B45 The Evaluation of a Novel Massively Parallel Sequencing (MPS) Panel of 74 Microhaplotypes for Ancestry Prediction of Four Major United States Population Groups

Aishwaryaa Subramanian, The George Washington University, Washington, DC 20052; Fabio Oldoni, PhD, The George Washington University, Washington, DC 20007; Sathya Prakash Harihar, The George Washington University, Washington, DC 20007; Leena Yoon, Tysons, VA 22102; Sharon C. Wootton, PhD, South San Francisco, CA 94080; Robert Lagacé, BS, Thermo Fisher Scientific, South San Francisco, CA 94080; Ryo Hasegawa, BS, Foster City, CA 94404; Joseph P. Chang, BS, Thermo Fisher Scientific, South San Francisco, CA 94080; Moses S. Schanfield, PhD, The George Washington University, Washington, DC 20007; Kenneth Kidd, PhD, Yale University School of Medicine, New Haven, CT 06520; Daniele S. Podini, PhD, The George Washington University, Washington, DC 20007*

Learning Overview: After attending this presentation, attendees will better understand the advantages of using microhaplotype (MH) markers in addition to Single Nucleotide Polymorphism (SNP) and Insertion Deletion (InDel) panels for predicting the biogeographic ancestry of individuals.

Impact on the Forensic Science Community: This presentation will contribute to the forensic science community by providing an innovative MPS panel of MH loci, for improved ancestry prediction.

Recent advancements in high-throughput sequencing technologies have enabled exploring a new type of genetic marker: microhaplotypes (MHs).[1] These new forensic DNA markers are based on two or more SNPs within less than 300 bp from each other and can be genotyped using MPS platforms. The conventional Sanger sequencing method does not allow determining the *cis/trans* relationship of the SNP alleles (i.e., phase) within the same amplicon while MPS enables distinguishing the parental haplotypes by clonal sequencing of each individual strand of DNA, thus providing unambiguous SNP phase information at each locus. The small amplicon size, absence of stutter peaks along with lower mutation rates than conventional short tandem repeat (STR) loci are features that make MHs a promising marker for addressing relevant forensic challenges including ancestry prediction and mixture deconvolution.[2] In this study, the authors generated allele frequency databases for four American population groups and explored the potential for MPS-based MH analysis to provide biogeographic ancestry information.

A novel forensic panel of 74 MH loci was developed and implemented on the Ion Chef™ and Ion S5™ MPS (Thermo Fisher Scientific) platform.[3] A total of 100 European American (EA), 100 African American (AA), 100 South West Hispanic (SWHIS) and 100 East Asian American (EAA) population samples were selected and genotyped using the 74-plex MPS forensic assay. Allele frequencies were further generated to create allele frequency databases specific for each population group tested. In addition, a set of 10 unknown testing-samples representative of each population group was genotyped and related biogeographic ancestry inferred by calculating the random match probability (RMP) in the four corresponding American populations. The RMP calculated for the full set of samples was found remarkably higher for all those populations where individuals self-identified as such. Moreover, likelihood ratio (LR) was also calculated by dividing the highest RMP value obtained for the four tested populations by the second highest RMP value. The resulting LR value provides an indication of how much more likely it is to observe the MH profile of interest if it originated from an individual from the population at the numerator than if it originated from an individual from the population at the denominator. The biogeographic ancestry of the full set of testing-samples was correctly predicted using the allele frequencies generated from the four available in-house genotyped population groups. The level of heterozygosity of each MH locus was also calculated along with the power of exclusion (PE) to determine how good the MH markers are at discriminating individuals and excluding a random person as a possible contributor of an allele at a given locus, respectively. The heterozygosity ranged from 0.40 to 0.85 and the PE from 0.10 to 0.50. Finally, STRUCTURE-based iterative Bayesian clustering software was used for the analysis of population structure and to further confirm the origin of unknown testing-population samples.

These preliminary results suggest that the novel MPS 74-plex MH assay is an effective forensic DNA analysis tool, which provides enhanced biogeographic ancestry inference capabilities while supplementing the accuracy of existing ancestry prediction tools.

Reference(s):

1. Pakstis AJ, Fang R, Furtado MR, Kidd JR, Kidd KK. Mini-haplotypes as lineage informative SNPs and ancestry inference SNPs. *European Journal of Human Genetics* (2012) 20(11): 1148-1154.
2. Kidd KK, Speed WC, Pakstis AJ, Podini DS, Lagacé R, Chang J, Wootton S, Haigh E, Soundararajan U. Evaluating 130 microhaplotypes across a global set of 83 populations. *Forensic Science International: Genetics* (2017) 29:29-37.
3. Oldoni F, Hart R, Long K, Maddela K, Cisana S, Schanfield M, Wootton S, Chang J, Lagace R, Hasegawa R, Kidd K, Podini D, Microhaplotypes for ancestry prediction, *Forensic Science International: Genetics Supplement Series* (2017) 6:e513-e515.

Microhaplotypes, MPS, Ancestry Inference