Comparison of ForenSeq® Kintelligence and Whole Genome Sequencing in Searching for Relatives in GEDmatch PRO™

S.M. Radecke[1], J. Antunes[1], G. Padmabandu[1], K.M. Stephens[1]

[1]QIAGEN, San Diego, CA USA

Introduction

- Whole genome sequencing is a popular alternative to microarrays for highly degraded and/or low input samples for FGG[1]
- ForenSeq Kintelligence 10,230 SNP panel was developed concurrently with GEDmatch PRO's One-To-Many Kinship Tool that maximizes kinship SNPs to identify up to 5th degree relationships[2-3]

Materials and Methods

- WGS and ForenSeq Kintelligence libraries were generated with V024 Samples (Fig. 1): 1ng intact, 1ng with DI 6, 100pg intact, or 100pg with DI 6 (Table 1)
- Sequencing was done on the NovaSeq® at 2x151 with a target coverage of 30x (WGS) or 3-plex on MiSeq® FGx 2x151 (ForenSeq Kintelligence)
- WGS samples were aligned with bwa mem and SNPs called using bcftools
- For WGS data, a custom utility was used to type SNPs in loci that are accepted in GEDmatch (GM) and in a custom set of loci
- ForenSeq Kintelligence data was analyzed with UAS automatic analysis
- Heterozygosity for ForenSeq Kintelligence was taken from UAS Sample Result
- Specific loci were filtered across all bioinformatics workflows (Fig. 3)

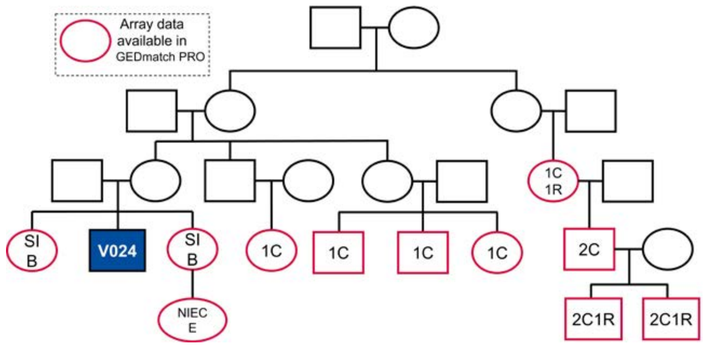
Results

- Bioanalyzer traces show degraded DNA fragments (<400bp in length) (Fig. 2)
- WGS coverage remained close to 30x for all samples (Table 1)
- WGS Degraded 100pg sample had the lowest number of SNPs called (Fig. 4A), which required removal of all variant calling filters (Fig. 4B)
- All ForenSeq Kintelligence results had >9000 SNPs typed (Fig. 4C)
- WGS heterozygosity ratio was close to expected heterozygosity ratio for Europeans4 (Table 1)
- WGS with segment matching (Table 2)
- The custom utility performed best overall and was able to identify true matches but only up to 1st degree for Degraded using GM loci and up to 4th for Intact
- For Degraded 100pg, applying no filters and filtering to array loci returned highest number of true positives
- ForenSeq Kintelligence with kinship (Table 2)
- 1ng samples matched all expected 1st-5th relationships and one 6th degree
- Degraded 100pg matched up to 4th degree
- False positives remained low (Fig. 5)

Conclusions

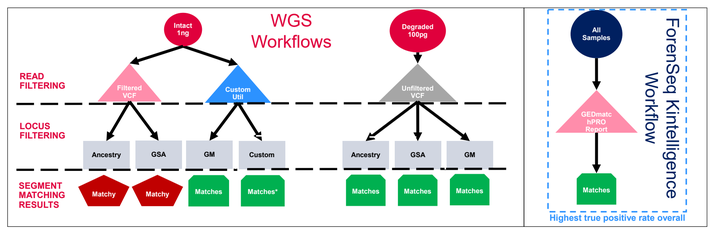
- Increased computation time and resources to process WGS data compared to ForenSeq Kintelligence. Sample-specific WGS data processing; not reproducible across samples. Unknown expectations for identifying high order relationship degrees
- ForenSeq Kintelligence workflow is simple and is reproducible across samples. Expectations for matches are known (e.g., up to 5th degree). Consistent results across sample types.
- Disease SNPs are excluded from the ForenSeq Kintelligence assay, but must be removed bioinformatically with WGS

Figure 1.



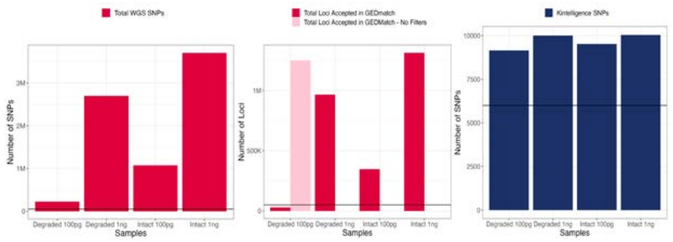
Pedigree of a family with various kit types available in GEDmatch PRO

Figure 3.



