A94 Molecular "Eyewitness": Predicting Phenotype and Geographic Ancestry via SNPs Analysis for Forensic Applications

Katherine L. Butler, MS, 716 Otis Place Northwest, Washington, DC 20010; Michelle A. Peck, BS, 4004 47th Street Northwest, Apartment 1, Washington, DC 20016; Jessica A. Hart, BS, 1204 South Washington Street, Apartment 403W, Alexandria, VA 22314; Moses S. Schanfiled, PhD, The George Washington University, 2100 Foxhall Road Northwest, Washington, DC 20007; and Daniele S. Podini, PhD, The George Washington University, Department of Forensic Science, 2100 Foxhall Road Northwest, Washington, DC 20007*

After attending this presentation, attendees will be able to illustrate the potential for the prediction of physical traits and geographic ancestry of an individual using SNPs analysis.

This presentation will impact the forensic science community by understanding DNA analysis targeted at inferring the possible ancestral origin and phenotypic characteristics (i.e., hair color, skin color, and eye color) of the possible perpetrator could yield information valuable to the investigators.

Often an STR DNA profile obtained from crime scene evidence does not match identified suspects or profiles from available databases. In such cases, further DNA analyses targeted at inferring the ancestral origin and phenotypic characteristics (i.e., hair color, skin color, and eye color) of the possible perpetrator could yield valuable investigative information. Such a tool would aid in prioritizing suspect processing, corroborating witness testimony, determining the relevance of a piece of evidence to a crime, and ultimately increase the ability to identify individuals related to the crime scene. The completion of the Human Genome Project and the International HapMap Project have provided the scientific community with a repository of reference information for the human nuclear genome, and efforts such as the 1000 Genome Project continue adding to this wealth of data. Numerous SNPs have been identified as having alleles associated with certain populations and/or correlated to specific physical characteristics.

The method chosen to develop a SNP based assay for ancestry and phenotype prediction is the Single Base primer Extension (SBE). This technique allows for the simultaneous typing of over 30 SNPs. Once an assay is optimized, it is possible to obtain robust results over a broad range of both quality and quantity of genomic DNA template. The sensitivity is in the range of STRs (down to 300 pg) and the method utilizes the same Capillary Electrophoresis equipment typically available in Forensic DNA laboratories.

The SBE technology has been used to develop panels which include 100 ancestry and phenotype markers selected from recent literature. Over 270 DNA samples along with corresponding ancestry/ phenotype survey information, and spectrophotometric skin color data have been collected from anonymous volunteers of varying ethnicity, gender, and age. These DNA samples, along with additional samples of known ancestry (without phenotype data), have been screened with the SBE panels. The genotypes and corresponding known characteristics are being evaluated to assess the predictive value of the candidate SNPs with the goal of identifying the optimal panel of SNPs to efficiently assess an unknown individual's characteristics. Different statistical approaches are being evaluated for effective ancestry and physical trait inference. *STRUCTURE 2.3* is a population genetics and anthropology software package, based on Bayesian statistics, which was developed to analyze the genetic composition of individuals and populations. It can be used for various purposes including, but not limited to, assigning individuals to populations. While this software is powerful, other tools are being evaluated which are potentially more appropriate for estimating the phenotype and ancestry of a particular individual. These include analysis of molecular variance (AMOVA), multinomial logistic regression, and principle component analysis (PCA). The ideal statistical tool would allow for a more complex model that could appropriately incorporate the ancestry information contained within different types of markers such as mtDNA and the Y chromosome.

The final goal is the selection of the most informative ~ 30 SNPs that will be incorporated into a robust and sensitive SBE assay for ancestry and somatic trait prediction. This analytical tool, utilizing technology currently available in forensic DNA laboratories, could be implemented in a kit form and used on casework as needed. Preliminary results show good correlation between a small set of SNPs and eye color, making blue or brown (light versus dark) eyes highly predictable. Furthermore, PCA analysis shows clear separation between black and red hair, while dark brown, light brown and blond hair are more difficult to separate. Similar results have been obtained with skin pigmentation and ancestry, indicating that developing models for the prediction of skin and hair pigmentation together with ancestry will be a challenging process.

SNPs, Phenotype Prediction, Ancestry Prediction

 ** Presenting Author*