

Email: michael.coble@nist.gov  
Phone: 301-975-4330

Michael D. Coble, Carolyn R. (Becky) Hill, and John M. Butler

National Institute of Standards and Technology (NIST), 100 Bureau Drive, Gaithersburg, MD 20899-8314

Poster available for download from STRBase:  
[http://www.cstl.nist.gov/strbase/pub\\_pres/Coble-ISHI2012-PPY23.pdf](http://www.cstl.nist.gov/strbase/pub_pres/Coble-ISHI2012-PPY23.pdf)

Chromosome Y-STR testing has found a useful niche in population genetics studies, historical investigations and forensic applications. For forensic investigations, Y-STRs can be especially useful for DNA mixture samples where low level male DNA is present in a high female DNA background. Large databases from the 17 markers of the Y-File STR kit (1) are available for the forensic community and provide relatively high discrimination among unrelated individuals with a discrimination capacity greater than 97% (2).

One limitation of Y-STR testing is the haploid nature of the Y-chromosome: the evidence and the accused is evaluated in terms of how frequently the haplotype is observed in a relevant database. For common Y-STR haplotypes, additional markers can be useful for additional discrimination (3).

The PowerPlex Y23 (PP Y23) kit was recently introduced to the forensic community and provides an additional 6 markers (DYS576, DYS481, DYS549, DYS533, DYS570, and DYS643) in addition to the 17 markers found in Y-File. We have evaluated the PP Y23 using a set of over 600 NIST population samples representing three US groups: Caucasian, African American, and Hispanics. In addition, nearly 400 Father-Son samples (4) among U.S. Caucasian, African American, Asian and Hispanics were also evaluated with PP Y23 to characterize mutation rates. Finally, a concordance check was performed on all of the samples which were previously typed with Y-File. Here we present a summary of the kit performance, population genetic parameters, and mutation rates for the new markers in the PP Y23 kit.

**References:**

- Mulero JJ, Chang CW, Calandro LM, Green RL, Li Y, Johnson CL, Hennessy LK. (2006) Development and validation of the AmpFISTR® Yfiler PCR amplification kit: a male specific, single amplification 17 Y-STR multiplex system. *J Forensic Sci.* 51(1):64-75.
- Butler JM, Hill CR, Decker AE, Kline MC, Reid TM, Vallone PM (2007) New autosomal and Y-chromosome STR loci: characterization and potential uses. *Proceedings of the Eighteenth International Symposium on Human Identification*. Available at <http://www.promega.com/geneticidproc/>
- Decker AE, Kline MC, Vallone PM, Butler JM (2007) The impact of additional Y-STR loci on resolving common haplotypes and closely related individuals. *Forensic Sci Int Genet.* 1:215-217.
- Decker AE, Kline MC, Redman JW, Reid TM, Butler JM (2008) Analysis of mutations in father-son pairs with 17 Y-STR loci. *Forensic Sci Int Genet.* 2: 631-5.
- Butler JM, Hill CR, Coble MD (2012) Variability of new STR loci and kits in U.S. population groups. *Profiles in DNA*. (2012) Available at <http://www.promega.com/profiles/>
- Davis C, Ge J, Sprecher C, Chidambaram A, Thompson J, Ewing M, Fulmer P, Rabbach D, Stora D, Budowie B. (2012) Prototype PowerPlex® Y23 System: A concordance study. *Forensic Sci Int Genet.* (in press)
- Coble MD, Hill CR, Butler JM (2012) Haplotype Data for 23 Y-chromosome markers in four U.S. population groups. *Forensic Sci Int Genet.* (manuscript submitted).
- Sauriner JL, Irwin JA, Just RS, O'Callaghan J, Parsons TJ (2008) Mitochondrial control region sequences from a U.S. "Hispanic" population sample. *Forensic Sci Int Genet.* 2(2):e19-23.
- Ballantyne KN, Goedbloed M, Fang R, Schaap O, Lao O, Wolstein A, Choi Y, van Duijn K, Vermeulen M, Brauer S, Decore R, Poetsch M, von Wurmb-Schwark N, de Knijff P, Labuda D, Vézina H, Knoblach H, Lessig R, Rowler L, Płoski R, Dobosz T, Henke L, Henke J, Furtado MR, Kayser M. (2010) Mutability of Y-chromosomal microsatellites: rates, characteristics, molecular bases, and forensic implications. *Am J Hum Genet.* 87(3):341-353.

