.6/11.1 4.51 (SNPs)	Alias 23 48/13.2 4.11 (71587 SNPs)
Alias 16 28.6/10.7 4.48 (53578 SNPs)	Alias 16 28.6/10.7 4.48 (53628 SNPs)				Alias 4 47.8/11.4 4.12 (52630 SNPs)
					Alias 3 47.5/10.5 4.12 (52764 SNPs)
					Alias 24 47.2/14.4 4.12 (71673 SNPs)

Key:	Kit Alias Shared cM/Longest Segment Generation Estimate (SNP Overlap)

Note: Kit IDs have been redacted for privacy. Identical kit IDs are designated by matching color. Kit IDs displayed in order presented in GEDmatch match list. GSA = GSA BeadChip.

- Anomalous observations when using the donor sample of non-European ancestry and with no known close relatives in database
  - Genome Sequencing matches unusable at <40 cM
  - primarily because of low representation in the database

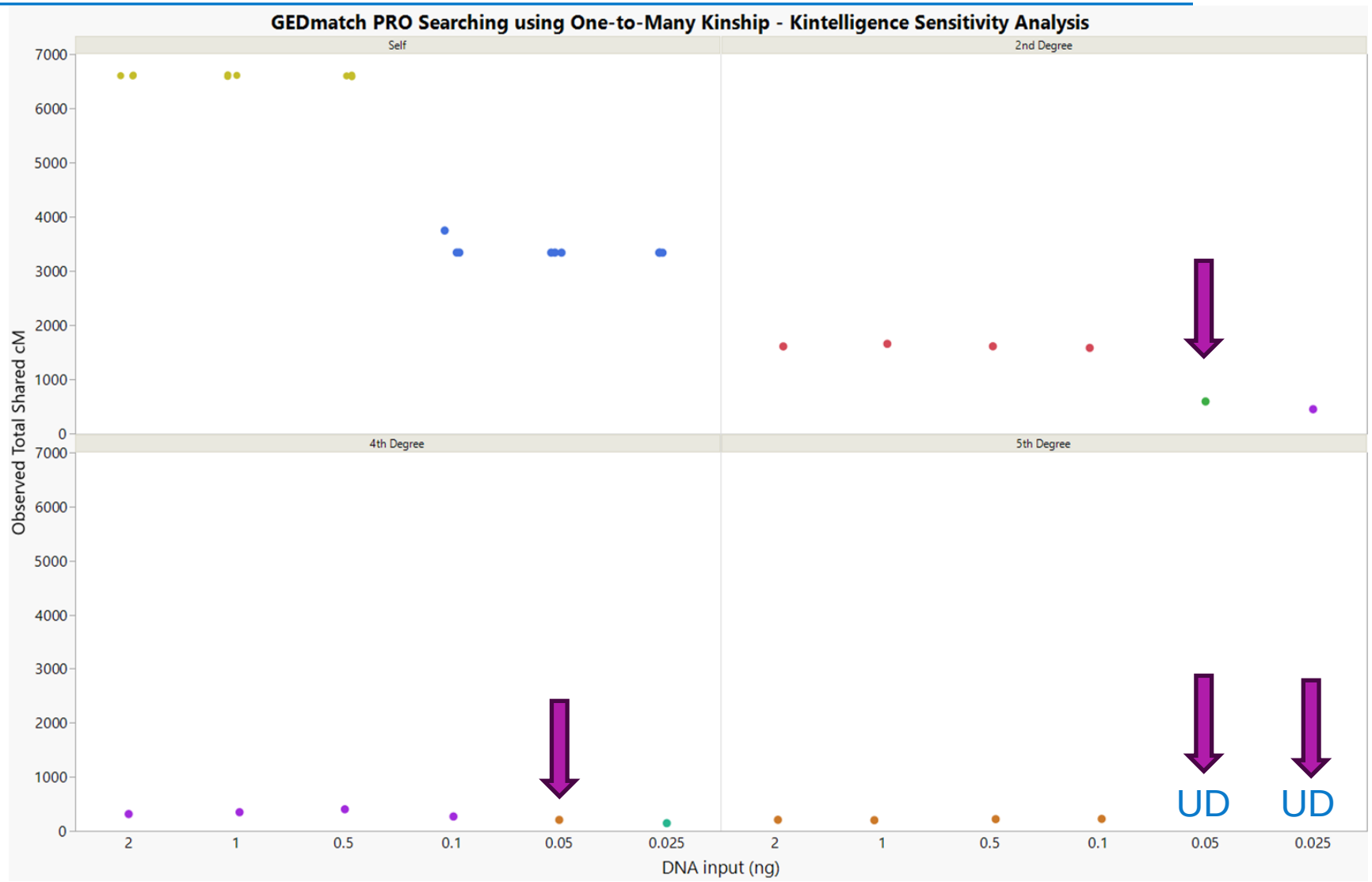
GEDmatch PRO GENEALOGICAL COMPARISONS – Kintelligence Sensitivity Samples One-to-Many Kinship Matching



