Variability of New STR Loci and Kits in US Population Groups

ABSTRACT

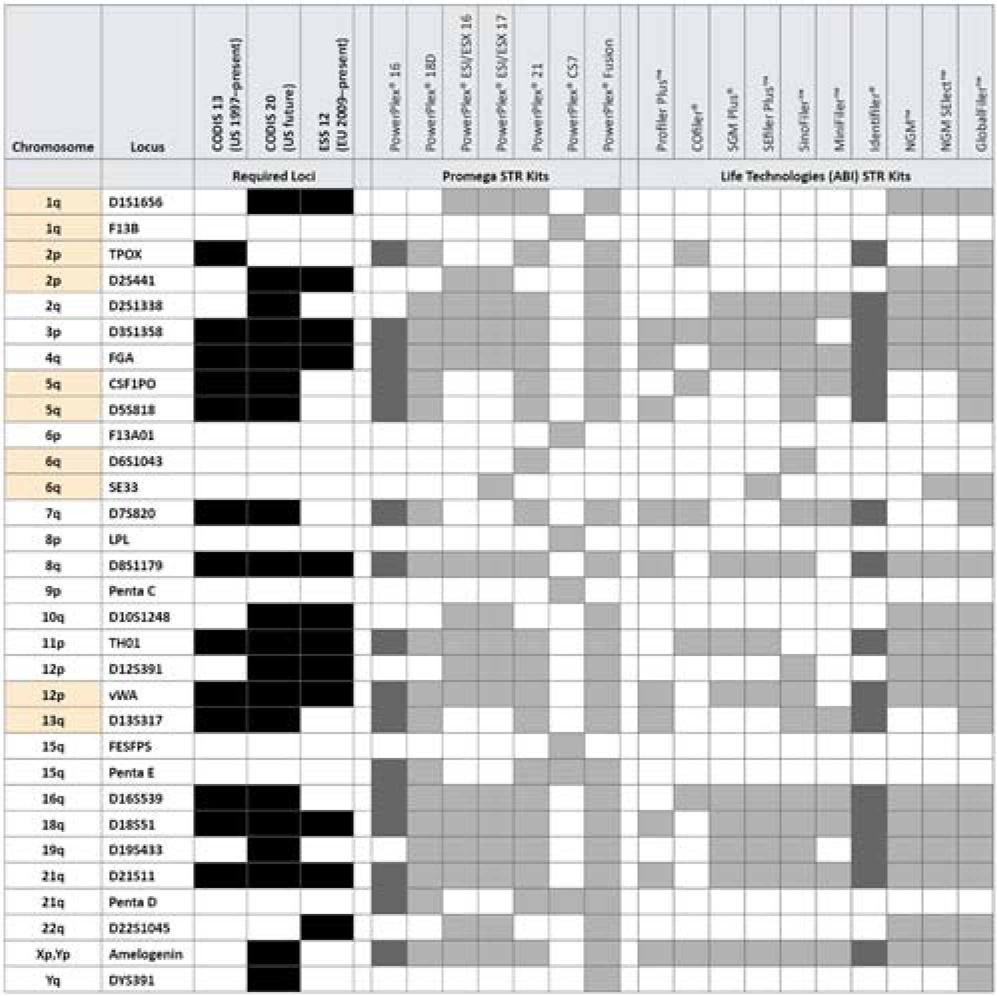
Over the past few years, Promega Corporation, Applied Biosystems (now Life Technologies Corporation) and Qiagen have released a number of new short tandem repeat (STR) typing kits to aid worldwide efforts in forensic DNA testing and human identification. The most recent STR kits were developed primarily due to expansion of the European Standard Set of STR loci and the U.S. core loci used for the Combined DNA Index System. These new commercially available STR kits utilize different combinations and subsets of 24 autosomal STR loci: CSF1PO, FGA, TH01, TPOX, vWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, D2S1338, D19S433, Penta D, Penta E, SE33, D1S1656, D12S391, D2S441, D10S1248, D22S1045 and D6S1043. In addition, Promega supplies five additional autosomal STR loci as part of the PowerPlex® CS7 System, which is primarily used in relationship testing laboratories. The CS7 loci include F13A01, F13B, FESFPS, LPL and Penta C. In July 2012, Promega released the PowerPlex® Y23 System, which includes six additional Y-chromosome STRs beyond the Applied Biosystems Yfiler® kit: DYS481, DYS533, DYS549, DYS570, DYS576 and DYS643. The Applied Genetics Group at the National Institute of Standards and Technology has examined a standard set of U.S. population samples across the 29 autosomal STR loci and the 23 Y-STRs available in commercial STR kits. This article describes the observed variability of STR loci in 1036 unrelated (primarily male) samples from U.S. Caucasian, Hispanic, Asian and African American self-declared ancestries.

John M. Butler, Carolyn R. Hill and Michael D. Coble
Applied Genetics Group, National Institute of Standards and Technology, Gaithersburg, Maryland, USA
Publication Date: 2012

Introduction

Over the past few years, Europe and the United States have worked to expand core sets of genetic markers used for routine forensic DNA testing and human identification. This expansion will aid international data-sharing capabilities with more loci in common and reduce the potential for adventitious matches as DNA databases grow in size. The European Standard Set (ESS) of short tandem repeat loci increased from seven to 12 core loci in 2009 [(1)] [(2)] [(3)] . The FBI Laboratory has announced plans to expand from a core set of 13 autosomal STRs used in its Combined DNA Index System (CODIS) to at least 18 loci in the near future [(4)] [(5)] [(6)] . Commercial STR typing kits have been released by Promega Corporation (Madison, WI) and Life Technologies/Applied Biosystems (Foster City, CA) to meet these core loci expansion needs.

Table 1 lists the U.S. and European core loci and the 29 autosomal STR loci found in commercial kits involved in the present study, which is focused on recently released kits from Promega and Applied Biosystems. For more information and specific details regarding the loci and kit configurations, see *Advanced Topics in Forensic DNA Typing: Methodology* [(7)] or our recent article in *Forensic Science Review* [(8)] . As noted in a recent study, genotypes from these autosomal STR loci are not associated with any known phenotypes and therefore do not reveal any potentially sensitive or medically relevant information [(9)] .



Chromosome	Locus	CODIS 13 (US 1997–present)	CODIS 20 (US future)	ESS 12 (EU 2009–present)	PowerPlex® 16	PowerPlex® 18D	PowerPlex® ESI/ESX 16	PowerPlex® ESI/ESX 17	PowerPlex® 21	PowerPlex® CS7	PowerPlex® Fusion	Profiler Plus™	COfiler®	SGM Plus®	SEfiler Plus™	Sinofiler™	MiniFiler™	Identifiler®	NGM™	NGM SElect™	GlobalFiler™
		Required Loci			Promega STR Kits							Life Technologies (ABI) STR Kits									
1q	D1S1656		■	■			▒	▒	▒		▒								▒	▒	▒
1q	F13B									▒											
2p	TPOX	■			▓	▒			▒		▒		▒					▓			▒
2p	D2S441		■	■			▒	▒			▒								▒	▒	▒
2q	D2S1338		■			▒	▒	▒	▒		▒			▒	▒	▒	▒	▓	▒	▒	▒
3p	D3S1358	■	■	■	▓	▒	▒	▒	▒		▒	▒	▒	▒	▒	▒		▓	▒	▒	▒
4q	FGA	■	■	■	▓	▒	▒	▒	▒		▒	▒		▒	▒	▒	▒	▓	▒	▒	▒
5q	CSF1PO	■	■		▓	▒			▒		▒		▒			▒	▒	▓			▒
5q	D5S818	■	■		▓	▒			▒		▒	▒				▒		▓			▒
6p	F13A01									▒											
6q	D6S1043								▒							▒					
6q	SE33							▒							▒					▒	▒
7q	D7S820	■	■		▓	▒			▒		▒	▒	▒			▒	▒	▓			▒
8p	LPL									▒											
8q	D8S1179	■	■	■	▓	▒	▒	▒	▒		▒	▒		▒	▒	▒		▓	▒	▒	▒
9p	Penta C									▒											
10q	D10S1248		■	■			▒	▒			▒								▒	▒	▒
11p	TH01	■	■	■	▓	▒	▒	▒	▒		▒		▒	▒	▒			▓	▒	▒	▒
12p	D12S391		■	■			▒	▒	▒		▒					▒			▒	▒	▒
12p	vWA	■	■	■	▓	▒	▒	▒	▒		▒	▒		▒	▒	▒		▓	▒	▒	▒
13q	D13S317	■	■		▓	▒			▒		▒	▒				▒	▒	▓			▒
15q	FESFPS									▒											
15q	Penta E				▓	▒			▒	▒	▒										
16q	D16S539	■	■		▓	▒	▒	▒	▒		▒		▒	▒	▒	▒	▒	▓	▒	▒	▒
18q	D18S51	■	■	■	▓	▒	▒	▒	▒		▒	▒		▒	▒	▒	▒	▓	▒	▒	▒
19q	D19S433		■			▒	▒	▒	▒		▒			▒	▒	▒		▓	▒	▒	▒
21q	D21S11	■	■	■	▓	▒	▒	▒	▒		▒	▒		▒	▒	▒	▒	▓	▒	▒	▒
21q	Penta D				▓	▒			▒	▒	▒										
22q	D22S1045			■			▒	▒			▒								▒	▒	▒
Xp,Yp	Amelogenin		■		▓	▒	▒	▒	▒		▒	▒	▒	▒	▒	▒	▒	▓	▒	▒	▒
Yq	DYS391		■								▒										▒

Table 1. STR loci and kits examined in this study along with United States and European Union required loci. Loci are listed by chromosomal position, with syntenic pairs highlighted in yellow. The GlobalFiler™ kit also includes a Yindel to confirm AMEL Y.

Click to Enlarge

The recently released PowerPlex® Fusion (Promega Corporation) and GlobalFiler™ (Life Technologies) STR typing kits include loci that are unfamiliar to many U.S. crime laboratories such as D12S391, D1S1656, D10S1248, D2S441, D22S1045 and SE33. Although this article does not cover any direct performance work with the PowerPlex® Fusion or GlobalFiler™ kits, we discuss the combination of loci present in these kits and their relative variability in U.S. population samples. We also examine five additional STR loci that are only found in the PowerPlex® CS7 System, which is primarily used by relationship testing laboratories. The PowerPlex® CS7 additional loci are F13A01, F13B, FESFPS, LPL and Penta C (see www.promega.com/products/genetic-identity/str-analysis-for-forensic-and-paternity-testing/powerplex-cs7-system_-custom/). In July 2012, Promega released the PowerPlex® Y23 System (see www.promega.com/products/pm/genetic-identity/powerplex-y23), which includes six additional Y-chromosome STRs (Y-STRs) beyond the Applied Biosystems Yfiler® kit: DYS481, DYS533, DYS549, DYS570, DYS576 and DYS643. In this article, we also explore the benefits of these additional Y-STRs.

Over the past decade, the Applied Genetics Group at the National Institute of Standards and Technology (NIST) has studied a standard set of U.S. population samples across numerous genetic markers of interest to the human identity testing community (see www.cstl.nist.gov/strbase/NISTpop.htm). This article describes the relative variability of individual STR loci—and thus the potential resolving power for combinations of loci found in commercial STR kits—through examining 1036 unrelated samples from U.S. Caucasian, Hispanic, Asian and African American sources across 29 autosomal STR loci and 23 Y-STRs.

Materials and Methods

Sources of DNA Samples: The DNA samples and their sources have been described in previous publications (10) (11) (12) . Anonymous donors with self-identified ethnicities were obtained from Interstate Blood Bank (Memphis, TN), Millennium Biotech, Inc. (Ft. Lauderdale, FL) and DNA Diagnostics Center (Fairfield, OH). Our complete data set of almost 1450 samples (12) contains related individuals including father and son samples for studying mutation rates (11) .

In the present collection of unrelated individuals, only the sons were used from the father-son sample set. In addition, all matching, single-mismatch or double-mismatch Y-STR haplotypes produced from PowerPlex® Y23 were subjected to kinship calculations across the autosomal STR loci to help identify any potential close (first or second degree) relatives in the data set. After this screening process, a final set of 1036 unrelated individuals was selected: 361 U.S. Caucasians, 342 African Americans, 236 U.S. Hispanics and 97 U.S. Asians. There are 1032 males and four females (two Caucasian, one African American and one Asian). The female sample names are C82B, C38A, MT94831 and UA16927.

DNA Results: For the past decade, the NIST U.S. population and father-son samples have been repeatedly examined with over 20 different commercial STR typing kits (and in some cases, prerelease prototype materials) for concordance study purposes (13) . Some of these studies have been published (12) (14) (15) (16) (17) (18) (19) , and others will be made available in the near future.

Details regarding any primer-binding-site mutations discovered during our concordance testing have been shared directly with the collaborating company, which has enabled improved products for the international DNA community. In some cases, primers may have been redesigned (and samples retested to verify concordance in allele calls) prior to kit release. In other instances, extra degenerate primers have been included (and samples retested) in final commercial kit multiplexes to enable PCR amplification of the regular and null or variant allele possessing a primer-binding-site mutation (7) (20) .

For work with commercial STR kits, manufacturer amplification and capillary electrophoresis protocols were followed with the exception of running half-volume PCRs for selected STR kits. Data were collected on Applied Biosystems® 3130*xl* and 3500 Genetic Analyzers as described in previous publications (12) .

Profiles also have been generated on these same DNA samples (or subsets) using in-house-developed multiplexes such as various miniSTR assays (14) (21) (22) (23) and the NIST 26plex (24) . Y-chromosome haplotypes have been produced with PowerPlex® Y (25) , Yfiler® (11) , in-house Y-STR assays (26) (27) , rapidly mutating Y-STRs (28) and PowerPlex® Y23. Portions of our Y-STR data are included in the YHRD (www.yhrd.org) and US Y-STR (www.usystrdatabase.org) databases.

