In recent years, an increasing number of technologies and platforms suitable for short tandem repeat (STR) and single nucleotide polymorphism (SNP) genotyping have become available. However, since information directly comparing the different technologies is limited, members of the DNA Sequencing (DSRG) and Fragment Analysis (FARG), and Nucleic Acid (NARG) research groups of the ABRF, launched a joint Genotyping Pilot Study. This study involved participants from all three research groups, and the technologies surveyed in the members’ laboratories included microfluor/STR-PCR analysis, direct DNA sequencing, real-time PCR, SNaPshot, Scorpion/Amplifluor, pyrosequencing, and heteroduplex analysis.

In this paper we will focus specifically on the DNA sequencing and SNaPshot aspects of the study. Results will be evaluated for the quality and concordance of the genotyping calls reported by each participant. A detailed cost analysis and a comparison of the advantages and limitations of these two methods will also be provided. The information gained during this study will be used to assess the feasibility, and optimize the implementation, of a larger study upon the genotyping community at large.

For more information on the entire study, please attend the Joint Research Group presentation, and/or visit the accompanying poster.

METHODS

Human genomic DNA from nine individuals was used to produce eleven different samples, as shown in Figure 1. Samples 1-7 were taken directly from the respective individual. Sample 7 was further subjected to 6 rounds of GenomiPhi (Pharmacia) to produce sample 8. Samples 9-12 represent a mix of DNA from individuals II & I, ratios as indicated.

Prior to analysis by DNA sequencing and SNaPshot, samples were PCR amplified to produce fragments ranging from 350-740 bp. PCR primers were tagged with M13 universal forward and reverse primer sites that were used for the subsequent sequencing reactions. A separate set of primers was used for the SNaPshot reactions. Reactions were performed in both forward and reverse directions. SNaPshot reactions were carried out either individually or multiplexed, as per each participant’s preference. All samples, primers, and a standardized reaction set-up protocol was provided for each method, however, samples were ran according to each participant’s usual instrument parameters.

The samples were genotyped in a known polymorphic position in 4 different genes: AR, MTHFR, TNF and TYMS. The known genotype for each sample is shown in Table 1.

Concordance was calculated as the percentage of samples whose called genotype was identical to the known genotype.

RESULTS

A comparison of different analysis methods, obtained for (A) direct sequencing and (B) SNaPshot. The percentage of results that are in concordance with the expected genotype is reported across all samples from a single participant for each analysis method.

DISCUSSION

For the majority of the samples, participants were able to accurately assign the correct genotype using either DNA sequencing or SNaPshot with ≥80% accuracy. There are, however, some noteworthy exceptions. The AR locus was particularly problematic using DNA sequencing, due to the repetitive region occurring just upstream of the region of interest in the forward direction. This made it very difficult to assign a correct base-call in the forward direction, and some participants were unwilling or unable to assign a genotype to these samples based only on the reverse direction sequence. As only a single-base extension is performed in the SNaPshot reaction, this region had no bearing on the results.

As expected, sample mixtures that deviated from the normal 50/50 mix of bases found in a heterozygous individual were more problematic across all loci (samples 9-12) (refer to Figures 2 & 3). Those samples in which the 2nd base was less than 10% of the first (sample 9) were particularly difficult for both techniques (Figure 4).

An example of typical results obtained from (A) direct sequencing on the MTHFR SNP, for mixed samples 9-12, and (B) SNaPshot analysis of all SNPs for mixed samples 9-12.

Figure 6. Cost comparison between DNA sequencing and SNaPshot. Note that the amounts are for individually-processed samples, and cost reduction is possible with SNaPshot for multiplexed samples.

ACKNOWLEDGEMENTS

Thank you to all of our collaborators in the FARG and NARG research groups, and to our co-workers who helped to support this endeavor.

We would also like to acknowledge Mike Klein, University of Utah, for preparation and verification of all PCR reactions used in this study.