- Robust and consistent matching to “Self” and documented relatives at expected relationship level when inputs >100 pg

UD: Undetected, Relative no longer in match list

GEDmatch PRO GENEALOGICAL COMPARISONS – Kintelligence Sensitivity Samples One-to-Many Kinship Matching



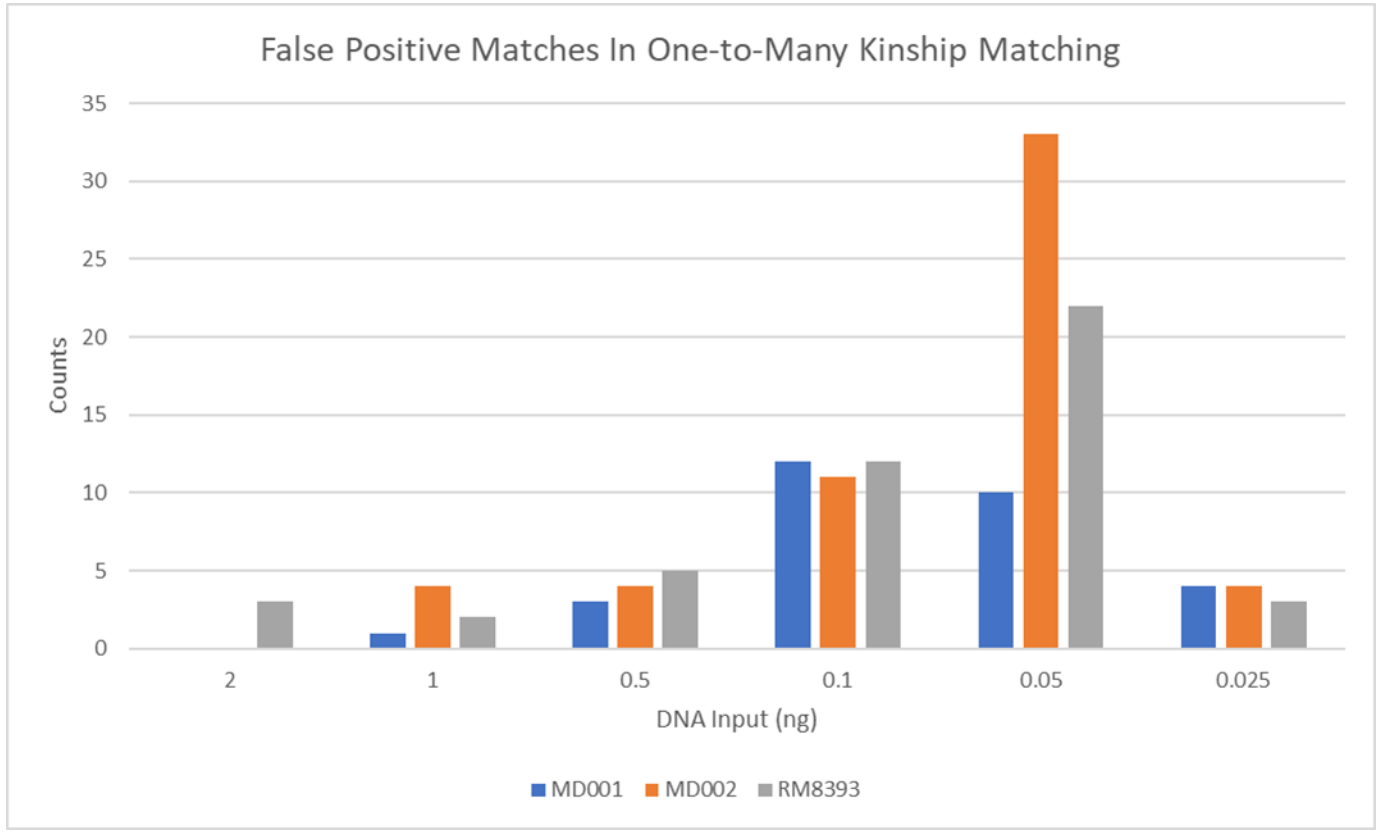
Highest Relationship Probability

- Self
- 1st Degree
- 2nd Degree
- 3rd Degree
- 4th Degree
- 5th Degree
- Expanded Not Detected/Present in All Matches