Over 100 autosomal single nucleotide polymorphism (SNP) markers have been assessed on a subset of these samples (29) (30) (31) as have 50 Y-SNPs (32) , 68 insertion/deletion (Indel) markers (33) , 16 X-chromosome miniSTR loci (34) and the Argus X-12 kit (35) . In addition, the Armed Forces DNA Identification Laboratory has conducted mitochondrial DNA control region sequencing (36) (37) and complete mitochondrial genome analysis (38) on a fraction of these samples. Many of these results are available at: www.cstl.nist.gov/strbase/NISTpop.htm

Data Analysis: An Excel®-based software tool was developed by David Duewer (NIST Analytical Chemistry Division) to perform calculations of allele and genotype frequencies as well as the probability of identity values reported in this article. This software is planned for future release on the NIST STRBase web site:

www.cstl.nist.gov/strbase/software.htm. Hardy-Weinberg equilibrium assessment of the autosomal STR loci was performed using PowerMarker v3.25 [(39)] .

Results and Discussion

Benefits of Our Repeated Study of the Same DNA Samples: Three aspects that can affect overall quality and reliability of DNA results include 1) the presence of primer-binding-site mutations (that can lead to absent alleles failing to be called), 2) incorrect categorization of a sample into its appropriate population group (that can lead to a slight skewing of frequencies if true differences exist between population groups) and 3) the presence of relatives in a data set when unrelated individuals are being assumed (again this scenario can skew frequencies due to an overabundance of allele sharing among related individuals). Each of these issues—albeit in typically a very minor amount—can potentially shift the observed allele frequency from the "true" allele frequency that exists in a population.

Although usually rare (≈0.1% [(17)]), primer-binding-site mutations can give rise to typing results that do not reflect the true underlying alleles present in a DNA sample. Most university and forensic labs conduct population genetic research or casework with only a single STR typing kit. Therefore, investigators do not typically have the opportunity to cross-check results with different PCR primer pairs. It is worth noting that, by using consistent primers within a laboratory, question and known samples will agree if both contain the same primer-binding-site mutation. In other words, false homozygotes caused by allele dropout in a heterozygote when a primer fails to anneal properly and extend are not a problem *within* a laboratory. It is with interlaboratory comparisons and DNA databases supplied with results from multiple laboratories using different kits where potential null alleles can cause artificial mismatches. Of course, some DNA database search strategies incorporate reduced stringency to avoid false negatives that can arise with null alleles from a specific STR kit [(7)] .

As noted above, our NIST population samples have been well characterized for human identity testing markers and repeatedly examined with autosomal, Y-chromosome, X-chromosome and mtDNA markers. Some of the commonly used STR loci, such as D18S51, have been examined with more than a dozen different PCR primer pairs [(17)] . Detailed characterization of our DNA samples with redundant typing results, which were obtained due to probing the same region of DNA multiple times on the same samples with different PCR primer sets, increases confidence that our final data set is free of potential null alleles from specific primer-binding-site mutations. Thus, concordance studies aid typing result quality assurance.

Known Null Alleles Have Been Removed from Our Data Set: In our final data set, we report the correct underlying data in spite of the fact that a specific STR kit may not amplify all of the alleles reported. For example, African American sample GT37178 has a mutation in the amelogenin X allele primer-binding site for PowerPlex® 16 and Identifiler® STR kits, such that only the amelogenin Y allele is amplified [(40)] . However, in our final genotypes table, GT37178 is listed as containing an X,Y type for amelogenin because it amplifies correctly with other PCR primer pairs. In this same manner, with Hispanic sample ZT80731, D18S51 is reported as a 13,15 genotype in spite of the fact that all Applied Biosystems kits (Identifiler®, NGM SElect™, etc.) do not detect the allele 13 [(17)] due to a primer-binding-site mutation G→A 172 nucleotides downstream of the STR region that affects the reverse primer [(41)] .

Self-Identified Sample Ancestries Have Been Confirmed with Ancestry-Informative Markers: Many of our DNA samples have been studied with ancestry-informative markers [(30)] [(31)] . In studies where autosomal, Y and mtDNA loci were compared [(30)] , we obtained appropriate haplogroups for the lineage markers, which aided our confidence in the sample clustering and self-declared ancestries of our DNA donors.

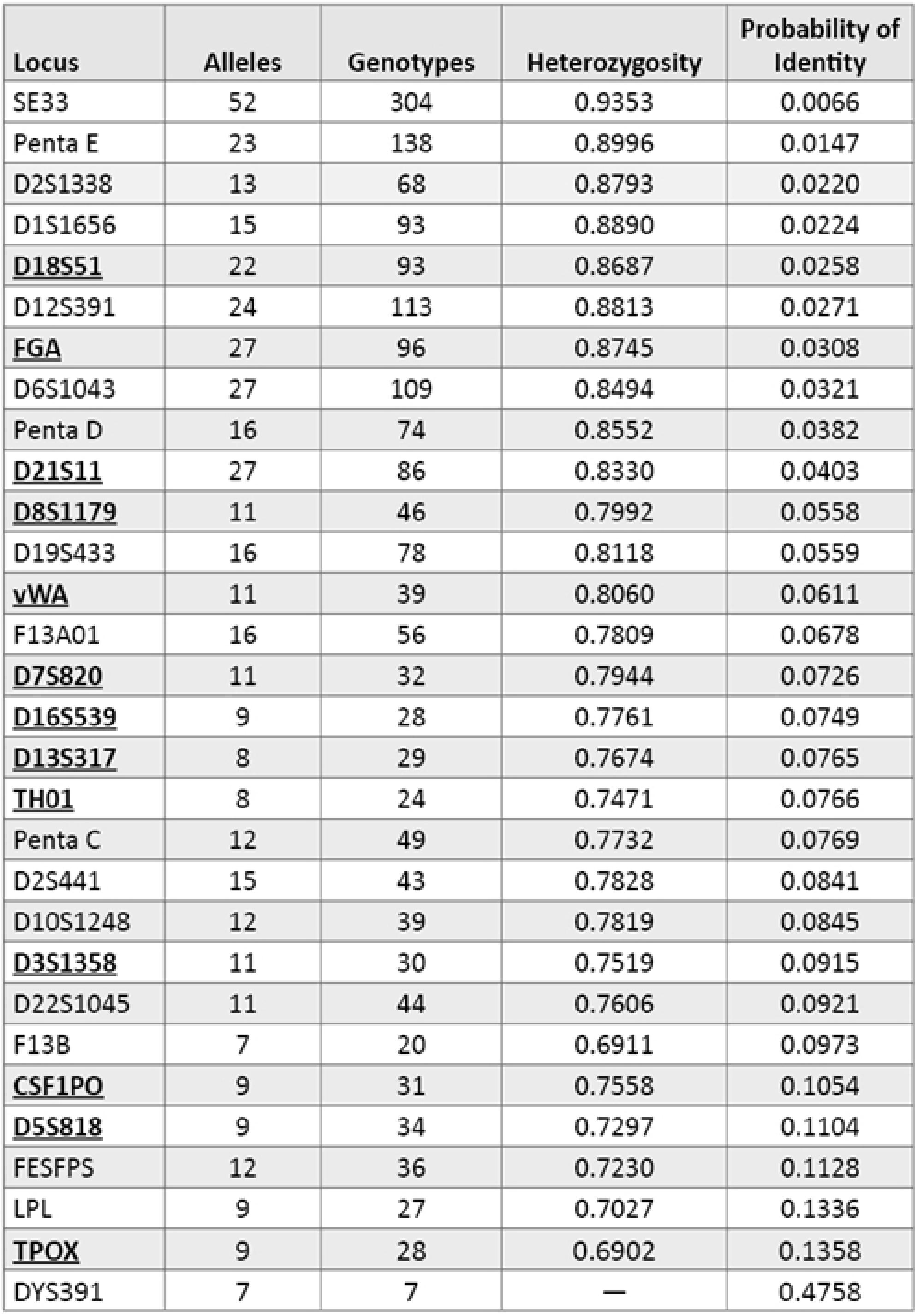
Related Samples Have Been Removed from Our Data Set: Only sons were included from our full father-son set of information. In addition, we have used the Y-STR (and in some cases mtDNA) lineage marker results to help screen out potential relatives for our final unrelated data set. An advantage of examining our DNA sample sets with many different loci and STR typing kits is the ability to identify potential relatives [42] . An example helps illustrate our screening process. Hispanic samples ZT79994 and ZT79995 share a total of 22 alleles at 22 loci out of 24 autosomal loci studied (only D12S391 and Penta D have nonoverlapping heterozygous alleles). This sample pair also exhibits a full match across the 23 Y-STR loci in PowerPlex® Y23 and possesses the same mtDNA control region sequence. A full Y-STR match suggests that either a father-son or a brother-brother relationship is possible. However, identical mtDNA control region sequences make it more likely that these sample sources are brothers. Kinship calculations produced likelihood ratios of zero for a parent-child relationship, 56300 for full siblings, 5690 for half-siblings (or uncle-nephew, grandfather-grandson) and 264 for first cousins. Thus, the genetic results are much more likely if ZT79994 and ZT79995 are brothers than unrelated individuals or even other close relatives. With this additional information, we have pulled ZT79995 from our data set, leaving ZT79994 to represent this individual’s family in our final unrelated data set.

Autosomal STR Markers

Relative Performance of the 29 Autosomal STR Loci: For human identification applications, loci that exhibit a high degree of variability among individuals are typically preferred so that samples from different sources can be appropriately separated from one another. There are several measures of genetic markers to help assess their relative variability. These include the number of observed alleles, the number of observed genotypes, the heterozygosity (which is a fractional evaluation of the total number of heterozygotes in a set of samples) and the probability of identity.

The probability of identity (P_I) is the chance that two unrelated people selected at random will have the same genotype [43] . The P_I value of a single locus is determined by summing the square of the observed genotype frequencies. Thus, the lower the P_I value, the more variable the genetic marker is in the measured population because there are more genotypes occurring at a lower frequency. Individual locus P_I values can be multiplied together with independently inherited loci to create a profile P_I (i.e., the product rule). The P_I value is a better measure of locus performance than the total number of observed alleles or genotypes because specific alleles may occur in a relatively high frequency and reduce the overall variability, especially if a number of rare alleles occur at this locus. Ideally, we would like to observe a fairly even level of variation across many genotypes at each locus so that there is a greater chance of finding a difference between two unrelated individuals selected at random.

Table 2 ranks the variation observed in our 1036 U.S. population samples across all 29 autosomal STR loci and DYS391 that are part of commercial STR kits as noted in Table 1. The information in Table 2 is listed with the most variable loci at the top (by probability of identity values). Note that there are several loci that are better than D18S51, the top performer for the current CODIS 13 loci. The most variable locus in our data set is SE33 with 52 observed alleles, 304 different genotypes, 93.5% heterozygosity and a P_I value of 0.0066. The least variable autosomal STR is TPOX with 9 alleles, 28 genotypes, 69% heterozygosity and a P_I value of 0.1358.



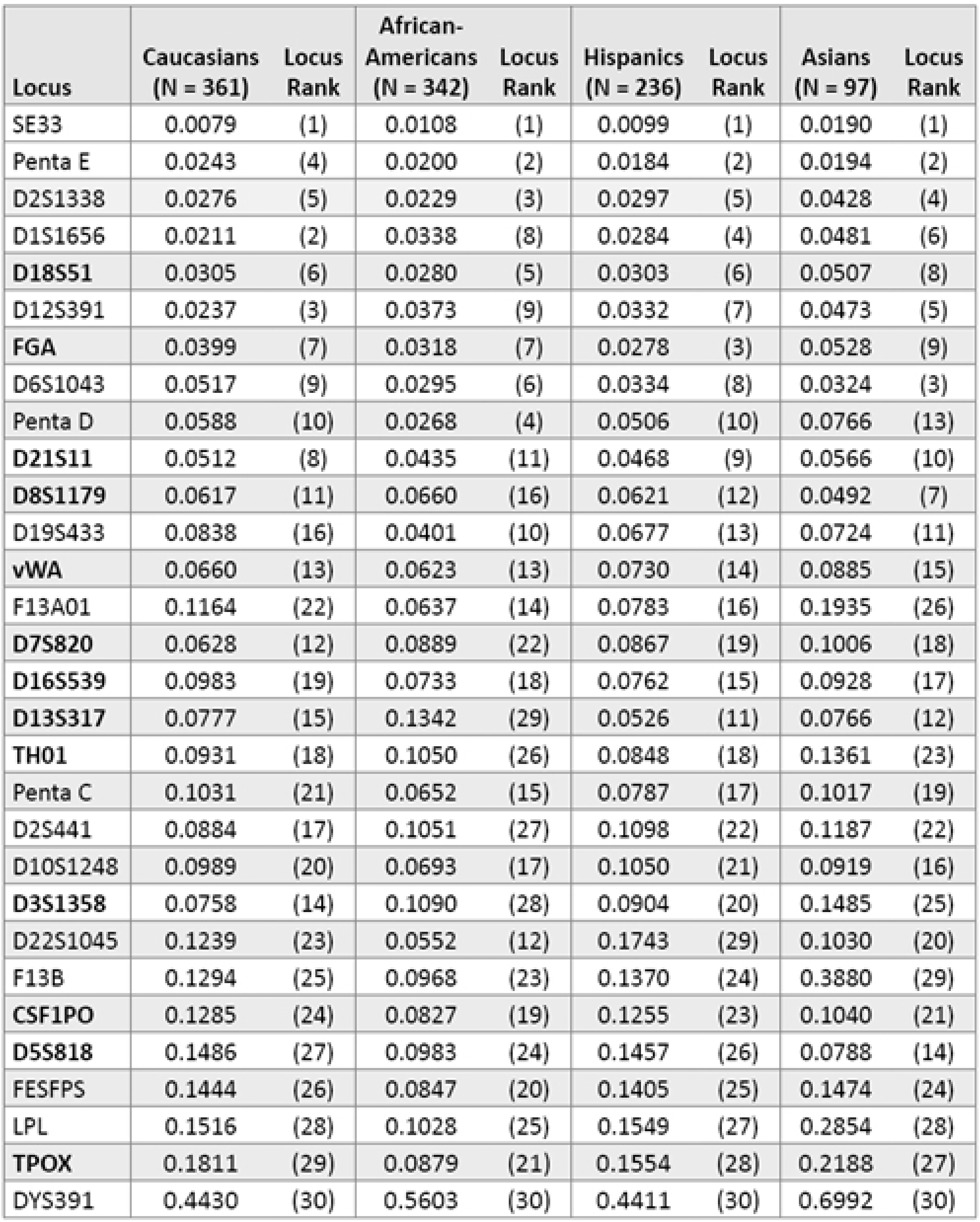
Locus	Alleles	Genotypes	Heterozygosity	Probability of Identity
SE33	52	304	0.9353	0.0066
Penta E	23	138	0.8996	0.0147
D2S1338	13	68	0.8793	0.0220
D1S1656	15	93	0.8890	0.0224
D18S51	22	93	0.8687	0.0258
D12S391	24	113	0.8813	0.0271
FGA	27	96	0.8745	0.0308
D6S1043	27	109	0.8494	0.0321
Penta D	16	74	0.8552	0.0382
D21S11	27	86	0.8330	0.0403
D8S1179	11	46	0.7992	0.0558
D19S433	16	78	0.8118	0.0559
vWA	11	39	0.8060	0.0611
F13A01	16	56	0.7809	0.0678
D7S820	11	32	0.7944	0.0726
D16S539	9	28	0.7761	0.0749
D13S317	8	29	0.7674	0.0765
TH01	8	24	0.7471	0.0766
Penta C	12	49	0.7732	0.0769
D2S441	15	43	0.7828	0.0841
D10S1248	12	39	0.7819	0.0845
D3S1358	11	30	0.7519	0.0915
D22S1045	11	44	0.7606	0.0921
F13B	7	20	0.6911	0.0973
CSF1PO	9	31	0.7558	0.1054
D5S818	9	34	0.7297	0.1104
FESFPS	12	36	0.7230	0.1128
LPL	9	27	0.7027	0.1336
TPOX	9	28	0.6902	0.1358
DYS391	7	7	—	0.4758

