Developmental Validation of the Illumina Infinium Assay using the Global Screening Array (GSA) on the iScan System for use in Forensic Laboratories

David A. Russell, MS; Carmen Reedy, PhD ; Stephen D. Turner, PhD; Stephanie Guertin, PhD; Mary Heaton, MS; Jessica Bouchet, BS; Michelle Peck, MFS; Erin Gorden, MFS; Elayna Ciuzio, MS; Christina Neal, MS

Scan for more information

ABSTRACT

This developmental validation study aimed to provide the forensic community with a validated microarray-based genome-wide SNP genotyping workflow for DNA inputs consistent with forensic sample types. Most labs performing microarray-based SNP genotyping are clinical and lack a forensic framework. Our approach was to evaluate the Infinium Global Screening Array-24 v3.0 Kit [GSA, Illumina] on the iScan array system [Illumina], focusing on shifting applicability from clinical laboratories to the forensic community. To date, there is no existing guidance on how to validate microarray-based genome-wide SNP genotyping within the forensic community. This validation was performed guided by the standards (FBI QAS, SWGDAM) accepted by the forensic community with the intended application to challenging cases crime labs encounter. This will allow greater confidence in the investigative leads developed by law enforcement agencies, as the data used to develop such leads will have been generated under the same scientific standards already established within other human identification methods.

The following studies were performed as part of the developmental validation: sensitivity, precision and accuracy (repeatability and reproducibility), mixtures, and degradation. As well as: species specificity, contamination, case-type samples, and array stability. Throughout all studies, except for the case-type study, extensively characterized human genomes for which high-confidence variant calls are known were used. Concordance was assessed using the NIST/Genome-in-a-bottle (GIAB) sample call sets as truth data for each sample.

For the sensitivity study DNA inputs from 200 ng to 0.2 ng were assessed using NA12878 [Coriell]. This evaluation demonstrated call rates of >99% and >95% for inputs ≥1 ng and 0.2 ng, respectively. Additionally, results were highly concordant at DNA inputs as low as 0.2 ng: <0.001% discordance for all replicates down to 1 ng, and <0.5% discordance at 0.2 ng. For the precision and accuracy (repeatability and reproducibility) studies, when comparing sample genotypes to the NIST/GIAB sequencing data, average concordance rates were 99.2% across all samples. Comparing duplicate samples to each other showed a concordance rate of >99.8% across all samples. The repeatability and reproducibility studies demonstrated reliable and consistent call rates and high concordance, regardless of operator.

In the mixture study, the profile generated from samples of a 3:1 mixture of NA24631 to NA12878 (major to minor contributor) was on average 98.85% concordant with gold standard data for NA24631; the 9:1 mixture ratio was on average 99.99% concordant. This demonstrated that at a mixture ratio of 3:1 or greater between the major and minor contributors, the resulting major contributor's profile is accurate. The data generated during the degradation study was unexpected; samples known to have been severely degraded produced genotyping data with call rates similar to pristine samples. However, examination of concordance data showed that the genotype calls were less accurate as the samples became more degraded. Lower heterozygosity was observed in the degraded samples (average of 7%) compared to the GIAB population set (average of 17.3%). This suggests that allelic dropout is responsible for the noticeable shift to false homozygous calls, accounting for the high call rate yet discordant SNP data.

This validation characterized the performance of forensic samples on the GSA platform, providing valuable insight into what forensic practitioners can expect from microarray-based genotyping data. Thresholds for call rate, heterozygosity, and overall signal (fluorescence) were established to assess the data and determine suitability of samples for the workflow. Ultimately, this work will provide a better understanding of how microarray systems will operate in the forensic community as the application of this technology to more cases continues to grow.

Acknowlegements and Contact Information

This work is supported by Signature Science, LLC through the Center for Advanced Genomics.

David Russell, MS
Signature Science, LLC
1670 Discovery Drive
Charlottesville, VA 22911
drussell@signaturescience.com

RESULTS

Precision & Accuracy

Accuracy statistics show the call rate and concordance for all sample replicates (Table 1). Call rate for all samples was >**99%**. Concordance at called sites between the test and GIAB samples was >**99.9%** for all samples.

Sample ID	Rep	Call Rate (%)	SNPS Called	Concordance at Called Sites	Missing SNPs
HG001	rep1	99.29%	593,783	99.94%	388
HG001	rep2	99.30%	593,783	99.94%	370
HG002	rep1	99.93%	548,279	99.93%	305
HG002	rep2	99.94%	548,279	99.93%	293
HG005	rep1	99.94%	556,487	99.95%	260
HG005	rep2	99.93%	556,487	99.95%	311

Table 1: Call rates for each input level, averaged across all three replicates.