- Not Detected

- 5th degree/2nd cousin relationship undetected with ≤ 50 pg
- High confidence matches (>200 cM, >6000 SNP overlap) to non-relatives detected at inputs ≤ 50 pg

UD: Undetected, Relative no longer in match list

GEDmatch PRO GENEALOGICAL COMPARISONS – Kintelligence Sensitivity Samples One-to-Many Kinship Matching



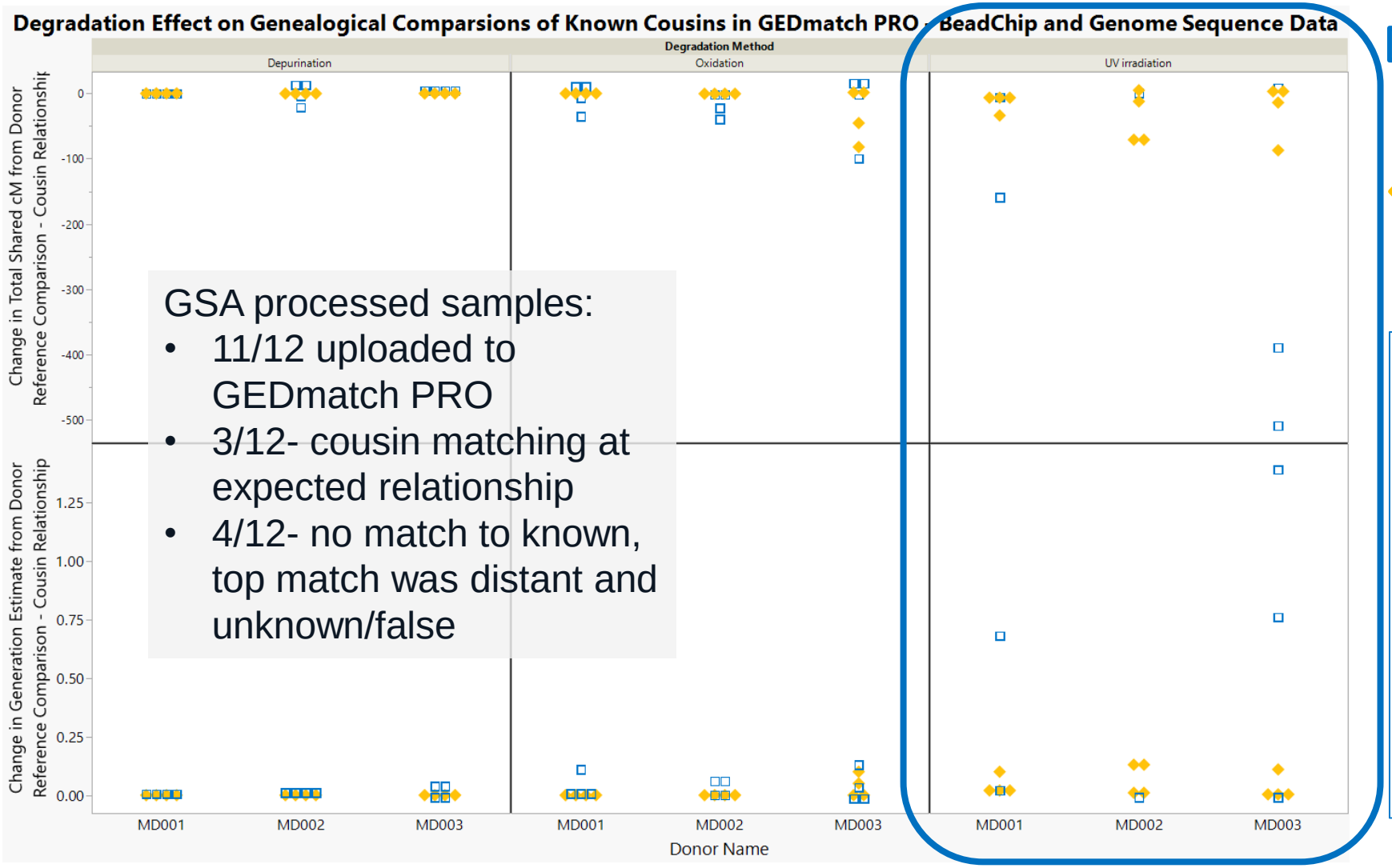
- Counts of “false positive” matches identified in Expanded Match List