Table 2. Variation observed across 1036 U.S. population samples with 29 autosomal STR loci and the Y-STR locus DYS391. Loci are rank ordered by their probability of identity values. The current CODIS 13 core loci are shown in bold font and are underlined. DYS391, as a single-copy Y-STR marker, has no heterozygosity and exhibits the poorest ability to resolve samples from one another.

Click to Enlarge

Comparing and contrasting these marker metrics can be instructive. For example, D2S1338 displays only 13 alleles and 68 genotypes yet has the third best probability of identity score and the fifth best heterozygosity value (87.9%). This means that D2S1338 has a more even distribution of frequencies across the observed alleles compared to a marker like D3S1358 that has 11 alleles but only 30 genotypes with about 75% heterozygosity and a P_I of 0.0915. Almost 80% of our D3S1358 results contain alleles 15, 16, or 17, while the top three D2S1338 alleles 17, 19, and 20 account for only 42% of the observed variation at the more polymorphic D2S1338 locus. On a per-locus basis, a marker such as D2S1338, which has a broader distribution of variation across its alleles, will be more effective at detecting DNA mixtures from multiple contributors than D3S1358 even though the total number of possible alleles is not that different. Thus, with locus variability it is the distribution of genotypes that matters rather than the total number of alleles observed.

Table 3 examines relative rankings of the 29 autosomal STRs and DYS391 with our four different population groups. In all four groups, SE33 is the top-performing locus. Penta E has the second highest P_I value in three of the four populations—only exceeded by D1S1656 and D12S391 in our Caucasian data. Most of the other top ten performing loci (down to D21S11) remain in the top ten of each population group. It is interesting to note some population-specific differences such as D6S1043, which ranks 8th overall but 3rd with Asians. In reality, D6S1043 is not performing that much better in Asians in terms of absolute P_I value across our population groups but rather the other loci such as D2S1338, D1S1656, D18S51 and D12S391 are not quite as variable in our Asian samples compared to the other population groups. Our Penta C data compares well to previous data reported on African American, Asian, Caucasian and Hispanic samples (44) .



Locus	Caucasians (N = 361)	Locus Rank	African-Americans (N = 342)	Locus Rank	Hispanics (N = 236)	Locus Rank	Asians (N = 97)	Locus Rank
SE33	0.0079	(1)	0.0108	(1)	0.0099	(1)	0.0190	(1)
Penta E	0.0243	(4)	0.0200	(2)	0.0184	(2)	0.0194	(2)
D2S1338	0.0276	(5)	0.0229	(3)	0.0297	(5)	0.0428	(4)
D1S1656	0.0211	(2)	0.0338	(8)	0.0284	(4)	0.0481	(6)
D18S51	0.0305	(6)	0.0280	(5)	0.0303	(6)	0.0507	(8)
D12S391	0.0237	(3)	0.0373	(9)	0.0332	(7)	0.0473	(5)
FGA	0.0399	(7)	0.0318	(7)	0.0278	(3)	0.0528	(9)
D6S1043	0.0517	(9)	0.0295	(6)	0.0334	(8)	0.0324	(3)
Penta D	0.0588	(10)	0.0268	(4)	0.0506	(10)	0.0766	(13)
D21S11	0.0512	(8)	0.0435	(11)	0.0468	(9)	0.0566	(10)
D8S1179	0.0617	(11)	0.0660	(16)	0.0621	(12)	0.0492	(7)
D19S433	0.0838	(16)	0.0401	(10)	0.0677	(13)	0.0724	(11)
vWA	0.0660	(13)	0.0623	(13)	0.0730	(14)	0.0885	(15)
F13A01	0.1164	(22)	0.0637	(14)	0.0783	(16)	0.1935	(26)
D7S820	0.0628	(12)	0.0889	(22)	0.0867	(19)	0.1006	(18)
D16S539	0.0983	(19)	0.0733	(18)	0.0762	(15)	0.0928	(17)
D13S317	0.0777	(15)	0.1342	(29)	0.0526	(11)	0.0766	(12)
TH01	0.0931	(18)	0.1050	(26)	0.0848	(18)	0.1361	(23)
Penta C	0.1031	(21)	0.0652	(15)	0.0787	(17)	0.1017	(19)
D2S441	0.0884	(17)	0.1051	(27)	0.1098	(22)	0.1187	(22)
D10S1248	0.0989	(20)	0.0693	(17)	0.1050	(21)	0.0919	(16)
D3S1358	0.0758	(14)	0.1090	(28)	0.0904	(20)	0.1485	(25)
D22S1045	0.1239	(23)	0.0552	(12)	0.1743	(29)	0.1030	(20)
F13B	0.1294	(25)	0.0968	(23)	0.1370	(24)	0.3880	(29)
CSF1PO	0.1285	(24)	0.0827	(19)	0.1255	(23)	0.1040	(21)
D5S818	0.1486	(27)	0.0983	(24)	0.1457	(26)	0.0788	(14)
FESFPS	0.1444	(26)	0.0847	(20)	0.1405	(25)	0.1474	(24)
LPL	0.1516	(28)	0.1028	(25)	0.1549	(27)	0.2854	(28)
TPOX	0.1811	(29)	0.0879	(21)	0.1554	(28)	0.2188	(27)
DYS391	0.4430	(30)	0.5603	(30)	0.4411	(30)	0.6992	(30)

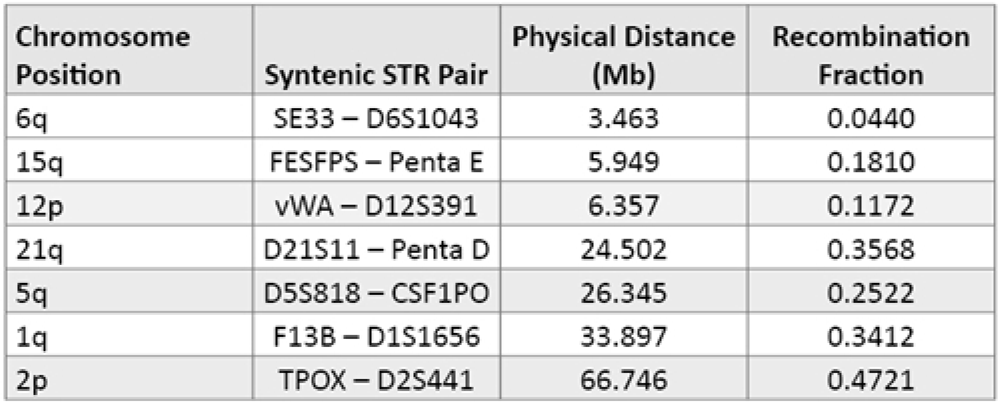
Table 3. Probability of identity values for individual population groups (listed in same order as Table 2).

Click to Enlarge

Penta D ranks 4th in African Americans with a P_I value of 0.0268 and 13th in Asians because of a poorer P_I value of 0.0766. The new loci D1S1656 and D12S391 rank higher in our Caucasian samples (2nd and 3rd) compared to our African Americans (8th and 9th). Some of the lower-ranking loci such as D5S818 and TPOX demonstrate fairly dramatic population variability. D5S818 has a P_I value of 0.0788 with Asians, where it ranks 14th, but a P_I of 0.1486 in Caucasians, where it ranks 27th. TPOX is 21st in African Americans with a P_I of 0.0879 yet it is 29th in Caucasians with a P_I of 0.1811. One of the typically moderate performing markers D13S317 was 11th with our Hispanic data (P_I = 0.0526) but 29th in African Americans (P_I = 0.1342). On the other hand, the new STR locus D22S1045 did much better in African Americans (P_I = 0.0552; 12th overall) than Hispanics (P_I = 0.1743; 29th overall).

This type of variability across different population groups (Table 3) illustrates the need to examine locus variability across populations of interest rather than trying to rely solely on a combined population ranking of genetic markers such as shown in Table 2.

Chromosomal Location of the 29 Autosomal STR Loci: With an increase in the number of autosomal STR loci combined in a multiplex assay comes the potential to have loci that are located in close proximity in the human genome on a specific chromosome—and thus not always inherited independently. Without independent genotype assortment, the simple product rule does not accurately estimate genotype and profile frequencies. Perhaps in an ideal world we would only use loci located on separate chromosomes so that we could have full confidence in independent genotypes from autosomal loci. However, with the prior work on specific STR loci and the availability of primer sets in commercial kits, we sometimes have to work in a less than ideal world. As seen in Tables 1 and 4, there are seven pairs among the 29 autosomal STRs that are located on the same arm of a chromosome (i.e., syntenic loci) (45) .



Chromosome Position	Syntenic STR Pair	Physical Distance (Mb)	Recombination Fraction
6q	SE33 – D6S1043	3.463	0.0440
15q	FESFPS – Penta E	5.949	0.1810
12p	vWA – D12S391	6.357	0.1172
21q	D21S11 – Penta D	24.502	0.3568
5q	D5S818 – CSF1PO	26.345	0.2522
1q	F13B – D1S1656	33.897	0.3412
2p	TPOX – D2S441	66.746	0.4721

Table 4. Syntenic STR loci (i.e., those markers found on the same arm of various chromosomes) ranked by their physical distance. The physical distance in megabases (Mb) and recombination fraction are from reference 45. As noted by Phillips, the recombination activity between two genetic markers is more important than the physical distance in terms of ensuring independent genotype assortments in multiplex DNA profiles.

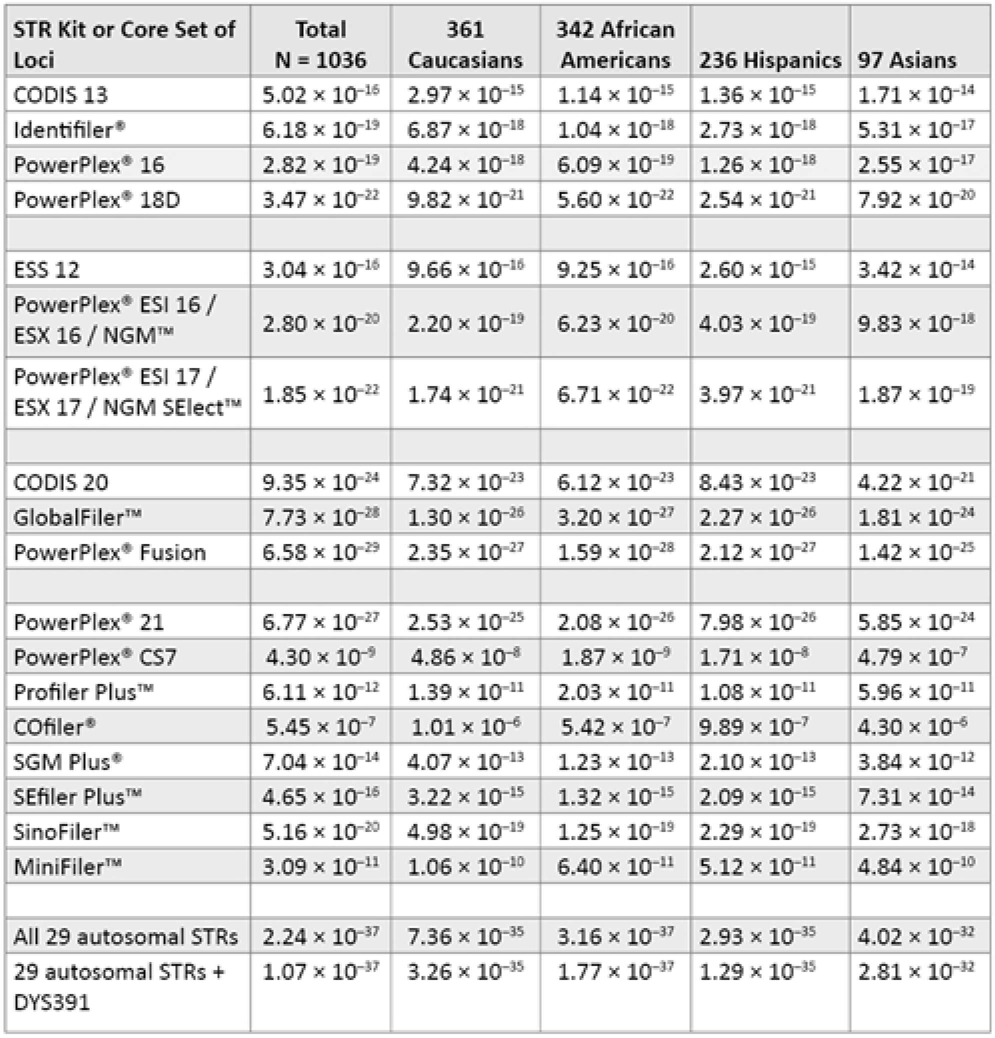
Click to Enlarge

Loci that originally were characterized and developed for separate kits have now been combined as core locus expansion has occurred. The first of these locus “conflicts” was the adoption of D12S391 during the 2009 ESS expansion. D12S391 is located just over 6 million nucleotides away from vWA on the short arm of human chromosome 12 (Table 4). The classical assumption is that markers should be at least 50 megabases (Mb) apart to ensure recombination between the loci (i.e., effective shuffling of the genome deck of cards with each generation). However, recombination rates are not fixed across the human genome, and recombination hotspots have been discovered (46) . Several studies have shown that the D12S391-vWA loci are in linkage disequilibrium and genotype frequencies can be multiplied because there is no strong association between the loci when examining unrelated individuals (47) (48) (49) (50) . However, when examining closely related individuals it can be safer to work with linked D12S391-vWA haplotype frequencies rather than multiplying the individual genotype frequencies (51) .