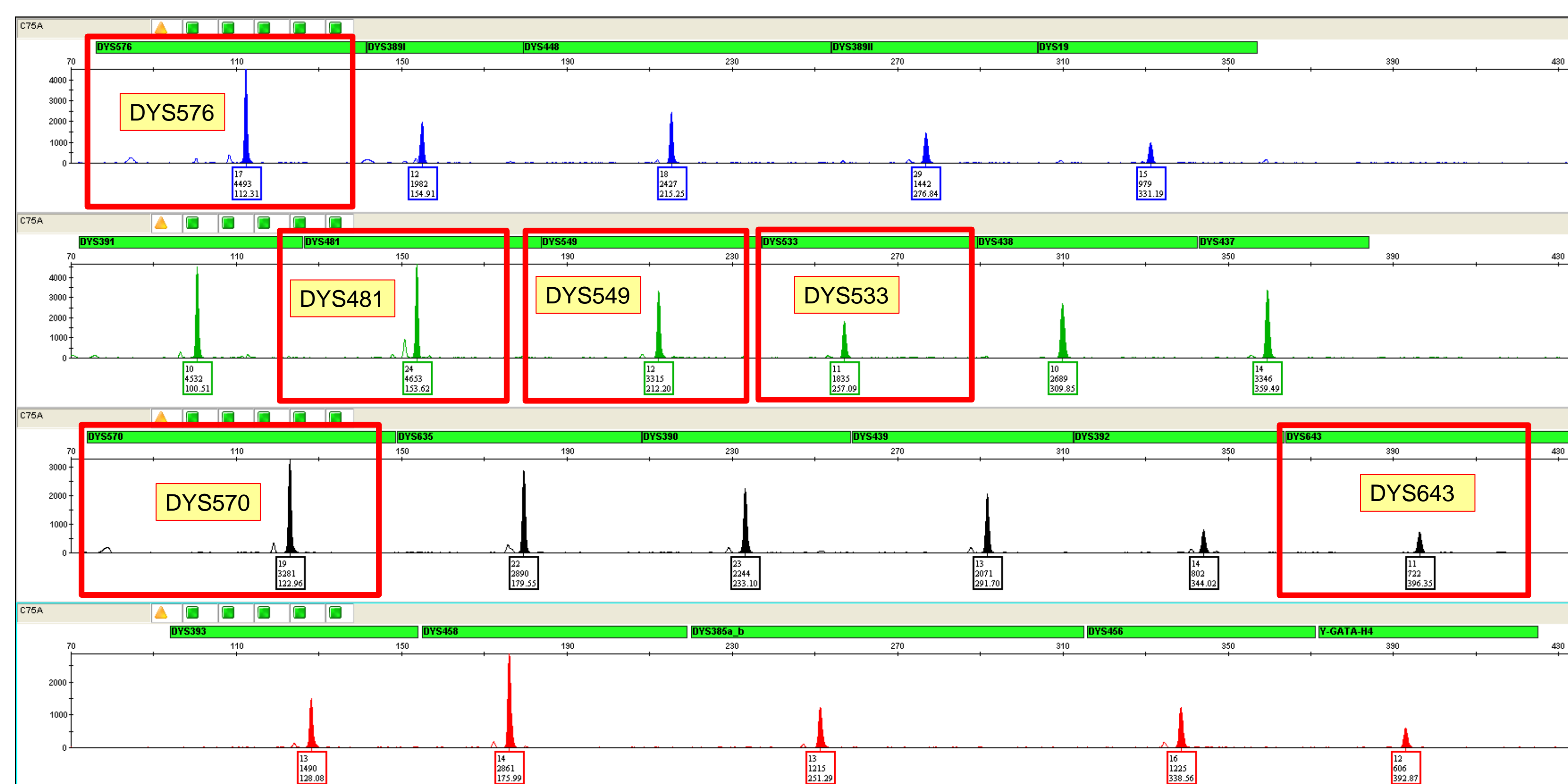
**NIST Funding:** Interagency Agreement 2010-DN-R-121 between the National Institute of Justice and NIST Office of Law Enforcement Standards.