  Expanded Match Thresholds:
  140 cM/8000 SNP overlap
  120 cM/9000 SNP overlap

- One-to-One Q matching of match list Kit IDs to donor GSAv2 reference kit indicates no relationship

- None of these kit IDs observed in One-to-Many Segment Based match lists for respective donors

GEDmatch PRO GENEALOGICAL COMPARISONS – GSA and Genome Sequencing Degradation Samples



GSA processed samples:

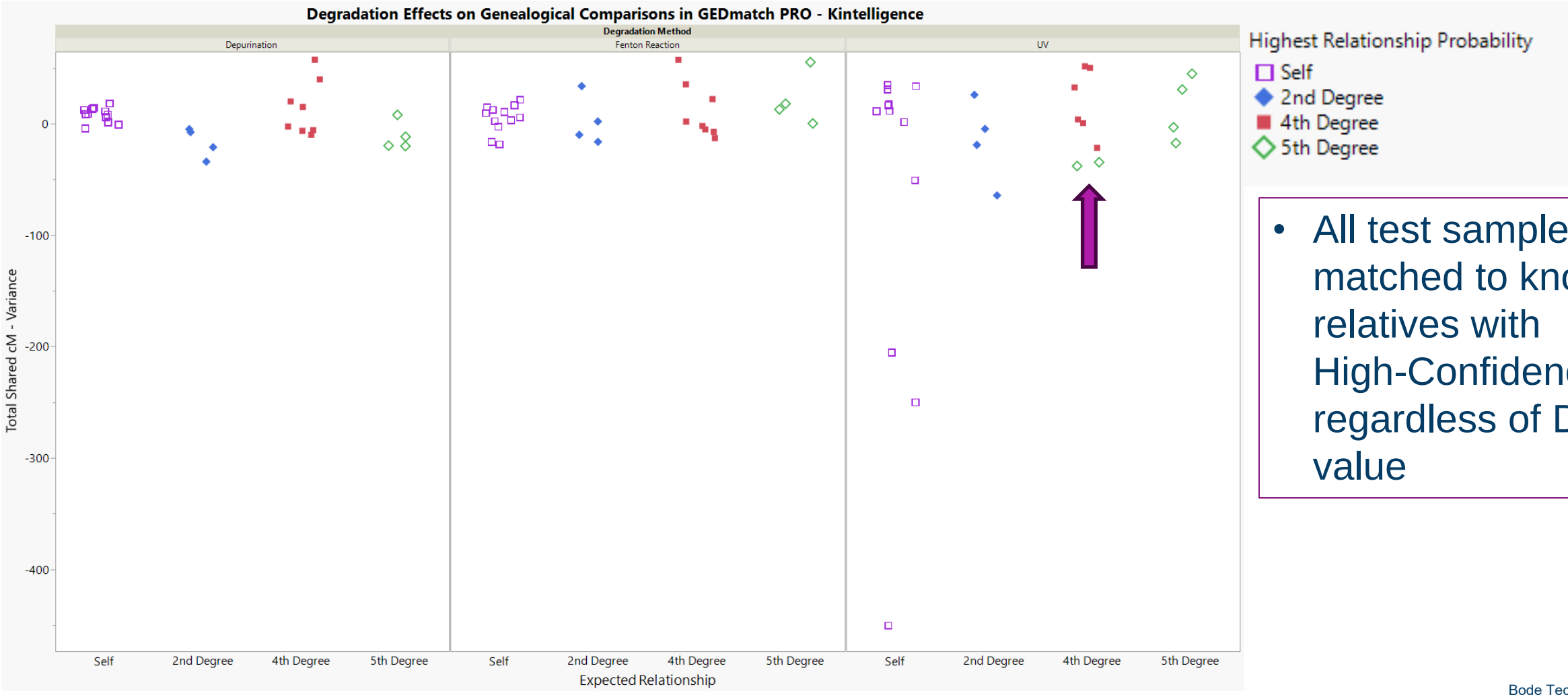
- 11/12 uploaded to GEDmatch PRO
- 3/12- cousin matching at expected relationship
- 4/12- no match to known, top match was distant and unknown/false

