

## Notice of release of the 2015 FBI Population Data for the expanded CODIS core STR loci

The FBI Laboratory recently announced an expansion of the original thirteen short tandem repeat (STR) loci that have been the core of the National DNA Index System (NDIS) since 1997 [See D.R. Hares (2015) Selection and Implementation of Expanded CODIS Core Loci in the United States. *Forensic Sci. Int. Genet.*, Volume 17, page 33-34 (available at <http://dx.doi.org/10.1016/j.fsigen.2015.03.006>)]. Seven additional STR loci were selected by the CODIS Core Loci Working Group and, following an implementation phase concluding on January 1, 2017, will also be required for upload and searching of DNA profiles at NDIS. Collectively, these loci provide greater discrimination potential for human identification applications and enhance kinship analyses typically used in missing person inquiries. Because many of these loci are included in databases worldwide, the expanded STR locus set facilitates international law enforcement and counterterrorism endeavors. The twenty STR loci (the original set: D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, CSF1PO, FGA, TH01, TPOX and vWA; and the additional set: D1S1656, D2S441, D2S1338, D10S1248, D12S391, D19S433 and D22S1045) can be simultaneously genotyped with either the AmpFISTR® GlobalFiler® (GlobalFiler, Life Technologies, Inc., Carlsbad, CA) or PowerPlex® Fusion™ (Fusion, Promega Corporation, Madison, WI) multiplex amplification systems. These kits also enable the genotyping of SE33 and a Y indel locus (GlobalFiler), Penta D and Penta E (Fusion), and DYS391 (GlobalFiler and Fusion), as well as Amelogenin for sex determination.

The 2015 Expanded FBI STR Population Data includes allele frequencies for the autosomal STR loci as determined with both the GlobalFiler and Fusion kits in African Americans, Caucasians, Southeastern Hispanics, Southwestern Hispanics, Bahamians, Jamaicans, Trinidadians, Apaches, Navajos, Chamorros and Filipinos. Concordance studies on these populations demonstrated genotyping accuracy and identified instances of non-concordance due to rare kit-specific primer binding site variants. Additionally, the results of population genetic analyses supported the usage of these loci and the associated allele frequencies for estimating match statistics in human identity testing. Once the manuscript completes the internal FBI Laboratory internal review process, this population data will be submitted for publication under the title “*Population data on the expanded CODIS core STR loci for eleven populations of significance for forensic DNA analyses in the United States.*”

Please note that the 2015 Expanded FBI STR Population Data and the Amended 1999/2001 FBI STR Population Data were generated using different STR typing kits on the FBI population samples. The Amended 1999/2001 Data corrects the 1999/2001 FBI STR Population Data originally published in *Population data on the thirteen CODIS core short tandem repeat loci in African Americans, U.S. Caucasians, Hispanics, Bahamians, Jamaicans and Trinidadians*, *Journal of Forensic Sciences* 1999 44(6):1277-86. The Amended data is described in the July 2015 Erratum (available at <http://onlinelibrary.wiley.com/doi/10.1111/1556-4029.12806/abstract>). The Amended 1999/2001 Data corrects the errors and the duplicated samples present within the original 1999/2001 data, and required no retesting of population samples to make the corrections described in the Erratum.

The 2015 Expanded FBI STR Population Data is the retyping of all of the remaining FBI population samples using the new expanded loci kits. The biggest difference between the 2015 FBI STR Population Data and the Amended 1999/2001 Populations is that the 2015 expanded data combines the African American, Bahamian, and Jamaican populations due to the small genetic distance found among these populations as demonstrated by the population pair-wise *F<sub>st</sub>* values. Additionally, because some of the

original population samples were not available for testing in 2015 and because others yielded partial or no profiles using the new expanded loci kits, the N for some populations in the 2015 Expanded FBI STR Population Data and for some loci within these populations are different from the N values for the 1999/2001 Amended Populations. For example, the Amended 1999/2001 Trinidadian population contains 85 individuals (with an across loci range of from 78 to 85 individuals) while the 2015 expanded data consists of 79 Trinidadian individuals (with an across loci range of 77 to 79, excluding male-specific markers). Other similar differences exist across the two datasets, but the 2015 Expanded FBI STR Populations greatly reduce the instances of within-population, across-loci variation in the N (for the autosomal loci of a given population). The 2015 Expanded FBI STR Populations are being provided as a new data set with more loci for use by laboratories as they migrate to the new expanded loci PCR typing kits in anticipation of the January 2017 deadline for the implementation of the 20 CODIS core loci. In summary, the 1999/2001 Amended and 2015 Expanded FBI STR Populations are independently generated datasets that contain accurate data for different numbers of STR loci and these different datasets contain differences in sample size at both the population and locus levels. BOTH datasets contain sample sizes appropriate for STR DNA databases and all meet population testing criteria, so either the amended or the expanded populations are suitable for use.

Additionally, the FBI Laboratory is providing herein the 2015 Expanded FBI STR allele frequency tables for use by anyone interested in implementing this dataset prior to its formal publication. Please note that while these pre-publication population data tables have undergone extensive review during their compilation by the FBI, they have not yet undergone the formal peer review process required for publication. While the FBI has complete confidence in the accuracy of this data and has implemented the 2015 FBI STR Populations for its own DNA casework examinations, the Laboratory believes it necessary to make potential users aware of this fact so that laboratories can take it consideration during any internal implementation discussions.

If you have any questions, please contact the FBI's DNA Support Unit at 703-632-7572.

African American Expanded 2015 STR Allele Frequency Table

Caucasian Expanded 2015 STR Allele Frequency Table

Southeastern Hispanic Expanded 2015 STR Allele Frequency Table

Southwestern Hispanic Expanded 2015 STR Allele Frequency Table

Apache Expanded 2015 STR Allele Frequency Table

Navajo Expanded 2015 STR Allele Frequency Table

Trinidadian Expanded 2015 STR Allele Frequency Table

Chamorro Expanded 2015 STR Allele Frequency Table

Filipino Expanded 2015 STR Allele Frequency Table



