Table 4 provides more detail regarding the seven pairs of syntenic STR loci. Some of these syntenic loci have been used extensively without any known adverse effects, such as D5S818 and CSF1PO, which are 26Mb apart on the long arm of chromosome 5. Both D5S818 and CSF1PO are required for the current CODIS 13 core. Likewise, D21S11 and Penta D, which are separated by 24Mb on the long arm of chromosome 21, are part of the PowerPlex® 16 System. TPOX and D2S441, which are on the short arm of chromosome 2, are part of the expanded U.S. loci but separated by more than 50Mb (Table 4). FESFPS and Penta E are both part of the PowerPlex® CS7 System, and FESFPS-Penta E haplotypes have been studied to aid kinship analysis (52) . Currently, there are no STR kits that utilize SE33 and D6S1043 or F13B and D1S1656 in the same multiplex assay.

Exploring Capabilities of Sets of Loci Present in STR Kits: With the probability of identity values for different loci and an understanding of potential challenges with syntenic loci, we can examine how various sets of markers would perform across the four U.S. population groups and our combined set of data.

Table 5 contains the probability of identity values found by multiplying individual locus P_I values together to meet the CODIS 13, ESS 12 and CODIS 20 requirements (see Table 1 for the loci present in each set). Each of these core sets of loci represents a minimum requirement in terms of loci to include in a commercial kit. It is important to keep in mind that decisions have been made by manufacturers regarding the specific combinations of loci beyond the core requirements. In addition, relative sizes of the same STR loci may be shifted so that loci that are small in one kit may be larger in another kit [12] . Thus, equivalent loci may not be successfully recovered from badly degraded DNA samples. With these caveats, we can discuss the relative performance of loci present in various STR typing kits.



STR Kit or Core Set of Loci	Total N = 1036	361 Caucasians	342 African Americans	236 Hispanics	97 Asians
CODIS 13	5.02×10^{-16}	2.97×10^{-15}	1.14×10^{-15}	1.36×10^{-15}	1.71×10^{-14}
Identifiler®	6.18×10^{-19}	6.87×10^{-18}	1.04×10^{-18}	2.73×10^{-18}	5.31×10^{-17}
PowerPlex® 16	2.82×10^{-19}	4.24×10^{-18}	6.09×10^{-19}	1.26×10^{-18}	2.55×10^{-17}
PowerPlex® 18D	3.47×10^{-22}	9.82×10^{-21}	5.60×10^{-22}	2.54×10^{-21}	7.92×10^{-20}
ESS 12	3.04×10^{-16}	9.66×10^{-16}	9.25×10^{-16}	2.60×10^{-15}	3.42×10^{-14}
PowerPlex® ESI 16 / ESX 16 / NGM™	2.80×10^{-20}	2.20×10^{-19}	6.23×10^{-20}	4.03×10^{-19}	9.83×10^{-18}
PowerPlex® ESI 17 / ESX 17 / NGM SElect™	1.85×10^{-22}	1.74×10^{-21}	6.71×10^{-22}	3.97×10^{-21}	1.87×10^{-19}
CODIS 20	9.35×10^{-24}	7.32×10^{-23}	6.12×10^{-23}	8.43×10^{-23}	4.22×10^{-21}
GlobalFiler™	7.73×10^{-28}	1.30×10^{-26}	3.20×10^{-27}	2.27×10^{-26}	1.81×10^{-24}
PowerPlex® Fusion	6.58×10^{-29}	2.35×10^{-27}	1.59×10^{-28}	2.12×10^{-27}	1.42×10^{-25}
PowerPlex® 21	6.77×10^{-27}	2.53×10^{-25}	2.08×10^{-26}	7.98×10^{-26}	5.85×10^{-24}
PowerPlex® CS7	4.30×10^{-9}	4.86×10^{-8}	1.87×10^{-9}	1.71×10^{-8}	4.79×10^{-7}
Profiler Plus™	6.11×10^{-12}	1.39×10^{-11}	2.03×10^{-11}	1.08×10^{-11}	5.96×10^{-11}
COfiler®	5.45×10^{-7}	1.01×10^{-6}	5.42×10^{-7}	9.89×10^{-7}	4.30×10^{-6}
SGM Plus®	7.04×10^{-14}	4.07×10^{-13}	1.23×10^{-13}	2.10×10^{-13}	3.84×10^{-12}
SEfiler Plus™	4.65×10^{-16}	3.22×10^{-15}	1.32×10^{-15}	2.09×10^{-15}	7.31×10^{-14}
SinoFiler™	5.16×10^{-20}	4.98×10^{-19}	1.25×10^{-19}	2.29×10^{-19}	2.73×10^{-18}
MiniFiler™	3.09×10^{-11}	1.06×10^{-10}	6.40×10^{-11}	5.12×10^{-11}	4.84×10^{-10}
All 29 autosomal STRs	2.24×10^{-37}	7.36×10^{-35}	3.16×10^{-37}	2.93×10^{-35}	4.02×10^{-32}
29 autosomal STRs + DYS391	1.07×10^{-37}	3.26×10^{-35}	1.77×10^{-37}	1.29×10^{-35}	2.81×10^{-32}

Table 5. Probability of identity for various sets of STR loci, which illustrates the resolving power of different STR typing kits on the same DNA samples. See Table 1 for combinations of STR loci present in each kit. Amelogenin, which offers little to the probability of identity, was not included in the CODIS 20 calculations.

Click to Enlarge

The two most widely used kits currently in the United States are Identifiler® and PowerPlex® 16. These STR kits have the CODIS 13 core loci and the sex-typing marker amelogenin in common and differ by the D2S1338 and D19S433 loci in Identifiler® and the Penta D and Penta E loci in PowerPlex® 16. Note that there is very little difference in performance between Identifiler® and PowerPlex® 16. From the full 1036 data set, P_I values for Identifiler® (6.18×10^{-19}) and for PowerPlex® 16 (2.82×10^{-19}) are within the same order of magnitude. The addition of D2S1338 and D19S433 to the PowerPlex® 16 loci results in a P_I value that is about three orders of magnitude better ($P_I = 3.47 \times 10^{-22}$) for the PowerPlex® 18D System.

The STR kits developed to meet the ESS 12 requirements also include D2S1338, D19S433 and D16S539 [53]. Thus, the PowerPlex® ESI 16, PowerPlex® ESX 16 and NGM™ kits all contain the same 15 STRs and amelogenin and exhibit the same P_I values (Table 5). PowerPlex® ESI 17, PowerPlex® ESX 17 and NGM SElect™ kits all add the SE33 locus, which improves their power of discrimination by almost two orders of magnitude (2.80×10^{-20} to 1.85×10^{-22}). Developmental validation studies for these three kits have been published [20] [54] [55].

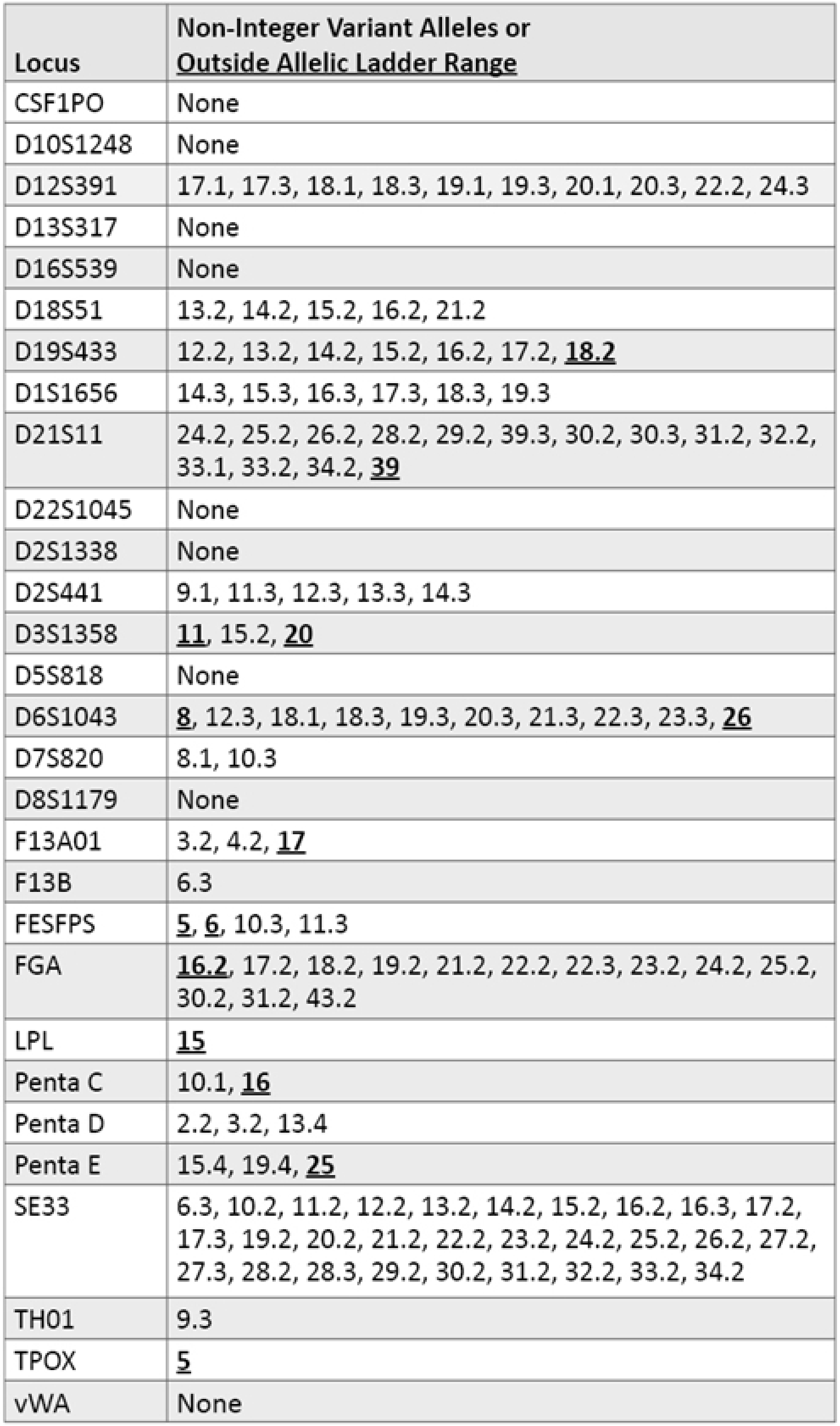
The goal of creating a multiplex to meet the proposed CODIS 20 loci (4) (5) has been met and exceeded with 24plex assays that were recently released. The GlobalFiler™ kit (Life Technologies, 2012; see www.slideshare.net/Lifetech_HID/globalfiler-webinar) is a 6-dye multiplex that includes the proposed CODIS 20 plus TPOX, D22S1045, SE33 and an insertion/deletion marker from the long arm of the Y chromosome to aid amelogenin result verification. PowerPlex® Fusion (Promega, 2012; see www.promega.com/products/pm/genetic-identity/powerplex-fusion) is a 5-dye multiplex covering the CODIS 20 plus TPOX, D22S1045, Penta D and Penta E. Thus, the STR differences between these 24plexes are SE33 for GlobalFiler™ and the two pentanucleotide loci for PowerPlex® Fusion. From our data set calculations, it appears that PowerPlex® Fusion with Penta D and Penta E will provide a slightly better probability of identity than the GlobalFiler™ kit with SE33 (Table 5).

Table 5 also includes the levels of variability that can be expected with other Promega and Applied Biosystems kits described in Table 1 as well as the entire set of 29 autosomal STRs plus DYS391 (under the assumption of independence between all loci).

As can be seen in Tables 2 and 3, SE33 is by far the most polymorphic locus among the 29 autosomal STR loci used in commercial STR kits currently supplied by Applied Biosystems and Promega (56) . A review of the SE33 literature has found more than 175 alleles when internal sequence rearrangements are included (7) . A number of single-base variants exist in SE33, which can require more care in the capillary electrophoresis separations needed to accurately type DNA samples at this locus. In addition, flanking-region mutations have affected amplicon mobility and led to discordant results with different PCR primer pairs. In particular, a G→A mutation 68bp downstream of the SE33 repeat region has been detected that can cause a mobility shift in PowerPlex® ESI 17 relative to PowerPlex® ESX 17 and NGM SElect™ SE33 alleles (57) (58) (59) . To remove this SE33 discordance issue, the PowerPlex® ESI 17 Pro System was developed with redesigned primers (18) .

The reason that SE33 is so variable is that it has a high mutation rate. At 0.64%, the rate is more than three times higher than the average mutation rate of most tetranucleotide STR loci (see www.cstl.nist.gov/strbase/mutation.htm). While the large numbers of alleles make SE33 excellent for mixture detection and interpretation, addressing relationship testing across one or more generations may be more challenging with the higher mutation rate of SE33. However, because SE33 is a core locus for the German national DNA database, an increasing number of labs in Europe are beginning to include it.