Degraded Samples Processed with GSA BeadChip

Degraded Samples Processed with Genome Sequencing

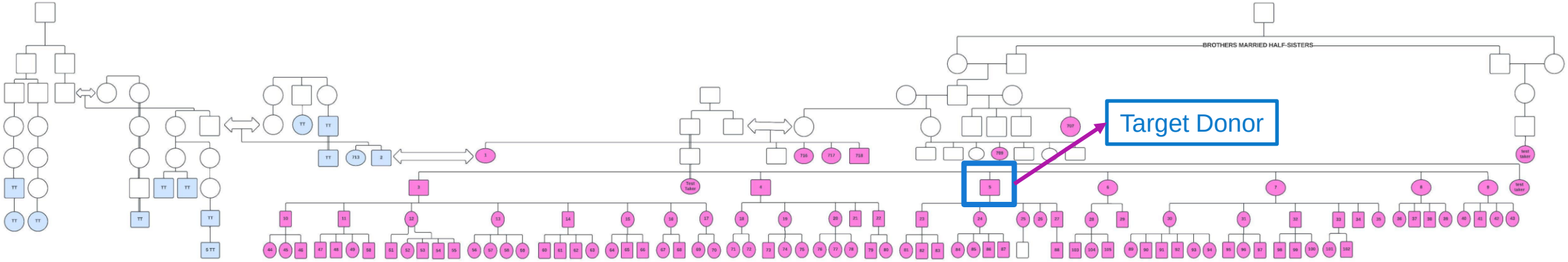
- All degraded samples processed with **genome sequencing** matched to known relatives out to 2nd cousin. Minimal loss in shared cM observed = no effect on relationship estimates.
- With **GSA** processing, UV degraded samples lose cousin matching when DI >4. Reductions in shared cM increased generation estimate, suggesting more distant relationship.

GEDmatch PRO GENEALOGICAL COMPARISONS – Kintelligence Degradation Samples



- All test samples matched to known relatives with High-Confidence, regardless of DI value

APPLICATION TO GENEALOGICAL PROOF PROJECT



		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

APPLICATION TO GENEALOGICAL PROOF PROJECT

		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	255	ND

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

CONCLUSIONS – SAMPLE PROCESSING CAPABILITY

- All three technologies sensitive to decreasing DNA inputs
- Degradation as measured by DI correlated with lower call rate recovery for both GSA and Genome Sequencing
  - GSA - DI >4 becomes problematic
  - Genome Sequencing - DI > 61, call rate drops, but led to minimal impact on matching results ***(> 1 million SNPs)***
- Kintelligence assay chemistry robust to DI with optimal input
- Corroborating the growing body of research for these methods[1-4]

1) DOI:10.1016/j.fsigen.2021.102625
2) DOI:10.1101/2022.10.28.514056
3) DOI:10.1101/2022.10.10.511614
4) DOI:10.1111/1556-4029.15469