Sensitivity

This study established a range of DNA inputs for the Infinium assay with pristine single source DNA. For inputs ranging from 1 ng to 200 ng, the call rate was greater than **99%**. At an input of 0.2 ng, more than **97%** of the SNPs were called, indicating high sensitivity. Further, comparison between each DNA input level and the manufacturer-recommended input of 200 ng produced highly concordant data (Table 2), with rates of >**99.9%** for inputs ≥ 1 ng and >**99.5%** for inputs of 0.2 ng. This study showed that high numbers of accurately called SNPs are generated for all inputs greater than or equal to 0.2 ng.

Input (ng)	Concordance at Called Sites	Missing SNPs	Missing SNPs Rate (%)
40	100.000%	816	0.128%
20	99.999%	861	0.135%
8	99.996%	1,484	0.233%
2	99.995%	1,896	0.298%
1	99.965%	4,338	0.682%
0.2	99.685%	18,366	2.888%

Table 2: Concordance for each input, averaged across all three replicates

Contamination Assessment

Assessing the DNA Negative Samples gave no indication that contamination was introduced during the procedure. If contamination did occur in the DNA blanks, the total intensity would be expected to be greater than the background threshold. All DNA negative samples fell below this threshold, indicating that the resulting calls were due to background noise.

Species Specificity

It was expected that non-human samples would have a low call rate, <90%. The call rates for the species study were similar in range compared to the DNA blanks used in the contamination study. Non-human samples had a total fluorescence intensity much less than that of human and human-like (RH monkey) samples. In Figure 1, the columns represent the total intensity of the fluorescence, and the line the call rate.

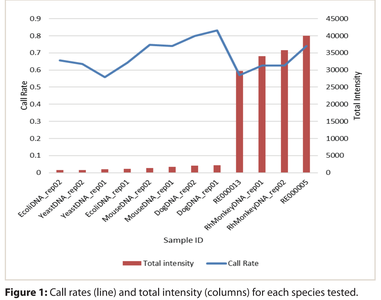

Figure 1: Call rates (line) and total intensity (columns) for each species tested.

Two human samples (RE00005, RE000013) that were low copy number (LCN) and degraded are included to demonstrate the expected intensity values for human DNA samples with similar call rates. As observed, the RH monkey DNA samples have similar intensities and call rate to human DNA samples.

Mixture Assessment

Assuming the unknown contributor is the major contributor, the 9:1 and 3:1 mixture ratios were compared against the neat sample NA24631(M) without doing any "subtraction" of the Known female contributor. The 9:1 ratio had a **99.99%** concordance with the neat sample, and the 3:1 ratio had a **97.9%–99.8%**. This study has shown that in instances where the mixture ratio is 3:1 or greater, the produced genotype of the major contributor is accurate as is, without removing the known minor contributor.

		Human	Male	
Mixture Ratio (M:F)	Sample ID	Quantity	Quantity	Mixture Index
1:0	mix_01	59.0704	41.6288	1.42
9:1	mix_02	53.1707	36.9877	1.44
3:1	mix_03	62.8300	32.0025	1.96
1:1	mix_04	63.6096	23.6591	2.69
1:3	mix_05	61.8043	10.8962	5.67
1:9	mix_06	59.7144	4.6121	12.95
0:1	mix_07	60.4859	0.00	0.00

Table 3: Mixture quantification metrics for each ratio SNP genotyped

Mock Casework Sample Types

In this study, the sample DNA input into SNP genotyping ranged from 0.13-62.9 ng. Many of these forensic samples yielded call rates <**95%**, which is unlike the results observed in the sensitivity study using pristine samples (0.2 ng DNA input). This is expected, as other factors afflict forensic casework samples (e.g. DNA damage, degradation, inhibition) that would not be observed in high quality reference samples and are likely impacting the call rates. Samples with call rates from **60-90%** may benefit from re-extraction and/or optimization of extraction protocols for specific sample types (e.g. bone samples); however, this represents future work and was not performed as part of this validation study.

Samples producing a call rate of 70% or above were uploaded into GEDmatch and the ability to obtain possible relatives for investigative leads in FGG were assessed. Potential relatives (representing investigative leads for law enforcement) were obtained for all the uploaded samples. The distance of relatives *[**M**ost **R**ecent **C**ommon **A**ncestor* (MRCA)] to the unknown forensic samples ranged from 2.4-4 generations. Biogeographical ancestry inference was performed for, the sample with the highest call rate (**94%**). The results suggest the individual is of Western/North European descent, which is consistent with likely ancestry of the matches obtained from GEDmatch.

