D53 A Computer Program for Calculating Forensic Population Study Parameters of STR Loci

Chang En Pu, MS, Ministry Justice Investigation Bureau, No.74 ChunHua Road, Hsin-Tien, Taipei, 231, Taiwan; and Adrian Linacre, PhD, Forensic Science Unit, Department of Pure and Applied Chemistry, University of Strathclyde, 16 Richmond Street, Glasgow, Scotland G1 1XQ, United Kingdom*

The goal of this presentation is to provide a program to help calculate a wide range of forensic population parameters and a FPP (false positive parentage) rate. This presentation will impact the forensic community and/or humanity by making the population study more efficient and provide an easy way to evaluate the false positive parentage rate to avoid false identification, especially for DNA database operation.

The computer program (STRstatistics 2005.1) is presented, which is capable of calculating a wide range of commonly used forensic population study parameters. These include: p value of G-tests for HWE proportion; the number of types of a particular allele; the occurrence frequency of alleles, expected and observed heterozygosity (H); polymorphism information content (PIC); power of discrimination (PD); probability of a match (PM); power of exclusion (PE) for trio and duo paternity tests; typical paternity index (PIt) and typical power of exclusion (PEt). The evaluation of data by these means is frequently a requirement in forensic practice, particularly when examining a new population.

At present, there are limitations to the computer programs that are available for forensic population studies, such as locus by locus handling (rather than batch handling), limited sample volume, and data format transformation. Many other genetic processing related computer programs were not designed exclusively for forensic evaluation of population study and therefore only provide analysis for a few forensic population parameters, therefore requiring additional calculation tools to be used. The STRstatistics 2005.1 program runs on the basis of the initial STR data such as that directly imported from Applied Biosystems Genotyper® software as well as an Excel format or by manual addition. Microsoft® Excel® Macros and built-in functions controlled by Visual Basic language written by the authors was used to handle the Hardy-Weingberg test and other forensic calculations. The application requires only that the users post or import their 15 STR genotypes from a population onto a Microsoft® Excel® worksheet, then press the hot key to activate the Macros. The allele frequency and forensic parameter table will be generated ready for publishing or for use as a population database. The program is capable of handling data of 1,000 individuals and 15 loci simultaneously, from which the informative forensic parameters will be tabulated automatically. The "STRstatistics 2005.1" Microsoft® Excel® template contains several worksheets. The "ori STR" worksheet provides brief instructions for using the template, and describes some limitations of the template. The genotype data for 15 STR loci, which comprises 30 columns with 2 columns for each locus, may be pasted onto the "ori STR" worksheet. Up to 40 alleles for each locus are acceptable. Genotypes containing text alleles (e.g., nc or 9.x) or with more than 2 alleles will be treated as in text (nonnumeric) mode and ignored in the auto-run program analyses. The final results table is produced as "publish tab" but can be modified manually by users to meet the required formats for publication. The program is freely available to any forensic scientist interested. Please e-mail requests to the corresponding author.

Forensic Science, STRs, Population Study

* *Presenting Author*