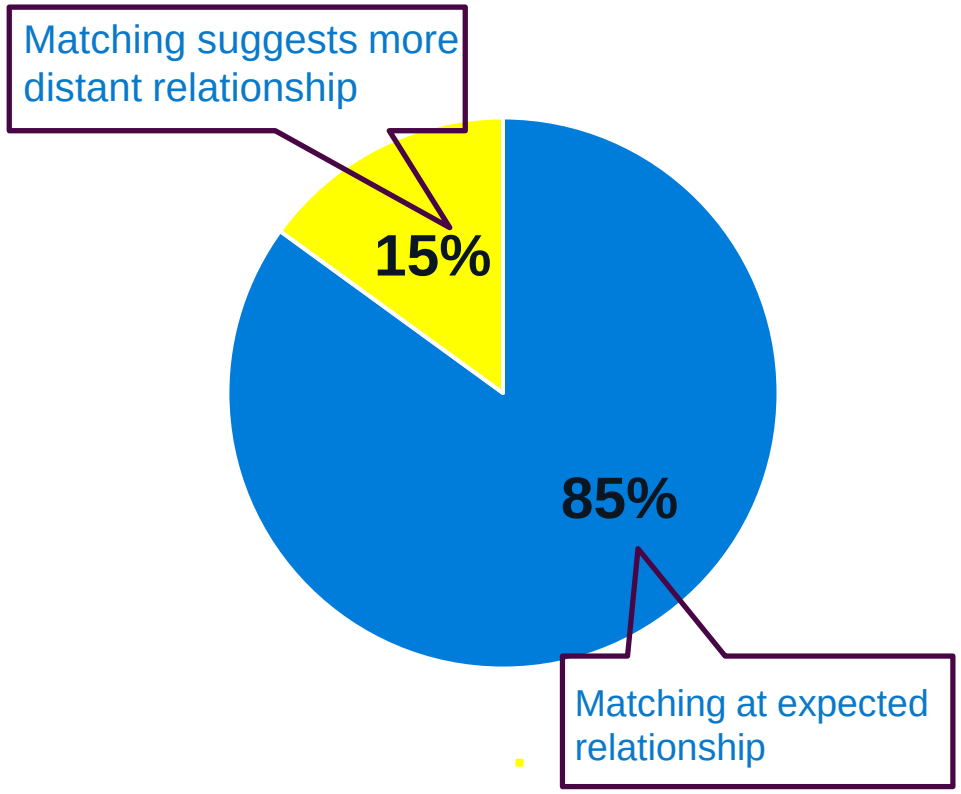
CONCLUSIONS – GENEALOGICAL VALUE ASSESSMENT

- When known relatives present in GEDMatch database:
  - Reliable matching out to 2nd cousin level with forensic level DNA inputs with all technologies
  - Reliable matching out to 1C1R with degraded samples when using Kintelligence High Confidence Thresholds
- Kits generated with GSA or Genome Sequencing functionally identical
- Limitations observed with GEDmatch searching/matching
  - Genome Sequencing SNP set appears to match best to 23andMe commercial kits
  - Inconsistencies between kits identified with Kintelligence vs GSA/sequencing

CONCLUSIONS – GENEALOGICAL VALUE ASSESSMENT

- Limitations observed with GEDmatch searching/matching
  - Kintelligence shared cM/longest segment length trended toward lower-than-expected values
    - Shared cM totals are a primary indicator of relationship level for genealogists, lower values can be misleading to researchers
  - Donors of Non-European biogeographical ancestry – no common top matching kits between GSA, Genome Sequencing, and Kintelligence in GEDmatch PRO

Kintelligence: 98% of comparisons matched to expected documented relatives



CONCLUSIONS – GENEALOGICAL VALUE ASSESSMENT

- Caution is advised with Kintelligence matching
  - 4th and 5th degree (1C1R and 2nd cousin+) matches at Expanded Match thresholds are unreliable,
    - One-to-One Comparisons against GSAv2 donor reference kits showed <10 cM, more often no shared content
  - Kintelligence kits are not yet compatible with third party tools for comparison
    - No visualization tools similar to chromosome browser are yet available for genealogical comparisons with Kintelligence kits

CHOSEN WORKFLOW:
MPS FIGG Offerings at Bode

New service offerings

FORENSIC INVESTIGATIVE GENETIC GENEALOGY

- QIAGEN ForenSeq Kintelligence system
- Element Biosciences AVITI System
- Internally validated to ISO 17025
  - And accredited under a scope expansion covering SNP analysis!!!

