Interpreting Y-STR Evidence

Bruce Budowle
FBI Laboratory, Quantico, VA 22135

The Y-chromosome genetic markers, particularly STR loci, are the most recent class of DNA-based markers to gain wide acceptance as tools for human identity testing. The Y-chromosome is the smallest chromosome of the human genome (about 60 million base pairs), and unlike the autosomal chromosomes, it is transmitted solely along the paternal line. Most of the DNA in the Y-chromosome is non-recombinant. Therefore, barring mutation, Y-linked DNA profiles are identical for all paternal relatives, including male siblings.

For the application of Y-STR loci, interpretation issues need to be considered. Many are the same as those for autosomal STR loci. These include: a quality evaluation of the data, whether the sample is comprised of multiple donors or a single source sample, and the significance of "matching" data. Because of the haploid nature of Y- STR loci, for most loci, only one allele per locus is displayed per individual. Thus, interpretation of profiles in mixed samples is simplified. Those Y-STR loci selected for forensic applications are male-specific, also simplifying interpretation of mixed samples. Stochastic affects due to too few template molecules in the PCR is less of an issue for interpretation of Y-STR loci than for autosomal loci for single source samples. Thus, the minimum template/threshold interpretation requirements for the amplification by PCR of Y-STR loci can be lower than for autosomal STR loci.

When a Y- haplotype obtained from a forensic specimen matches that of a suspect or victim (or cannot be excluded as arising from a paternally-related individual), some significance is placed on the probability of such an occurrence. The mode of Y- STR inheritance must be considered to decide the statistical approach to use when placing weight on an observed Y-STR profile. The rarity of a multilocus Y-STR profile cannot be estimated as the combined product of the allele frequencies at each locus, as is done for the autosomal STR loci. They must be evaluated as a haplotype of loci. The counting method is one approach to convey an estimate of the rarity of the Y-haplotype. Basically, the number of times a particular haplotype is observed in a reference database(s) is counted and divided by the number of profiles in the data set. Then, a correction for sampling error is applied.

The reference population database(s) used in forensics are generally divided into major population groups. The degree of population heterogeneity is expected to be greater among populations within a major population group than for autosomal markers, and Y-STR haplotype data clearly demonstrate population substructure more so than for the autosomal loci. Therefore, when a reference database is not representative of the area of the crime scene, correction for effects of substructure should be considered. Analyses support use of the haplotype population data for estimating Y-STR profile frequencies for populations residing in North America.

The International Y-STR User Group has recommended use of nine Y-STR loci for identity testing. These are the loci: DYS19, DYS385a, DYS385b, DYS389I, DYS389II, DYS390, DYS391, DYS392, and DYS393 (known as the European Minimal Haplotype). The SWGDAM also recommended the use of these nine loci as well as the loci DYS438 and DYS439. Commercial kits are available to analyze these loci in a single multiplex reaction. Other loci are available for analysis. But, on a practical level, increasing the size of reference population databases will increase the power of analysis more so than adding more loci to an analysis. There are a number of reference database sites including: Charité – Universitary

Medicine Berlin YHRD, Promega's Powerplex Y haplotype database, Reliagene Technologies'Y-STR database, DNA heritage Ybase (a genealogical database).