Sample Type	Quant (ng/µL)		DNA Input for Genotyping (ng)	Run 1		Closest GEDmatch Relative		Run 2	
	Quantiplex Pro Human Target	Quantifiler Trio Small Target		Call Rate	Heterozygosity	Total cM	Generations to MRCA	Call Rate	Heterozygosity
Proficiency sample (non-sperm fraction)	0.916	2.54	3.663	88%	19%	108.844	3.5	96%	19%
Proficiency sample (sperm fraction)	1.092	3.03	4.366	94%	17%	536.253	2.4	99%	17%
Differential RB	0	0							
Differential RB	0	0							
Blood stain (~10 years old)	0.213	0.67	0.852	69%	18%			74%	19%
Chewed gum (~5 years old)	0.461	1.57	1.845	87%	19%	71.611	3.8	85%	18%
Proficiency sample	2.403	6.15	9.61	83%	19%	113.23	3.5	98%	18%
Proficiency sample	1.409	4.2	5.638	80%	20%	106.973	3.5	95%	18%
Tooth	15.729	33.35	62.918	90%	18%	84.225	3.7	100%	17%
Fingernail	0.52	1.35	2.082	72%	18%	56.559	4	86%	17%
Fingernail	0.001	0							
Fingernail	0.001	0							
Buccal sample/Bode collector (non-FTA)	0.089	0.37	0.356	0%	17%			57%	19%

Table 4: Mock casework samples with quantification, Input metrics, and GEDmatch results.

DNA input for genotyping was calculated using the quantification values from the Investigator Quantiplex Pro human target. Grey boxes represent non-applicable fields. Run 2 statistics are from when the samples were genotyped again in the reproducibility study.

Degradation

The degradation study presented here shows that the call rate alone is not sufficient to assess the overall data quality. Furthermore, the observed heterozygosity of the samples also explained the rebound in the observed call rate. Concordance showed that the sample accuracy was decreasing as the sample became more degraded. The heterozygosity shows that as the sample became more degraded, the less heterozygous (more homozygous calls) it became. This supports a hypothesis that increased call rates corresponded with an enrichment of erroneous homozygous genotype calls or put differently, increase in a genotype switch to homozygous for the alternative allele, rather than a no call.

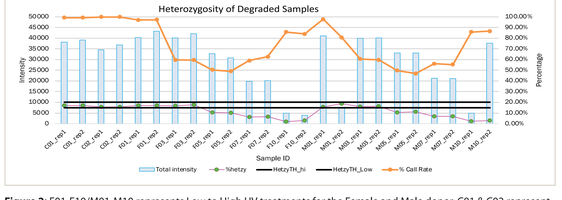

Figure 2: F01-F10/M01-M10 represents Low to High UV treatments for the Female and Male donor. C01 & C02 represent the control DNA of F&M at a target input of 20ng/µl

SAMPLE ASSESSMENT METRICS

Sample Signal Total Intensity

Sample assessment for this study was based on the call rates and signal intensities (p95 Red & p95 Grn). The "p95 Grn", and "p95 Red" metrics represent the 95th percentile of B allele and A allele intensity respectively. Total intensity is the sum of these metrics. Due to the way the iScan system uses relative fluorescence signal to call genotypes, non-specific binding or background noise may give rise to SNPs being called by chance as the resolution algorithm tries to maximize signal detection. In the absence of DNA, this study demonstrates that there is inherent background noise in the detection system and the call rate cannot be considered an absolute indication of sample presence/absence.

Using the 15 DNA blanks and RBs from the various studies in this validation, an average baseline noise (BLN) level was established by using the average total intensity (965). The intensity threshold was then calculated by adding three times the standard deviation (SD = 425.93) to the BLN (1,390.93), and then rounding up the value to the next hundred. Thus, the IT was set at 4200 (Figure 3).

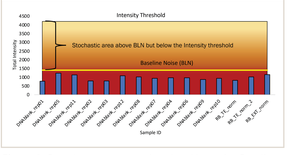

Figure 3: Sample assessment using signal intensity. Baseline noise threshold set using DNA blanks.

Observed Sample Heterozygosity

Using the 1000 Genomes study, heterozygosity statistics were calculated from 2,504 samples from 26 populations in order to have a proper dataset to establish an expected heterozygosity range for the human population. The calculated heterozygosity for humans on the GSA is on average around 17.3%, with a SD of 0.6844%. The range was 15.2% to 19.4%. The range was rounded to the next lowest or highest whole value, setting the upper and lower threshold limits at 20% and 15% respectively (Figure 4).

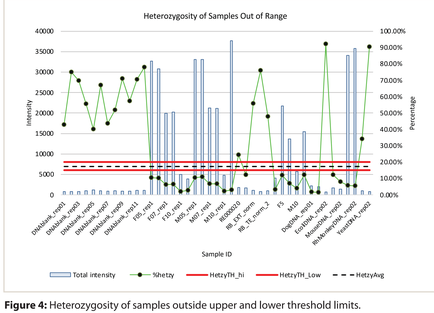

Figure 4: Heterozygosity of samples outside upper and lower threshold limits.