ANAB
ANSI National Accreditation Board

March 21, 2024

Erin Sweeney
Bode Cellmark Forensics, Inc.
dba Bode Technology
10430 Furnace Road
Lorton, VA 22079

Dear Director Sweeney,

Congratulations! On March 15, 2024, ANAB granted an extension of scope in the Field of Forensic Testing in the Biology discipline at the 10430 Furnace Road, Lorton, VA location. This decision was based upon the documentation provided in the assessment report and in accordance with the recommendation of the Team Leader. ANAB is satisfied that your organization has met or exceeded the accreditation requirements and requirements of your own documented management system for this extension of scope.

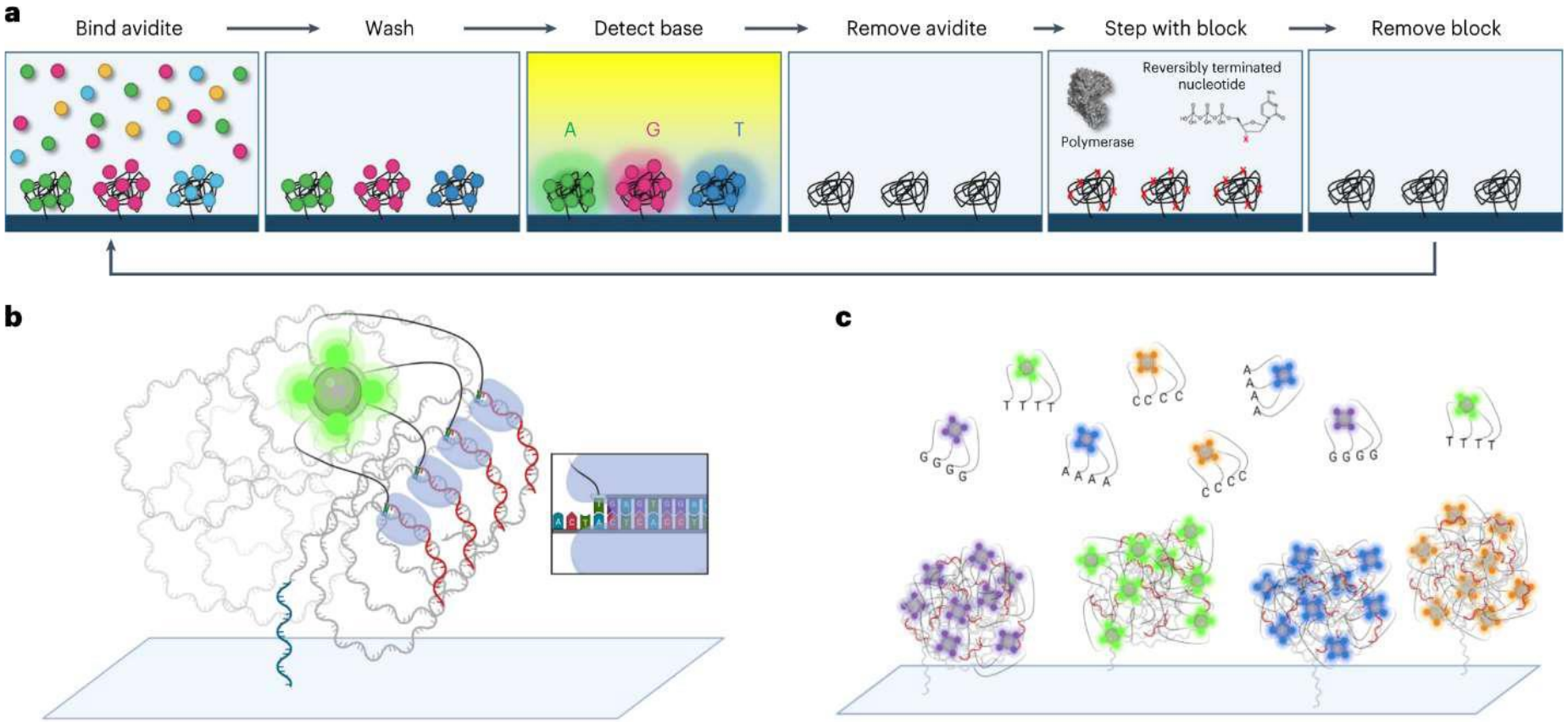
The report was provided to you during the assessment activity. An electronic version of the updated Scope of Accreditation document is included with this letter.

