B28 Evaluation of the Use of AmpF*l*STR® Identifiler™ STR Loci for the Inference of Ethnic Origins of Forensic Unknown DNA Profiles in Singapore

Hang Yee Wong, BSc, Simon Eng Seng Lim, BSc, Wai Fun Tan-Siew, MSc, and Shui Tse Chow, PhD, Health Sciences Authority, Centre for Forensic Science, DNA Database Laboratory, 7 Maxwell Road #05-04, MND Building Annex B, Singapore, Singapore 069111, Singapore*

After attending this presentation, attendees will have learned about the value of using the variations among the allele frequencies of each ethnic group to infer the ethnic origin of an unknown Identifiler™ developed DNA profile recovered from a crime scene.

This presentation will impact the forensic community and/or humanity by demonstrating the strengths and weaknesses of using the STR frequencies of Identifiler™ DNA markers to predict the ethnic origin of a given DNA profile, the accuracy of this method of inferring ethnic origin, and how this additional information can be useful in providing investigative leads for law enforcement agencies.

Singapore is a multi-ethnic country consisting of mainly Chinese, Malays, and Indians. Historically, these three major populations are migrants from the coastal regions of South China, the Malayan archipelago and South India more than two centuries ago. Based on the population database, it was observed that there are some unique and distinct differences in the STR allele frequency distribution between the three ethnic groups. By studying these dissimilarities, the laboratory has written a simple Excel® spreadsheet to predict the ethnic group of unknown forensic DNA profiles.

This presentation will impact the forensic community and/or provide humanity by demonstrating the possibility of inferring the ethnic origin of an unknown DNA profile using Identifiler™ STR allele frequencies.

The program was written based on two simple assumptions. Firstly, there is an unequal distribution of allele frequencies among the three populations in a few if not all loci. Hence, for an unknown DNA profile, different random match frequencies (RMF) will be obtained when using different ethnic population databases. Secondly, a large proportion within each ethnic population will have common alleles in a few if not all loci. Hence a DNA profile developed from a Chinese will yield the lowest RMF when using a Chinese population database compared to Malay or Indian database since the alleles encountered are more common in the Chinese population. With these understandings in mind, the laboratory has assumed that an unknown DNA profile always possess the highest composition of common alleles to its actual ethnic groups. Hence whichever population database gives the lowest RMF, the more probable the unknown DNA profiles belong to that ethnic group.

An Excel® spreadsheet was written based on the Identifiler™ STR loci allele frequencies of 500 individuals from the three ethnic groups. The function of the spreadsheet is to calculate the RMF for DNA profiles of unknown origin and to rank them accordingly; the database that gives the lowest RMF, the more likely that ethnic group it is. Testing of the excel spreadsheet was carried out with 200 DNA profiles from each ethnic groups. It has shown that the correct ethnic group is being inferred 58.8%, 70.9%, and 82.5% of the time for Malays, Chinese, and Indians respectively.

This in-house Excel® spreadsheet demonstrates the possibility of using Identifiler™ developed STR data to infer the most likely ethnicity of the DNA profile and to furnish such information to law enforcement agencies to serve as a potential investigation lead.

Ethnic Origin, Identifiler™, Random Match Frequency

 ** Presenting Author*