Variant Alleles Observed in Autosomal STR Loci: A number of variant (i.e., non-integer repeat) alleles were observed in our data set (Table 6). In some cases, these variant alleles are included within kit allelic ladders. Table 6 also describes alleles observed that migrated outside the range of an allelic ladder for at least one of the STR kits examined. Not surprisingly, the more variable loci have a greater number of variant alleles. For example, SE33 possesses a series of x.2 alleles ranging almost consecutively from 10.2 through 34.2. There is very little information described in the literature so far with D6S1043. We found a number of x.3 alleles with D6S1043 and an 18.1 variant in our data set. D12S391 had both x.1 and x.3 alleles and a 22.2 variant. In total, 20 of the 29 autosomal loci had some variant alleles.



Locus	Non-Integer Variant Alleles or Outside Allelic Ladder Range
CSF1PO	None
D10S1248	None
D12S391	17.1, 17.3, 18.1, 18.3, 19.1, 19.3, 20.1, 20.3, 22.2, 24.3
D13S317	None
D16S539	None
D18S51	13.2, 14.2, 15.2, 16.2, 21.2
D19S433	12.2, 13.2, 14.2, 15.2, 16.2, 17.2, **18.2**
D1S1656	14.3, 15.3, 16.3, 17.3, 18.3, 19.3
D21S11	24.2, 25.2, 26.2, 28.2, 29.2, 39.3, 30.2, 30.3, 31.2, 32.2, 33.1, 33.2, 34.2, **39**
D22S1045	None
D2S1338	None
D2S441	9.1, 11.3, 12.3, 13.3, 14.3
D3S1358	**11**, 15.2, **20**
D5S818	None
D6S1043	**8**, 12.3, 18.1, 18.3, 19.3, 20.3, 21.3, 22.3, 23.3, **26**
D7S820	8.1, 10.3
D8S1179	None
F13A01	3.2, 4.2, **17**
F13B	6.3
FESFPS	**5**, **6**, 10.3, 11.3
FGA	**16.2**, 17.2, 18.2, 19.2, 21.2, 22.2, 22.3, 23.2, 24.2, 25.2, 30.2, 31.2, 43.2
LPL	**15**
Penta C	10.1, **16**
Penta D	2.2, 3.2, 13.4
Penta E	15.4, 19.4, **25**
SE33	6.3, 10.2, 11.2, 12.2, 13.2, 14.2, 15.2, 16.2, 16.3, 17.2, 17.3, 19.2, 20.2, 21.2, 22.2, 23.2, 24.2, 25.2, 26.2, 27.2, 27.3, 28.2, 28.3, 29.2, 30.2, 31.2, 32.2, 33.2, 34.2
TH01	9.3
TPOX	**5**
vWA	None

Table 6. Variant alleles observed in autosomal STR loci.

Click to Enlarge

Y-Chromosome STR Markers

Y-STR Loci in PowerPlex® Y23: Until the recent release of PowerPlex® Y23, the Yfiler® kit from Applied Biosystems was the largest commercially available Y-STR multiplex. The Yfiler® kit amplifies 17 regions of the Y chromosome [(60)] although some researchers cite Yfiler® as having only 16 loci. The highly polymorphic Y-STR locus DYS385 is duplicated and produces two amplicons (termed “a” and “b”) with a single PCR primer pair. PowerPlex® Y23, which was released by Promega in July 2012, contains six additional loci and provides better resolution of haplotypes from unrelated males [(61)] . Figure 1 displays the relative locations along the Y chromosome of the 22 (or 23 if the duplicated DYS385 is counted twice) Y-STR loci present in the PowerPlex® Y23 System. The six new loci that go beyond the ones present in Yfiler® are indicated in bold font.

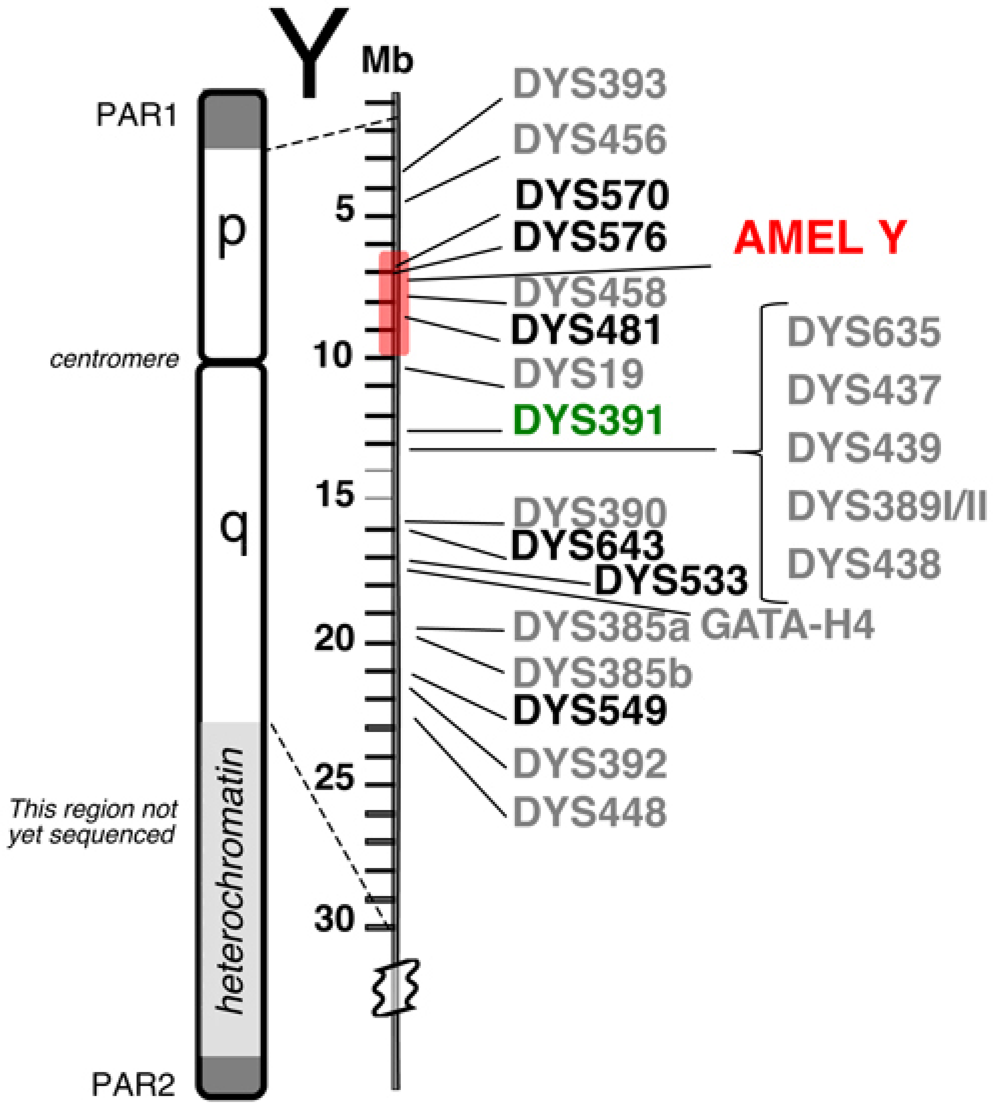
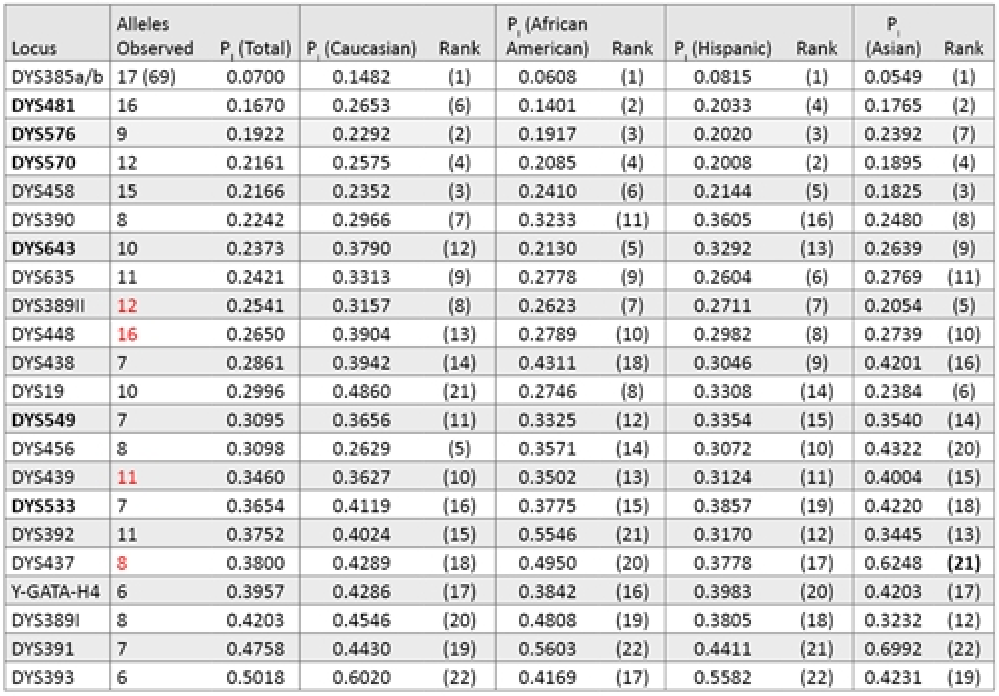


Figure 1. Relative positions of 23 Y-STR loci available in the PowerPlex® Y23 System. The six new loci are shown in bold font. PAR1 and PAR2 are pseudo-autosomal regions on the tips of the Y chromosome that recombine with the X chromosome. The shaded region around AMEL Y can sometimes be deleted (e.g., reference 63), causing loci such as DYS458 to be missing from an otherwise full Y-STR profile. DYS391 is located a sufficient distance away to avoid deletions affecting AMEL Y.

Relative Performance of PowerPlex® Y23 Loci: Table 7 lists the individual Y-STR locus probability of identity values. Typically the metric used for Y-STR markers is marker diversity, which is effectively one minus the probability of identity used here. The overall P_I values are much lower for Y-STR loci compared to autosomal markers because only a single allele is typically present at each locus—with the notable exception of DYS385a/b. Again, as observed in Table 3 with the autosomal loci, there are population group differences in marker variability. For example, even though DYS385a/b is the top-ranked Y-STR locus in all population groups, it is less polymorphic in Caucasians (P_I = 0.1482) compared to the other groups (e.g., Asian P_I = 0.0549). Likewise, DYS19 ranks lower in Caucasians compared to other groups. On the other hand, DYS456 ranks relatively higher in Caucasians. Of the six new loci, the trinucleotide DYS481 is the highest ranked single-copy Y-STR. DYS576 and DYS570 also rank near the top in every population (Table 7).

