EVALUATION OF STANDARD REFERENCE SAMPLE TYPES FOR NEXT GENERATION SEQUENCE-BASED GENOTYPING

Seth A. Faith, Mark E. Hester, Dan M. Bornman, Angela Minard-Smith, Brian A. Young
Battelle Memorial Institute, 505 King Avenue, Columbus, Ohio 43201, USA

Next-generation DNA sequencing (NGS) technologies are being evaluated by numerous investigators as a platform for forensic DNA analysis. An attractive feature of NGS is its ability to genotype multiple marker types in a single process. However, taking advantage of multiplex analysis by NGS will require custom sample preparation to achieve DNA preparations that are both suitable for NGS library preparation and for achieving appropriate depth of coverage for the different marker types. Genotyping methods previously developed in our laboratory employ either PCR enrichment for short tandem repeat (STR) loci or whole genome sequencing (WGS) for single nucleotide polymorphisms (SNPs). The current study evaluates WGS for both STRs and SNPs for several common reference sample types. Saliva, buccal swab, and venous blood preserved on FTA cards were collected from anonymous donors and subjected to DNA extraction. DNA extraction techniques were assessed for optimal methods to provide sufficient yield and quality requisite for NGS. Next, libraries were prepared for all three sample types using ultra-sonication and enrichment for producing 200-300 bps fragments. Commercially available library prep kits were used and sequencing was performed using an Illumina HiSeq2000 that generated over 200 million sequence reads (>40 GBps data) per lane. Raw sequence data was evaluated for the quantity of sequencer-ready DNA obtained and the percent human genome coverage achieved on a per-swab, punch or saliva sample basis. Further, because WGS is unbiased to input DNA, the proportion of human versus microbial DNA for buccal swabs and saliva was determined based on genome alignments of sequencer reads in order to express the amount of sequencer capacity consumed by microbial DNA. Lastly, an estimate of reference sample (swabs and punches) quantity required to provide high quality STR and SNP genotypes was calculated. Based on these results, we provide a foundational recommendation for methods to be used when evaluating reference samples by next-generation DNA sequencing technology. ⌘