**Disclaimer:** Points of view 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.

## NIST U.S. Population Samples

Approximately a decade ago, the NIST Human Identity Team (now the Applied Genetics Group) purchased a set of over 660 anonymous blood samples from two blood banks for characterizing forensic DNA markers. More recently, nearly 400 anonymous Father/Son samples were acquired from DNA Diagnostics Center for characterizing STR mutation rates among autosomal and Y-STR markers. A recent publication (5) has described the results of testing the 1036 NIST U.S. population samples among commonly tested forensic STR kits. Here we present results from the 1032 male samples in the NIST U.S. population data. For convention, only haplotypes from the "sons" were used for population genetics analysis. Self-described ancestries for the four population groups examined were: 359 European Caucasians, 341 African Americans, 236 Hispanics and 96 Asians

## The PowerPlex Y23 Kit (PP Y23)



An example profile from the PP Y23 kit. The kit contains all of the loci within the Y-File (Life Technologies, Foster City, CA) in addition to six new markers (in red boxes): DYS576, DYS481, DYS549, DYS533, DYS570 and DYS643.

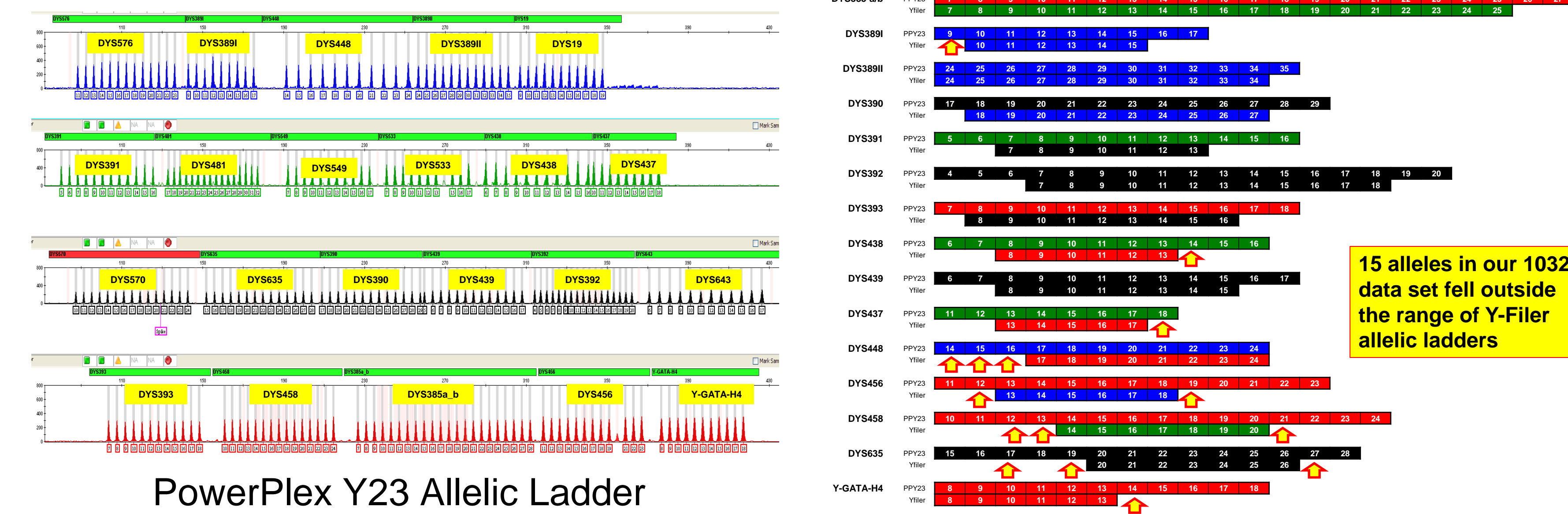
## Population Genetics

### Genetic Diversities of the markers

Locus	# Alleles	GD	African American		Caucasian		Hispanic		Asian	
			N = 341	Rank	N = 359	Rank	N = 236	Rank	N = 96	Rank
DYS385a/b	17(69)	0.929	1	0.939	1	0.850	1	0.919	1	0.944
<b>DYS481</b>	16	0.832	2	0.859	6	0.732	4	0.797	2	0.820
<b>DYS576</b>	9	0.806	3	0.807	2	0.768	3	0.798	7	0.756
<b>DYS570</b>	12	0.782	4	0.790	4	0.740	2	0.799	4	0.807
