

## POPULATION STATISTICS ON THE PROPOSED EXPANDED U.S. CORE LOCI

Carolyn R. (Becky) Hill, Margaret C. Kline, David L. Duewer, and John M. Butler  
National Institute of Standards and Technology, Applied Genetics Group, Gaithersburg, MD  
20899, USA

The original core set of 13 Combined DNA Index System (CODIS) autosomal short tandem repeat (STR) loci were selected in November 1997 and are required by the FBI for upload of DNA profiles to the national DNA database [1,2]. As the number of profiles stored in the National DNA Index System (NDIS) continues to increase each year (> 10 million total profiles), the likelihood of adventitious matches becomes greater [3]. Expanding beyond the 13 core loci is critical to reduce the potential of these types of matches occurring within the database, to increase international compatibility for data sharing, and to increase discrimination power in missing persons cases [3].

In November 2009, the European Union adopted five new autosomal STR loci as part of their expanded European Standard Set (ESS), including D12S391, D1S1656, D2S441, D10S1248, and D22S1045 [4,5]. All five of these loci are being considered for the expansion of the U.S. core set to provide greater capabilities for international comparisons when necessary. Also, D2S1338 and D19S443 are recommended as two new additions to the original 13 core loci because almost half of the U.S. national database already contains data for these loci. Finally, it has been suggested that the DYS391 locus be added to confirm amelogenin null alleles sometimes present in DNA profiles [3].

In the past few years, Promega Corporation and Life Technologies have released several new next generation STR multiplex kits that enable complete coverage of all of these additional loci. These multiplex kits have been extensively tested at National Institute of Standards and Technology (NIST), allowing the probability of identity (PI) calculations to be made with different sets of loci and population statistics, including allele frequencies and heterozygosity values for each locus, to be determined with our set of unrelated U.S. population samples. With this information, it has been possible to thoroughly characterize these new STR loci beyond the original 13 CODIS core loci to determine the impact that this additional information will have on database searches.

A summary of these results, including STR locus population statistics for the new STR loci, will be shown in order to help assess the benefits of adding additional loci to the current 13 CODIS core loci.

### References:

1. Budowle, B., et al. (1998). CODIS and PCR-based short tandem repeat loci: law enforcement tools. *Proceedings of the Second European Symposium on Human Identification*, pp. 73-88. Madison, Wisconsin: Promega Corporation. Available at <http://www.promega.com/geneticidproc/eusymp2proc/17.pdf>.
2. Butler, J.M. (2006). Genetics and genomics of core short tandem repeat loci used in human identity testing. *J. Forensic Sci.*, 51, 253-265.
3. Hares, D.R. (2012) Expanding the CODIS core loci in the United States. *Forensic Sci. Int. Genet.* 6(1):e52-4.
4. Gill, P., et al. (2006) The evolution of DNA databases-Recommendations for new European STR loci. *Forensic Sci. Int.* 156: 242-244.
5. Gill, P., et al. (2006) New multiplexes for Europe-amendments and clarification of strategic development. *Forensic Sci. Int.* 163: 155-157. ☞