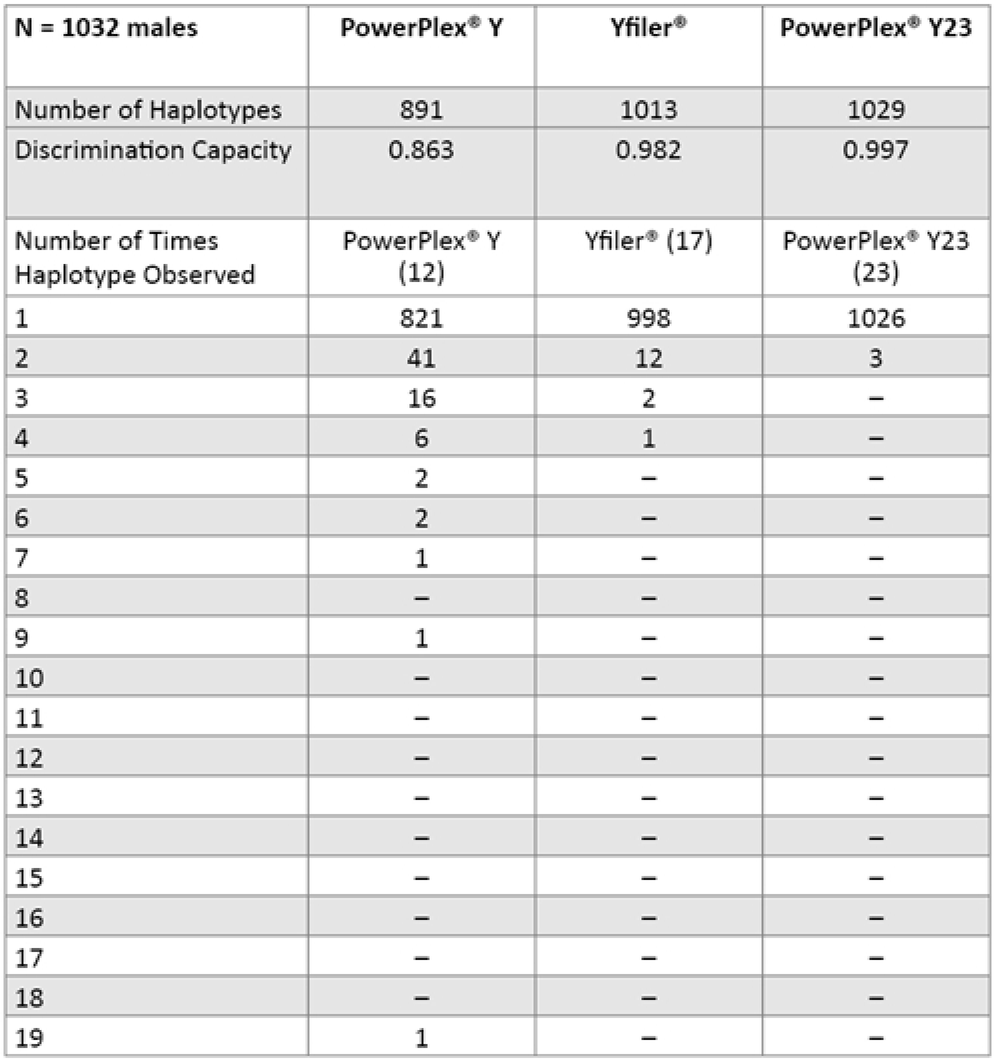


Locus	Alleles Observed	P_I (Total)	P_I (Caucasian)	Rank	P_I (African American)	Rank	P_I (Hispanic)	Rank	P_I (Asian)	Rank
DYS385a/b	17 (69)	0.0700	0.1482	(1)	0.0608	(1)	0.0815	(1)	0.0549	(1)
DYS481	16	0.1670	0.2653	(6)	0.1401	(2)	0.2033	(4)	0.1765	(2)
DYS576	9	0.1922	0.2292	(2)	0.1917	(3)	0.2020	(3)	0.2392	(7)
DYS570	12	0.2161	0.2575	(4)	0.2085	(4)	0.2008	(2)	0.1895	(4)
DYS458	15	0.2166	0.2352	(3)	0.2410	(6)	0.2144	(5)	0.1825	(3)
DYS390	8	0.2242	0.2966	(7)	0.3233	(11)	0.3605	(16)	0.2480	(8)
DYS643	10	0.2373	0.3790	(12)	0.2130	(5)	0.3292	(13)	0.2639	(9)
DYS635	11	0.2421	0.3313	(9)	0.2778	(9)	0.2604	(6)	0.2769	(11)
DYS389II	12	0.2541	0.3157	(8)	0.2623	(7)	0.2711	(7)	0.2054	(5)
DYS448	16	0.2650	0.3904	(13)	0.2789	(10)	0.2982	(8)	0.2739	(10)
DYS438	7	0.2861	0.3942	(14)	0.4311	(18)	0.3046	(9)	0.4201	(16)
DYS19	10	0.2996	0.4860	(21)	0.2746	(8)	0.3308	(14)	0.2384	(6)
DYS549	7	0.3095	0.3656	(11)	0.3325	(12)	0.3354	(15)	0.3540	(14)
DYS456	8	0.3098	0.2629	(5)	0.3571	(14)	0.3072	(10)	0.4322	(20)
DYS439	11	0.3460	0.3627	(10)	0.3502	(13)	0.3124	(11)	0.4004	(15)
DYS533	7	0.3654	0.4119	(16)	0.3775	(15)	0.3857	(19)	0.4220	(18)
DYS392	11	0.3752	0.4024	(15)	0.5546	(21)	0.3170	(12)	0.3445	(13)
DYS437	8	0.3800	0.4289	(18)	0.4950	(20)	0.3778	(17)	0.6248	**(21)**
Y-GATA-H4	6	0.3957	0.4286	(17)	0.3842	(16)	0.3983	(20)	0.4203	(17)
DYS389I	8	0.4203	0.4546	(20)	0.4808	(19)	0.3805	(18)	0.3232	(12)
DYS391	7	0.4758	0.4430	(19)	0.5603	(22)	0.4411	(21)	0.6992	(22)
DYS393	6	0.5018	0.6020	(22)	0.4169	(17)	0.5582	(22)	0.4231	(19)

Table 7. Y-STR locus variability across 1032 males in our data set (loci are ranked by their probability of identity values). There are 17 different alleles seen in our data set for DYS385 in 69 combinations of the “a” and “b” amplicons. Numbers highlighted in red include a deletion null allele (see Table 9). The six new loci present in PowerPlex® Y23 are in bold font.

Click to Enlarge

Haplotype Resolution with Various Y-STR Kits: An examination of how well subsets of data perform aids understanding of the relative value of adding new Y-STR loci to increase haplotype resolution (62) . In our data set of 1032 males, there were 70 sets of samples that contained two or more matching samples when examined with the 12 Y-STRs present in PowerPlex® Y (Table 8). Increasing the number of loci to 17 with the Yfiler® loci reduced the number of sample sets to 15 that could not be subdivided. Finally, when all 23 Y-STRs were applied to the data set, only three sets of sample pairs remained that could not be subdivided into individual sample haplotypes. There are 1026 PowerPlex® Y23 haplotypes occurring once in our dataset and three separate unresolved pairs. By comparison, the 17 Yfiler® loci produce 998 singletons, 12 pairs, two triplicates and one quadruplicate. Thus, in our dataset, PowerPlex® Y23 has a discrimination capacity of 99.7% (1029 types/1032 samples) compared to 98.2% (1013 types/1032 samples) with Yfiler® and 86.3% (891 types/1032 samples) with the original 12-locus PowerPlex® Y System (Table 8).



N = 1032 males	PowerPlex® Y	Yfiler®	PowerPlex® Y23
Number of Haplotypes	891	1013	1029
Discrimination Capacity	0.863	0.982	0.997
Number of Times Haplotype Observed	PowerPlex® Y (12)	Yfiler® (17)	PowerPlex® Y23 (23)
1	821	998	1026
2	41	12	3
3	16	2	–
4	6	1	–
5	2	–	–
6	2	–	–
7	1	–	–
8	–	–	–
9	1	–	–
10	–	–	–
11	–	–	–
12	–	–	–
13	–	–	–
14	–	–	–
15	–	–	–
16	–	–	–
17	–	–	–
18	–	–	–
19	1	–	–

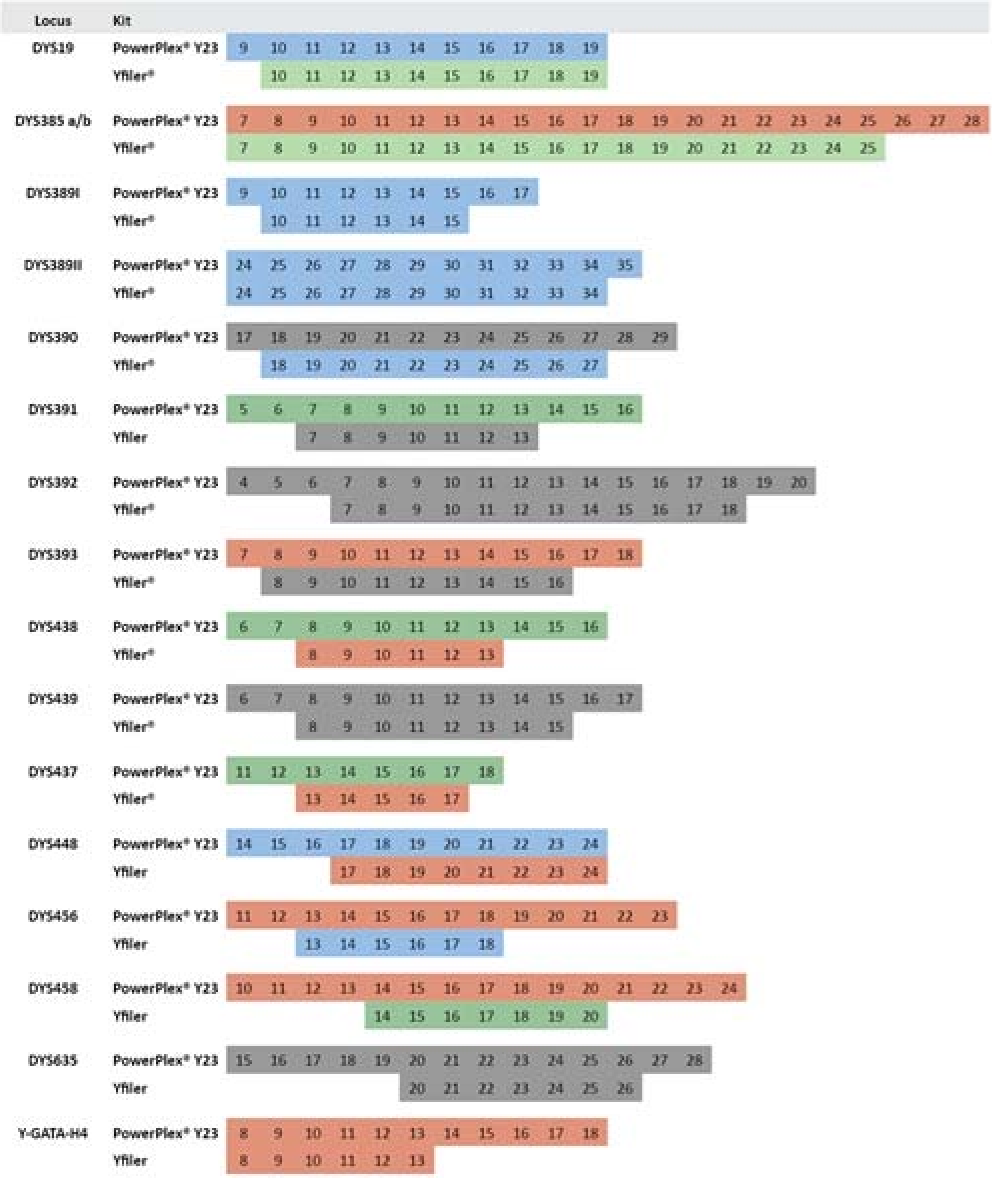
Table 8. Number of unique and shared haplotypes observed with various combinations of Y-STR loci across 1032 U.S. population samples.

Click to Enlarge

Further Examination of Matching PowerPlex® Y23 Samples: Three sets of sample pairs matched across all 23 Y-STR alleles amplified with PowerPlex® Y23: two Hispanic samples (PT84348 and ZT80369), another set of two Hispanics (C87H and GT37828), and an African American sample (ZT79304) and Hispanic sample (PT84253). These samples were carefully examined to see if they could be close male relatives.

In a previous study, the first set of Hispanic samples (PT84348 and ZT80369) could not be resolved with 37 Y-STRs [(40)] . However, recent examination with 14 rapidly mutating Y-STR loci described by Manfred Kayser's group [(28)] showed three mutations (data not shown). The mtDNA haplogroups differ, and autosomal STR kinship calculations did not indicate any conclusive close relationships. The highest likelihood ratio (LR) was 0.2 for a first cousin assumption. On the second set of Hispanic samples, the highest likelihood ratio using autosomal STR information was 0.228 for a first cousin hypothesis (parent-child and full-sibling LRs are zero). Kinship calculations with the ZT79304-PT84253 sample pair found a likelihood ratio of 0 for parent-child, 0 for full-sib, 0.029 for half-sib and 0.568 for first cousin hypotheses. The mtDNA haplogroups were also different between these two samples. Previous testing with 37 Y-STRs [(40)] showed that this sample pair could be subdivided with results from DYS505 (allele 13 vs. 11), DYS532 (12 vs. 11) or DYS540 (12 vs. 13). Additional testing with rapidly mutating Y-STR loci showed three mutations between these two samples (data not shown).

PowerPlex® Y23 and Yfiler® Allelic Ladders Compared: In addition to adding six more polymorphic Y-STR loci, the PowerPlex® Y23 System includes more alleles in each allelic ladder. Across the 16 loci in common with Yfiler®, there are 66 additional alleles present in the PowerPlex® Y23 allelic ladders (Figure 2). These additional alleles can be helpful in appropriately designating rare small or large Y-STR alleles. In our data set of 1032 males, there are 15 alleles that fall outside the range of the Yfiler® allelic ladders alleles but within the range of the PowerPlex® Y23 alleles. These Yfiler® off-ladder alleles include DYS389I allele 9; DYS437 allele 18; DYS438 allele 14; DYS448 alleles 14, 15 and 16; DYS456 alleles 12 and 19; DYS458 alleles 12, 13 and 21; DYS635 alleles 17, 19 and 27; and Y-GATA-H4 allele 14.



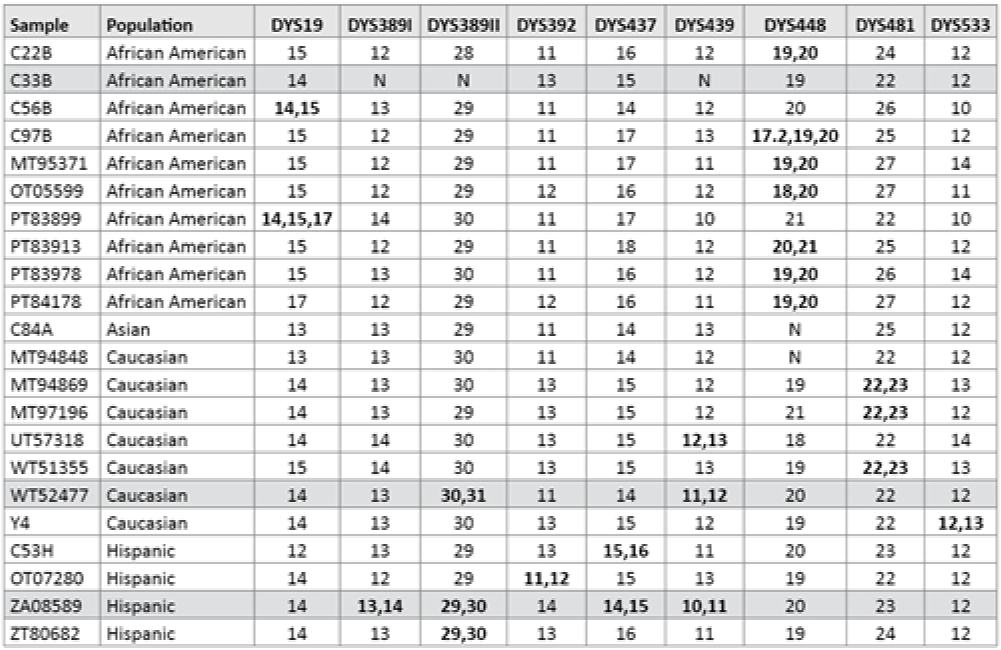
Locus	Kit	
DYS19	PowerPlex® Y23	🟦 9 10 11 12 13 14 15 16 17 18 19
	Yfiler®	🟩 10 11 12 13 14 15 16 17 18 19
DYS385 a/b	PowerPlex® Y23	🟥 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28
	Yfiler®	🟩 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25
DYS389I	PowerPlex® Y23	🟦 9 10 11 12 13 14 15 16 17
	Yfiler®	🟦 10 11 12 13 14 15
DYS389II	PowerPlex® Y23	🟦 24 25 26 27 28 29 30 31 32 33 34 35
	Yfiler®	🟦 24 25 26 27 28 29 30 31 32 33 34
DYS390	PowerPlex® Y23	⬛ 17 18 19 20 21 22 23 24 25 26 27 28 29
	Yfiler®	🟦 18 19 20 21 22 23 24 25 26 27
DYS391	PowerPlex® Y23	🟩 5 6 7 8 9 10 11 12 13 14 15 16
	Yfiler	⬛ 7 8 9 10 11 12 13
DYS392	PowerPlex® Y23	⬛ 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
	Yfiler®	⬛ 7 8 9 10 11 12 13 14 15 16 17 18
DYS393	PowerPlex® Y23	🟥 7 8 9 10 11 12 13 14 15 16 17 18
	Yfiler®	⬛ 8 9 10 11 12 13 14 15 16
DYS438	PowerPlex® Y23	🟩 6 7 8 9 10 11 12 13 14 15 16
	Yfiler®	🟥 8 9 10 11 12 13
DYS439	PowerPlex® Y23	⬛ 6 7 8 9 10 11 12 13 14 15 16 17
	Yfiler®	⬛ 8 9 10 11 12 13 14 15
DYS437	PowerPlex® Y23	🟩 11 12 13 14 15 16 17 18
	Yfiler®	🟥 13 14 15 16 17
DYS448	PowerPlex® Y23	🟦 14 15 16 17 18 19 20 21 22 23 24
	Yfiler	🟥 17 18 19 20 21 22 23 24
DYS456	PowerPlex® Y23	🟥 11 12 13 14 15 16 17 18 19 20 21 22 23
	Yfiler	🟦 13 14 15 16 17 18
DYS458	PowerPlex® Y23	🟥 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
	Yfiler	🟩 14 15 16 17 18 19 20
DYS635	PowerPlex® Y23	⬛ 15 16 17 18 19 20 21 22 23 24 25 26 27 28
	Yfiler	⬛ 20 21 22 23 24 25 26
Y-GATA-H4	PowerPlex® Y23	🟥 8 9 10 11 12 13 14 15 16 17 18
	Yfiler	🟥 8 9 10 11 12 13

Figure 2. Comparison of allelic ladder alleles between PowerPlex® Y23 and Yfiler® kits. Numbers within boxes indicate allele repeat number. Colors represent dye color for the locus label in the multiplex assay.