DYS458	15	0.782	6	0.758	3	0.762	5	0.786	3	0.814
DYS390	8	0.774	11	0.675	7	0.700	16	0.640	8	0.747
<b>DYS643</b>	10	0.761	5	0.786	12	0.617	13	0.671	9	0.731
DYS635	11	0.756	9	0.721	9	0.665	6	0.740	11	0.717
DYS389II	12	0.744	7	0.736	8	0.681	7	0.729	5	0.790
DYS448	16	0.733	10	0.719	13	0.605	8	0.702	10	0.721
DYS438	7	0.712	18	0.566	14	0.601	9	0.695	16	0.571
DYS19	10	0.698	8	0.724	21	0.509	14	0.669	6	0.757
<b>DYS549</b>	7	0.688	12	0.666	11	0.630	15	0.665	14	0.639
DYS456	8	0.688	14	0.641	5	0.734	10	0.693	20	0.559
DYS439	11	0.651	13	0.648	10	0.633	11	0.688	15	0.591
<b>DYS533</b>	7	0.632	15	0.620	16	0.583	19	0.614	18	0.569
DYS392	11	0.622	21	0.442	15	0.593	12	0.683	13	0.648
DYS437	8	0.617	20	0.502	18	0.566	17	0.622	21	0.362
Y-GATA-H4	6	0.601	16	0.614	17	0.567	20	0.602	17	0.571
DYS389I	8	0.576	19	0.516	20	0.540	18	0.619	12	0.670
DYS391	7	0.524	22	0.438	19	0.557	21	0.559	22	0.303
DYS393	6	0.494	17	0.579	22	0.391	22	0.442	19	0.568

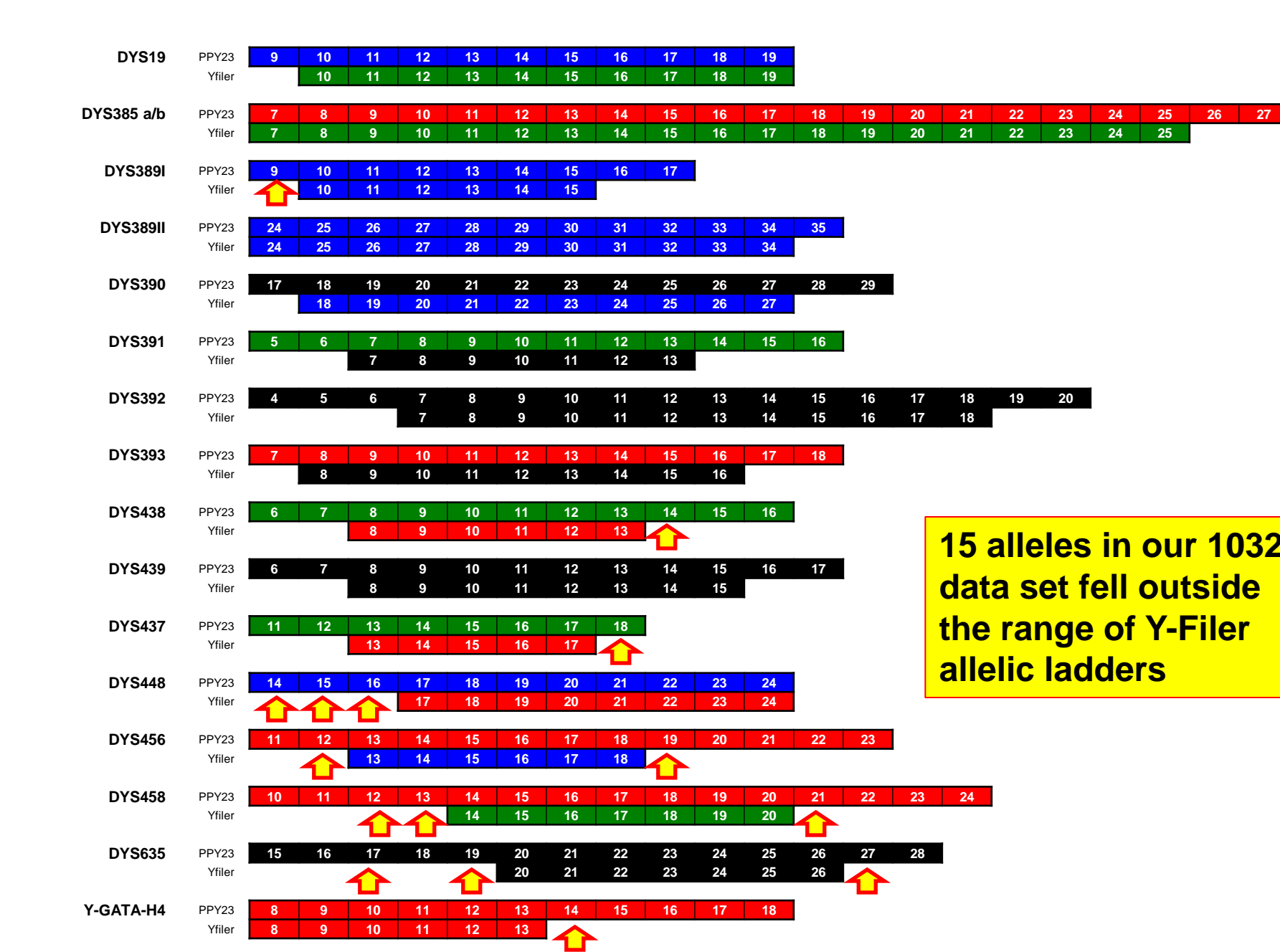
Genetic Diversities (GD) for Y-STR loci across all 1032 males and broken out among the four groups in our data set. Gene Diversity at each locus was calculated using the formula  $GD = N(1 - \sum p_i^2) / (N-1)$  where N = number of alleles at the locus and  $p_i$  is the frequency of the  $i^{th}$  allele at the locus. Loci are ranked in order of their GD values overall and then by group. The number of alleles observed at each marker is listed in the overall column. For the DYS385 marker, 17 separate alleles were observed in 69 combinations of the "a" and "b" amplicon products. The six new loci of the PP Y23 kit are underlined and in bold.