Element Biosciences AVITI

What is Aviti and why was it selected?

- Unparalleled performance and affordability in a benchtop sequencer
  - Industry leading accuracy >90% >Q30
    - Illumina NovaSeq >75% >Q30
    - Routinely see Q-scores >Q40 (99.99% accuracy)
    - Pushing Q50
  - Low duplication rate = greater library complexity
  - Amenable to the wide variety of library prep kits on the market
    - Can process previously constructed Illumina libraries
  - Offers longer sequencing reads compared to Illumina NovaSeq (2x300 vs 2x150)
  - Instrument cost and maintenance costs significantly lower than Illumina NovaSeq for equal amounts of data
  - Two parallel runs or independent operation

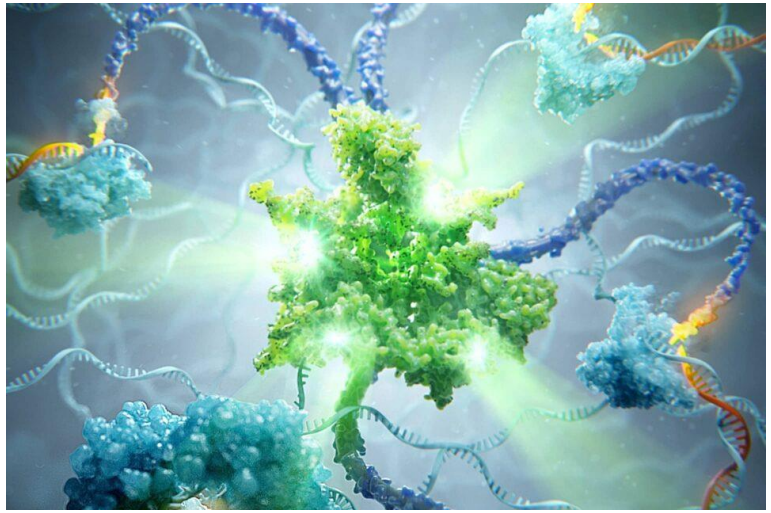
AVITI Sequencing Chemistry



- A) A complete cycle of avidity sequencing.
- B) A single avidite interaction with multiple DNA copies within a polony.
- C) Many avidites bound to multiple polonies on the flow cell surface.

AVITI Validation Highlights

- Validated according to internal FBI/SWGDAM guidelines
- >600,000 SNPs
- >95% concordant to known even in comprised samples
  - Trio Degradation Index of 99
  - Human DNA content down to 4% of total DNA
  - 19 pg limit of detection
- 12 sample multiplexing
- Casework Processing has begun!



ACKNOWLEDGEMENTS

- Bode's Genealogy Team
  - Melinde Byrne
  - Teresa Vreeland
- Gene by Gene Team
  - Arjan Bormans
  - Kendle Pryor
  - Lisa Longoria
  - Kate Leeman
- Qiagen/Verogen Technical Support
  - Melissa Kotkin
  - Emma Katzman
- Bode Validation Team
  - Anna Salmonsen
  - Sarah Schmitz
  - Kristen Naughton

CONTACT

Sarah Cavanaugh
sarah.cavanaugh@bodetech.com

Robert Bever
Robert.bever@bodetech.com

BODE TECHNOLOGY™

Answers With Confidence and Accuracy