Click to Enlarge

Variant Alleles Observed in Y-STR Loci: The following Y-STR microvariant alleles were detected within our sample set: DYS385 alleles 12.2 and 13.2; DYS448 alleles 17.2 and 18.4; DYS458 alleles 16.2, 17.2, 18.2, 19.2, and 21.2; DYS481 allele 25.1; DYS635 allele 21.3; and DYS643 allele 11.1. The sequence structure of the DYS635 21.3 and DYS643 11.1 were previously reported (27) . We have also previously reported on flanking region sequence variation in DYS385 that can exist outside of some primer pairs (26) .

Copy Number Variants (Duplications, Deletions and Tri-Allelic Patterns): As previously reported, we have several duplications and two triplications at various Y-STR loci within our data set (11) (25) . Of course, DYS385 a/b is a normal duplication where different sized alleles are regularly observed (7) (26) . In some cases, duplications occur at multiple loci within a sample (25) . Table 9 lists the Y-STR duplications, triplications and deletions observed in our data set. For example, within sample PT83899, a 14,15,17 tri-allelic pattern with well balanced peak heights was observed at DYS19. A PowerPlex® Y profile of this sample is available in a previous report [25] .



Sample	Population	DYS19	DYS389I	DYS389II	DYS392	DYS437	DYS439	DYS448	DYS481	DYS533
C22B	African American	15	12	28	11	16	12	**19,20**	24	12
C33B	African American	14	N	N	13	15	N	19	22	12
C56B	African American	**14,15**	13	29	11	14	12	20	26	10
C97B	African American	15	12	29	11	17	13	**17.2,19,20**	25	12
MT95371	African American	15	12	29	11	17	11	**19,20**	27	14
OT05599	African American	15	12	29	12	16	12	**18,20**	27	11
PT83899	African American	**14,15,17**	14	30	11	17	10	21	22	10
PT83913	African American	15	12	29	11	18	12	**20,21**	25	12
PT83978	African American	15	13	30	11	16	12	**19,20**	26	14
PT84178	African American	17	12	29	12	16	11	**19,20**	27	12
C84A	Asian	13	13	29	11	14	13	N	25	12
MT94848	Caucasian	13	13	30	11	14	12	N	22	12
MT94869	Caucasian	14	13	30	13	15	12	19	**22,23**	13
MT97196	Caucasian	14	13	29	13	15	12	21	**22,23**	12
UT57318	Caucasian	14	14	30	13	15	**12,13**	18	22	14
WT51355	Caucasian	15	14	30	13	15	13	19	**22,23**	13
WT52477	Caucasian	14	13	**30,31**	11	14	**11,12**	20	22	12
Y4	Caucasian	14	13	30	13	15	12	19	22	**12,13**
C53H	Hispanic	12	13	29	13	**15,16**	11	20	23	12
OT07280	Hispanic	14	12	29	**11,12**	15	13	19	22	12
ZA08589	Hispanic	14	**13,14**	**29,30**	14	**14,15**	**10,11**	20	23	12
ZT80682	Hispanic	14	13	**29,30**	13	16	11	19	24	12

Table 9. Samples with Y-STR duplications or triplications. N = null or no result due to deletion. Shaded rows contain multiple duplications or deletions in the same sample (see reference 25).

Click to Enlarge

In African American sample C33B, DYS389I, DYS389II and DYS439 fail to amplify (Table 9). This section of the Y chromosome is likely deleted, as the father of this individual also exhibits the same loss of this portion of the Y chromosome [(11)] . While the DYS533 12,13 duplication was previously reported [(27)] , the DYS481 22,23 duplication is new. We observed some form of duplication at nine of the loci amplified with PowerPlex® Y23 (Table 9).

Duplications or deletions were not observed to this degree with the autosomal STR loci examined. We only found a single tri-allelic pattern at TPOX among the 29 autosomal STR loci examined across the 1036 samples in this study. African American sample OT05588 had a fairly equal balanced TPOX 9,10,11 tri-allelic pattern when tested with multiple STR kits containing different PCR primer pairs (Figure 3).

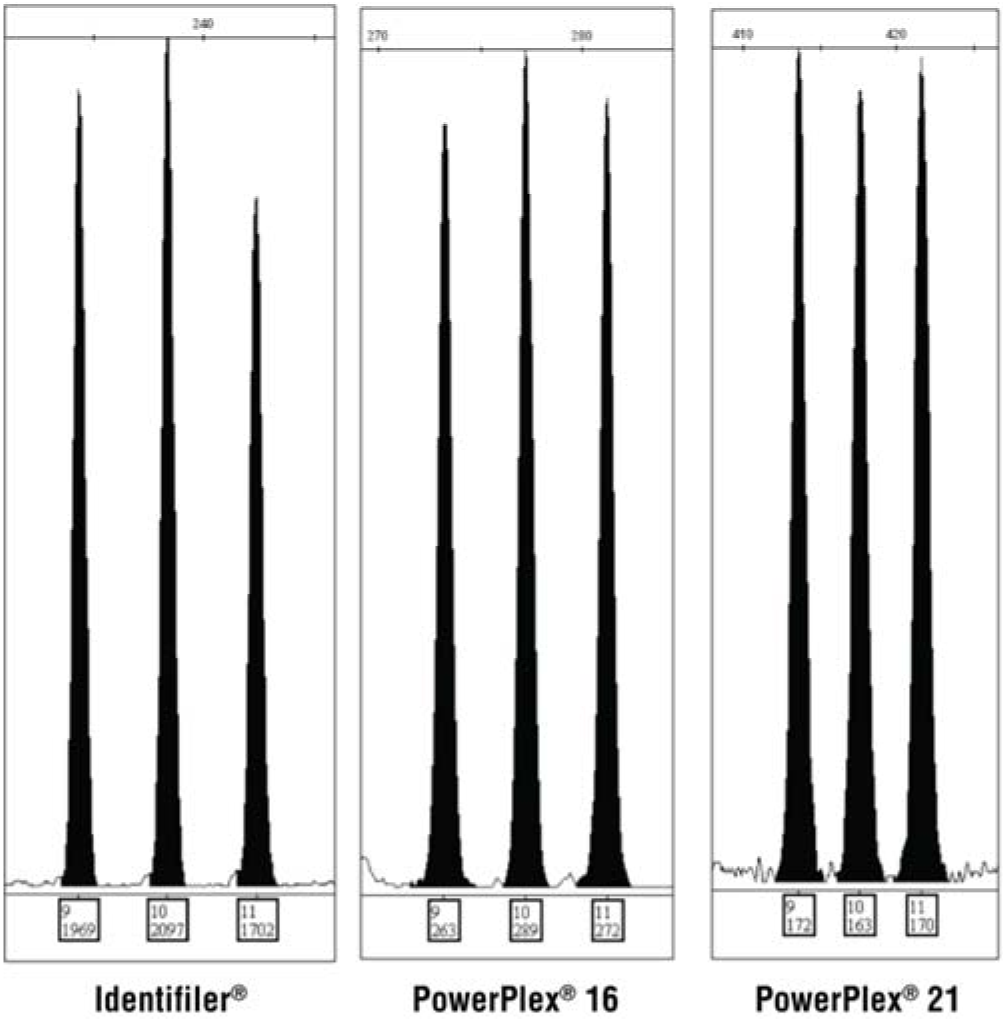


Figure 3. TPOX tri-allelic pattern observed in African American sample OT05588 with three STR kits. Labels under each peak display allele call and peak height.

Click to Enlarge

Amelogenin Y Confirmation and DYS391 Variation: We have four female DNA samples in our data set that exhibit an amelogenin X,X type and no Y-STR profile. None of our 1032 males possessed a missing amelogenin Y allele. Therefore, we were unable to examine the capabilities of gender-confirmation assays that involve amplifying a separate region of the Y chromosome like DYS391 (4) . Samples from other regions of the world will likely benefit more from amelogenin Y confirmation assays (63) (64) .

In all 1032 male DNA samples, DYS391 was successfully amplified. We measured seven different DYS391 alleles: 7 (observed once), 8 (observed twice), 9 (observed 33 times), 10 (observed 607 times), 11 (observed 370 times), 12 (observed 18 times) and 13 (observed once). Thus, it appears that about 95% of the time either allele 10 or allele 11 will be observed. Our DYS391 allele distributions were very similar to those found in the YHRD database, which involves over 700 worldwide populations and more than 104,000 samples (see www.yhrd.org/Research/Loci/DYS391/)—namely alleles 10 and 11 are seen most frequently. DYS391 was by far the worst performer in terms of locus variability in U.S. population groups (Table 3) and did not add much in terms of an increased power of discrimination (compare last two rows in Table 5). But, the purpose of DYS391 is for gender result verification (4) and was not intended for increasing the power of discrimination in the proposed CODIS 20 set of loci as has been suggested (65) .

Conclusions

The work described in this article continues efforts previously described at the 2007 International Symposium on Human Identification with examining the impact of new autosomal and Y-chromosome STR loci [(40)] . By creating a large data set of U.S. population data across 29 autosomal and 23 Y-STR loci available in STR typing kits from Life Technologies/Applied Biosystems and Promega Corporation, we hope to enable users of these kits with reliable population data to underpin their work. It will be interesting to see how U.S. forensic laboratories adopt these next-generation STR kits. Hopefully our U.S. population data set will be beneficial to future use of these new loci.

For calibration of sizing measurements and traceability of allele calls to a national metrology institute in accordance with ISO 17025 requirements, NIST provides Standard Reference Materials (SRMs). SRM 2391c contains six components that have been certified for numerous autosomal STRs and the 17 Y-STRs in the Yfiler® kit [(66)] . We plan to update the SRM 2391c certificate soon with reference values for the additional six Y-STRs found in PowerPlex® Y23.

Plans for Our Data Release: While this article describes some basic characteristics of our population data, we plan on releasing the full information in several phases. Announcement of population data will be submitted to *Forensic Science International: Genetics* for both the PowerPlex® Y23 information and the 29 autosomal STR loci present in commercially available STR kits to enable citation in a peer-reviewed journal. Autosomal STR genotypes and Y-STR haplotype data will be supplied in Excel® spreadsheets on the STRBase NIST Population Data section (see www.cstl.nist.gov/biotech/strbase/NISTpop.htm) to enable use of our full data set in additional calculations and further studies. Finally, the STR locus allele frequencies will be published in the forthcoming textbook *Advanced Topics in Forensic DNA Typing: Interpretation*, where worked examples will be supplied using this information.

Acknowledgments and Disclaimer

STR kits were kindly provided through collaborations with Applied Biosystems, Promega Corporation and Qiagen for the purpose of collecting concordance data between various commercial kits. This work would not have been possible without support of past and present members of the NIST Human Identity Project Team including Erica Butts, Amy Decker, David Duewer, Kevin Kiesler, Margaret Kline, Kristen Lewis O’Connor, Richard Schoske and Peter Vallone. In particular, David Duewer developed and regularly refined useful Excel® spreadsheets to enable rapid data analysis. Margaret Kline sequenced STR alleles and aided characterization of primer-binding-site mutations and variant alleles. Forensic DNA research conducted at NIST is supported by an interagency agreement between the National Institute of Justice and the NIST Office of Law Enforcement Standards. Points of view in this document are those of the authors and do not necessarily represent the official position or policies of the U.S. Department of Justice. Certain commercial equipment, instruments and materials are identified in order to specify experimental procedures as completely as possible. In no case does such identification imply a recommendation or endorsement by the National Institute of Standards and Technology nor does it imply that any of the materials, instruments or equipment identified are necessarily the best available for the purpose.

References

1. Gill, P. *et al.* (2006) The evolution of DNA databases—Recommendations for new European STR loci. *Forensic Sci. Int.* **156**, 242–4.
2. Gill, P. *et al.* (2006) New multiplexes for Europe—Amendments and clarification of strategic development. *Forensic Sci. Int.* **163**, 155–7.
3. Schneider, P.M. (2009) Expansion of the European Standard Set of DNA database loci—the current situation *Profiles in DNA* **12(1)**, 6–7.
4. Hares, D.R. (2012) Expanding the CODIS core loci in the United States. *Forensic Sci. Int. Genet.* **6**,

e52–e54.