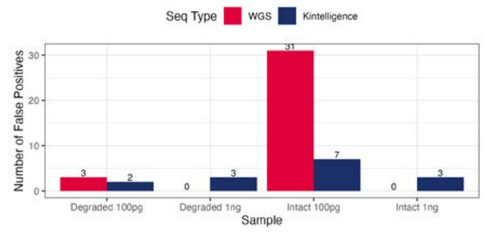
Bioinformatics workflows implemented for each type of sequencing (WGS [left] or ForenSeq Kintelligence [right]). ("Ancestry" includes Ancestry and 23AndMe loci; GSA=Global Screening Array; GM = Loci accepted in GEDmatch PRO). *One-To-Many Kinship Tool was used

Figure 4.



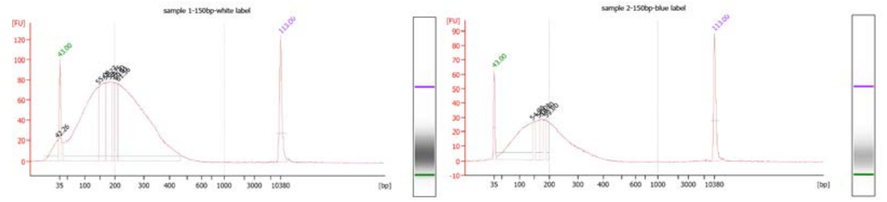
SNP counts per sequencing platform per sample (WGS is pink, ForenSeq Kintelligence is blue)

Figure 5.



False positive matches per sample, per sequencing platform

Figure 2.



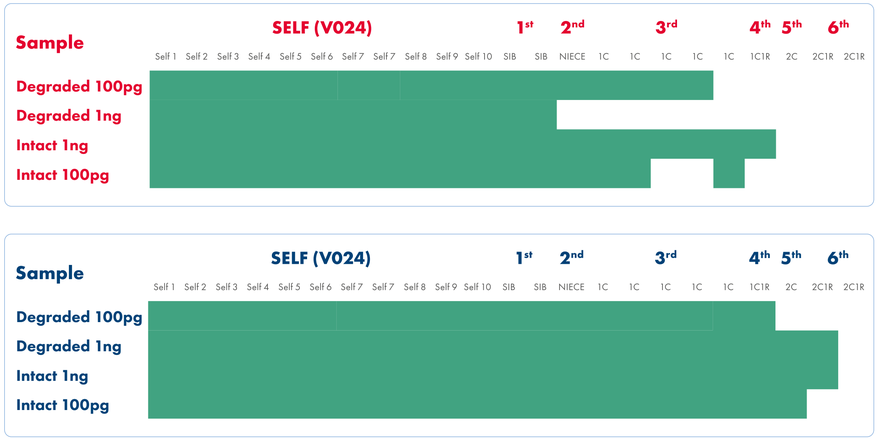
WGS bioanalyzer traces of Degraded 1ng (left) and Degraded 100pg (right)

Table 1.

Sample	Degradation Index (DI)	DNA Input Amount (ng)	WGS Average Coverage	WGS Heterozygosity Ratio[4]	ForenSeq Kintelligence Heterozygosity (%)
V024	1	1.0	30.0	1.79	47.5
		0.1	37.2	2.18	39.9
	6	1.0	21.0	1.84	46.9
		0.1	28.2	4.66	33.4

Average coverage and heterozygosity of four samples processed with whole-genome sequencing (WGS) or ForenSeq Kintelligence on the MiSeq FGx

Table 2.



True positive match matrix: WGS (pink) or ForenSeq Kintelligence (blue) sample per row vs expected matches per column. Green is a match, white is no match

REFERENCES
1. Tillmar A, et al. Whole-genome sequencing of human remains to enable genealogy DNA database searches - A case report. Forensic Sci Int Genet. 2020 May;46:102233. doi: 10.1016/j.fsigen.2020.102233. Epub 2020 Jan 17. PMID: 31981902.
2. Snedecor J, et al. Fast and accurate kinship estimation using sparse SNPs in relatively large database searches. Forensic Sci Int Genet. 2022 Nov;61:102769. doi: 10.1016/j.fsigen.2022.102769. Epub 2022 Aug 27. PMID: 36087514.
3. Verogen. SNP Typing in Universal Analysis Software and Kinship Estimation with GEDmatch PRO. Technical Note.March 2021.
4. Samuels DC, et al. Heterozygosity Ratio, a Robust Global Genomic Measure of Autozygosity and Its Association with Height and Disease Risk. Genetics. 2016 Nov;204(3):893-904. doi: 10.1534/genetics.116.189936.

Sample to Insight