## (A) An Improved Allelic Ladder



PowerPlex Y23 Allelic Ladder

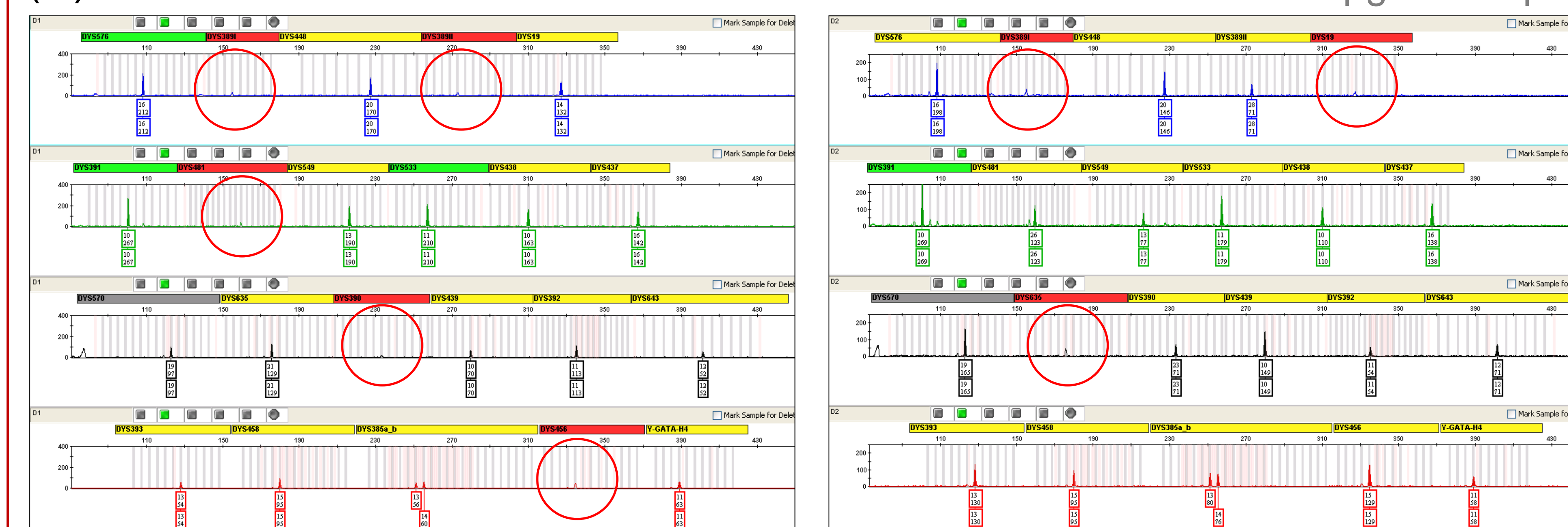
## Kit Characteristics, Performance and Applications