5. Hares, D.R. (2012) Addendum to expanding the CODIS core loci in the United States. *Forensic Sci. Int. Genet.* **6**, e135.
6. FBI (2012) Planned process and timeline for implementation of additional CODIS core loci. (www.fbi.gov/about-us/lab/codis/planned-process-and-timeline-for-implementation-of-additional-codis-core-loci).
7. Butler, J.M. (2012) *Advanced Topics in Forensic DNA Typing: Methodology*, Elsevier Academic Press, San Diego.
8. Butler, J.M. and Hill, C.R. (2012) Biology and genetics of new autosomal STR loci useful for forensic DNA analysis. *Forensic Sci. Rev.* **24**, 15–26.
9. Katsanis, S.H. and Wagner, J.K. (2012) Characterization of the standard and recommended CODIS markers. *J. Forensic Sci.* [Epub ahead of print].
10. Butler, J.M. *et al.* (2003) Allele frequencies for 15 autosomal STR loci on U.S. Caucasian, African American, and Hispanic populations. *J. Forensic Sci.* **48**, 908–11.
11. Decker, A.E. *et al.* (2008) Analysis of mutations in father-son pairs with 17 Y-STR loci. *Forensic Sci. Int. Genet.* **2**, e31–5.
12. Hill, C.R. *et al.* (2011) Concordance and population studies along with stutter and peak height ratio analysis for the PowerPlex® ESX 17 and ESI 17 Systems. *Forensic Sci. Int. Genet.* **5**, 269–75.
13. Hill, C.R. *et al.* (2010) Strategies for concordance testing. *Profiles in DNA* (www.promega.com/resources /articles/profiles-in-dna/2010/strategies-for-concordance-testing/).
14. Drábek, J. *et al.* (2004) Concordance study between miniplex assays and a commercial STR typing kit. *J. Forensic Sci.* **49**, 859–60.
15. Hill, C.R. *et al.* (2007) Concordance study between the AmpF/STR® MiniFiler™ PCR amplification kit and conventional STR typing kits. *J. Forensic Sci.* **52**, 870–3.
16. Ensenberger, M.G. *et al.* (2010) Developmental validation of the PowerPlex 16 HS System: An improved 16-locus fluorescent STR multiplex. *Forensic Sci. Int. Genet.* **4**, 257–64.
17. Hill, C.R. *et al.* (2011) Concordance testing comparing STR multiplex kits with a standard data set. *Forensic Sci. Int. Genet. Suppl. Ser.* **3**, e188–9.
18. McLaren, R.S. *et al.* (2012) Improved primer pair for the SE33 locus in the PowerPlex® ESI 17 Pro System. *Profiles in DNA* (www.promega.com/resources/articles/profiles-in-dna/2012/improved-primer-pair-for-the-se33-locus-in-the-powerplex-esi-17-pro-system/).
19. Oostdik, K. *et al.* (2012) Developmental validation of the PowerPlex® 18D System, a rapid STR multiplex for analysis of reference samples. *Forensic Sci. Int. Genet.* [Epub ahead of print].
20. Green, R.L. *et al.* (2012) Developmental validation of the AmpF/STR® NGM SElect™ PCR Amplification Kit: A next-generation STR multiplex with the SE33 locus. *Forensic Sci. Int. Genet.* [Epub ahead of print].
21. Butler, J.M., Shen, Y. and McCord, B.R. (2003) The development of reduced size STR amplicons as tools for analysis of degraded DNA. *J. Forensic Sci.* **48**, 1054–64.
22. Coble, M.D. and Butler, J.M. (2005) Characterization of new miniSTR loci to aid analysis of degraded DNA. *J. Forensic Sci.* **50**, 43–53.
23. Hill, C.R. *et al.* (2008) Characterization of 26 miniSTR loci for improved analysis of degraded DNA samples. *J. Forensic Sci.* **53**, 73–80.

24. Hill, C.R., Butler, J.M and Vallone, P.M. (2009) A 26plex autosomal STR assay to aid human identity testing. *J. Forensic Sci.* **54**, 1008–15.

25. Butler, J.M. *et al.* (2005) Chromosomal duplications along the Y-chromosome and their potential impact on Y-STR interpretation. *J. Forensic Sci.* **50**, 853–9.

26. Schoske, R. *et al.* (2004) High-throughput Y-STR typing of U.S. populations with 27 regions of the Y chromosome using two multiplex PCR assays. *Forensic Sci. Int.* **139**, 107–21.

27. Butler, J.M. *et al.* (2006) Allele frequencies for 27 Y-STR loci with U.S. Caucasian, African American, and Hispanic samples. *Forensic Sci. Int.* **156**, 250–60.

28. Ballantyne, K.N. *et al.* (2012) A new future of forensic Y-chromosome analysis: Rapidly mutating Y-STRs for differentiating male relatives and paternal lineages. *Forensic Sci. Int. Genet.* **6**, 208–18.

29. Vallone, P.M., Decker, A.E. and Butler, J.M. (2005) Allele frequencies for 70 autosomal SNP loci with U.S. Caucasian, African-American, and Hispanic samples. *Forensic Sci. Int.* **149**, 279–86.

30. Lao, O. *et al.* (2010) Evaluating self-declared ancestry of U.S. Americans with autosomal, Y-chromosomal and mitochondrial DNA. *Hum. Mutat.* **31**, E1875–93.

31. Fondevila, M. *et al.* (2012) Revision of the SNP*for*ID 34-plex forensic ancestry test: Assay enhancements, standard reference sample genotypes and extended population studies. *Forensic Sci. Int. Genet.* [Epub ahead of print].

32. Vallone, P.M. and Butler, J.M. (2004) Y-SNP typing of U.S. African American and Caucasian samples using allele-specific hybridization and primer extension. *J. Forensic Sci*. **49**, 723–32.

33. Fondevila, M. *et al.* (2012) Forensic performance of two insertion-deletion marker assays. *Int. J. Legal Med.* **126**, 725–37.

34. Diegoli, T.M. and Coble, M.D. (2011) Development and characterization of two mini-X chromosomal short tandem repeat multiplexes. *Forensic Sci. Int. Genet.* **5**, 415–21.

35. Diegoli, T.M. *et al.* (2011) Allele frequency distribution of twelve X-chromosomal short tandem repeat markers in four U.S. population groups. *Forensic Sci. Int. Genet. Suppl. Ser.* **3**, e481–3.

36. Saunier, J.L. *et al.* (2008) Mitochondrial control region sequences from a U.S. "Hispanic" population sample. *Forensic Sci. Int. Genet.* **2**, e19–23.

37. Diegoli, T.M. *et al.* (2009) Mitochondrial control region sequences from an African American population sample. *Forensic Sci. Int. Genet.* **4**, e45–52.

38. Just, R.S. *et al.* (2008) Complete mitochondrial genome sequences for 265 African American and U.S. "Hispanic" individuals. *Forensic Sci. Int. Genet.* **2**, e45–8.

39. Liu, K. and Muse, S.V. (2005) PowerMarker: An integrated analysis environment for genetic marker analysis. *Bioinformatics* **21**, 2128–9.

40. Butler, J.M. *et al.* (2007) New autosomal and Y-chromosome STR loci: Characterization and potential uses. *Proceedings of the Eighteenth International Symposium on Human Identification*

41. Kline, M.C. *et al.* (2011) STR sequence analysis for characterizing normal, variant, and null alleles. *Forensic Sci. Int. Genet.* **5**, 329–32.

42. Bodner, M. *et al.* (2011) Inspecting close maternal relatedness: Towards better mtDNA population samples in forensic databases. *Forensic Sci. Int. Genet.* **5**, 138–41.

43. Sensabaugh, G.F. (1982) Biochemical markers of individuality. In *Forensic Science Handbook*, Saferstein, R., ed., Prentice-Hall, New York. 338–415.

44. Maha, G.C. and Fuller, J.R. (2011) Allele Frequencies for Penta C in Caucasian, African American, American Indian, Asian and Hispanic Populations. *Profiles in DNA* (www.promega.com/resources/articles /profiles-in-dna/2011/allele-frequencies-for-penta-c/).

45. Phillips, C. *et al.* (2012) The recombination landscape around forensic STRs: Accurate measurement of genetic distances between syntenic STR pairs using HapMap high density SNP data. *Forensic Sci Int Genet.* **6**, 354–65..

46. Clark, A.G., Wang, X. and Matise T. (2010) Contrasting methods of quantifying fine structure of human recombination. *Annu. Rev. Genomics Hum. Genet.* **11**, 45–64.

47. Phillips, C. *et al.* (2011) Analysis of global variability in 15 established and 5 new European Standard Set (ESS) STRs using the CEPH human genome diversity panel. *Forensic Sci. Int. Genet.* **5**, 155–69.

48. Budowle, B. *et al.* (2011) Population genetic analyses of the NGM STR loci. *Int. J. Legal Med.* **125**, 101–9.

49. O'Connor, K.L. *et al.* (2011) Linkage disequilibrium analysis of D12S391 and vWA in U.S. population and paternity samples. *Forensic Sci. Int. Genet.* **5**, 538–40. Erratum in *Forensic Sci. Int. Genet.* **5**, 541–2.

50. Gill, P. *et al.* (2012) An evaluation of potential allelic association between the STRs vWA and D12S391: Implications in criminal casework and applications to short pedigrees. *Forensic Sci. Int. Genet.* **6**, 477–86.

51. O'Connor, K.L. and Tillmar, A.O. (2012) Effect of linkage between vWA and D12S391 in kinship analysis. *Forensic Sci. Int. Genet.* **6**, 840–4.

52. Worthen, C. *et al.* (2012) Evaluation of haplotypes to enhance the power of STR test kits in cases of questioned family relationships. *Profiles in DNA* (www.promega.com/resources/articles/profiles-in-dna /2012/haplotypes-in-family-relationship-testing/).

53. Sprecher, C.J. *et al.* (2009) The PowerPlex® ESX and ESI Systems: Meeting the new European standard. *Profiles in DNA* (www.promega.com/resources/articles/profiles-in-dna/2009/the-powerplex-esx-and-esi-systems-meeting-the-new-european-standard/).

54. Tucker, V.C. *et al.* (2011) Developmental validation of the PowerPlex® ESI 16 and PowerPlex® ESI 17 Systems: STR multiplexes for the new European standard. *Forensic Sci. Int. Genet.* **5**, 436–48.

55. Tucker, V.C. *et al.* (2012) Developmental validation of the PowerPlex® ESX 16 and PowerPlex® ESX 17 Systems. *Forensic Sci. Int. Genet.* **6**, 124–31.

56. Butler, J.M. *et al.* (2009) The single most polymorphic STR locus: SE33 performance in U.S. populations. *Forensic Sci. Int. Genet. Suppl. Ser.* **2**, 23–4.

57. Oldroyd, N. *et al.* (2011) Development of the AmpF/STR® NGM SElect™ Kit: New sequence discoveries and implications for genotype discordance. *Forensic News* [Internet] [cited: 2011, December, 2].

58. Wang, D.Y. *et al.* (2011) Identification and secondary structure analysis of a region affecting electrophoretic mobility of the STR locus SE33. *Forensic Sci. Int. Genet.* **6**, 310–6.

59. Butler, J.M. *et al.* (2011) SE33 variant alleles: Sequences and implications. *Forensic Sci. Int. Genet. Suppl. Ser.* **3**, e502–3.

60. Mulero, J.J. *et al.* (2006) Development and validation of the AmpFlSTR Yfiler PCR amplification kit: A male specific, single amplification 17 Y-STR multiplex system. *J. Forensic Sci.* **51**, 64–75.

61. Davis, C. *et al.* (2012) Prototype PowerPlex® Y23 System: A concordance study. *Forensic Sci. Int. Genet.* [Epub ahead of print].

62. Decker, A.E. *et al.* (2007) The impact of additional Y-STR loci on resolving common haplotypes and closely related individuals. *Forensic Sci. Int. Genet.* **1**, 215–7.

63. Chang, Y.M. *et al*. (2007) A distinct Y-STR haplotype for Amelogenin negative males characterized by a large Y(p)11.2 (DYS458-MSY1-AMEL-Y) deletion. *Forensic Sci. Int.* **166**, 115–20.

64. Takayama, T. *et al*. (2009) Determination of deleted regions from Yp11.2 of an amelogenin negative male. *Leg Med (Tokyo)* **11**, S578–80.

65. Ge, J., Eisenberg, A. and Budowle, B. (2012) Developing criteria and data to determine best options for expanding the core CODIS loci. *Investig. Genet.* **3**, 1.

66. Kline, M.C. *et al*. (2011) The latest and greatest NIST PCR-based DNA profiling standard: Updates and status of Standard Reference Material (SRM) 2391c. *Profiles in DNA* (www.promega.com/resources /articles/profiles-in-dna/2011/the-latest-and-greatest-nist-pcr-based-dna-profiling-standard/).

HOW TO CITE THIS ARTICLE

Scientific Style and Format, 7th edition, 2006

Butler, J.M., Hill, C.R. and Coble, M.D. Variability of New STR Loci and Kits in US Population Groups. [Internet] 2012. [cited: year, month, date]. Available from: http://www.promega.com/resources/articles/profiles-in-dna/2012/variability-of-new-str-loci-and-kits-in-us-population-groups/

American Medical Association, Manual of Style, 10th edition, 2007

Butler, J.M., Hill, C.R. and Coble, M.D. Variability of New STR Loci and Kits in US Population Groups. Promega Corporation Web site. http://www.promega.com/resources/articles/profiles-in-dna/2012/variability-of-new-str-loci-and-kits-in-us-population-groups/ Updated 2012. Accessed Month Day, Year.

Contribution of an article to Profiles in DNA does not constitute an endorsement of Promega products.

PowerPlex is a registered trademark of Promega Corporation.

AmpF/STR, COfiler, Identifiler and SGM Plus are registered trademarks of Applera Corporation. Excel is a registered trademark of Microsoft Corporation. GlobalFiler, MiniFiler, NGM, NGM SElect, Profiler Plus, SEfiler Plus and Sinofiler are trademarks of Life Technologies, Inc. Applied Biosystems and Yfiler are registered trademarks of Applied Biosystems.