DEVELOPMENTAL VALIDATION OF THE ILLUMINA INFINIUM ASSAY USING THE GLOBAL SCREENING ARRAY (GSA) ON THE ISCAN SYSTEM FOR USE IN FORENSIC LABORATORIES

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

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.