Comparison of Alleles in Y-STR Kit Allelic Ladders

In addition to adding six more polymorphic Y-STR loci, the PP Y23 kit includes more alleles in each allelic ladder (above, left). Across the 16 loci in common with Y-File, there are 66 additional alleles present in the PP Y23 allelic ladders (above, right). These additional alleles can be helpful in appropriately designating rare small or large Y-STR alleles. In our data set of 1032 males, we observed 15 alleles that fall outside the range of the Y-File allelic ladders alleles but within the range of the PP Y23 alleles (noted above, right with the arrows).

## (B) Robust Performance with Low-Level DNA



As part of the initial testing of the PPY 23 kit, a set of sensitivity samples were provided by Promega for quality control testing. Above are two amplification replicates at 62.5 pg/μL. Both replicates gave a high partial profile with five loci below the analytical threshold (50 RFUs, above left) and three loci below the analytical threshold (above right). Overall, we observed excellent sensitivity with the PP Y23 kit.

## (C) Increased Haplotype Discrimination

# Haplotypes	N = 1032 Combined Y-STR Haplotypes				
	Observed	MHT	SWG	PP Y12	Y-File
n=1 (unique)	671	787	821	998	1026
n=2	56	47	41	12	3
n=3	20	13	16	2	-
n=4	9	6	6	1	-
n=5	3	3	2	-	-
n=6	2	1	2	-	-
n=7	-	3	1	-	-
n=8	3	1	-	-	-
n=9	-	-	1	-	-
n=12	1	-	-	-	-
n=13	1	-	-	-	-
n=14	1	-	-	-	-
n=15	-	1	-	-	-
n=19	-	-	1	-	-
n=20	1	-	-	-	-
n=23	-	1	-	-	-
n=34	1	-	-	-	-
HD	0.99749	0.99884	0.99924	0.99995	0.99999
MP	0.00348	0.00213	0.00173	0.00101	0.00097
MP as 1 in...	288	470	579	986	1026
DC	0.74612	0.83624	0.86337	0.98159	0.99709

Number of observed haplotypes and descriptive statistics for 1032 U.S. population samples. Similar results for "unique" haplotypes among only the 6 new markers were observed by Davis *et al.* (in press, ref. 6). Additional statistical information on the NIST 1032 population samples can be found in references 5 and 7.

## Aiding the Assessment of Potential Family Relationships

During the quality control and validation of the NIST population data, the PP Y23 kit was used to identify potential close male relatives in the "unrelated" population samples. We found five pairs of samples that matched across all 23 Y-STR loci. Using autosomal STR genotypes from these samples we performed kinship analysis to determine if these pairs were potentially related. As an example, we compare samples ZT79994 and ZT79995.

Sample Name	DYS19	DYS389I	DYS389II	DYS390	DYS391	DYS392	DYS393a/b	DYS437	DYS438	DYS439	DYS448	DYS449	Y-GATA-H4	DYS533	DYS539	DYS548	DYS549	DYS553	DYS557	DYS570	DYS576	DYS643	DYS643	DYS643	DYS643	DYS643
ZT79994	14	13	29	24	10	13	13	11,15	13	12	16	15	17	23	19	12	20	16	10	12	23	12				
ZT79995	14	13	29	24	10	13	13	11,15	13	12	16	15	17	23	19	12	20	16	10	12	23	12				

Sample Name	Pos	Seq	ZT79995	Pos	Seq
ZT79994	16192	T-C	16192	T-C	
ZT79994	16311	A-G	16311	A-G	
ZT79994	16344	C-T	16344	C-T	
ZT79994	263	A-G	263	A-G	
ZT79994	315.1	C	315.1	C	
ZT79994	469	C-T	469	C-T	
ZT79994	523	A-del	523	A-del	
ZT79994	524	C-del	524	C-del	

MtDNA control region was performed at AFDIL (8) and matched exactly.

## Kinship Analysis

	Combined
Parent-Child	0
Full Sib	56327
Half-Sib*	5692
1st Cousin	264

\* is also the same stat for Aunt/Uncle/Niece/Nephew and Grandparent/Grandchild relationships

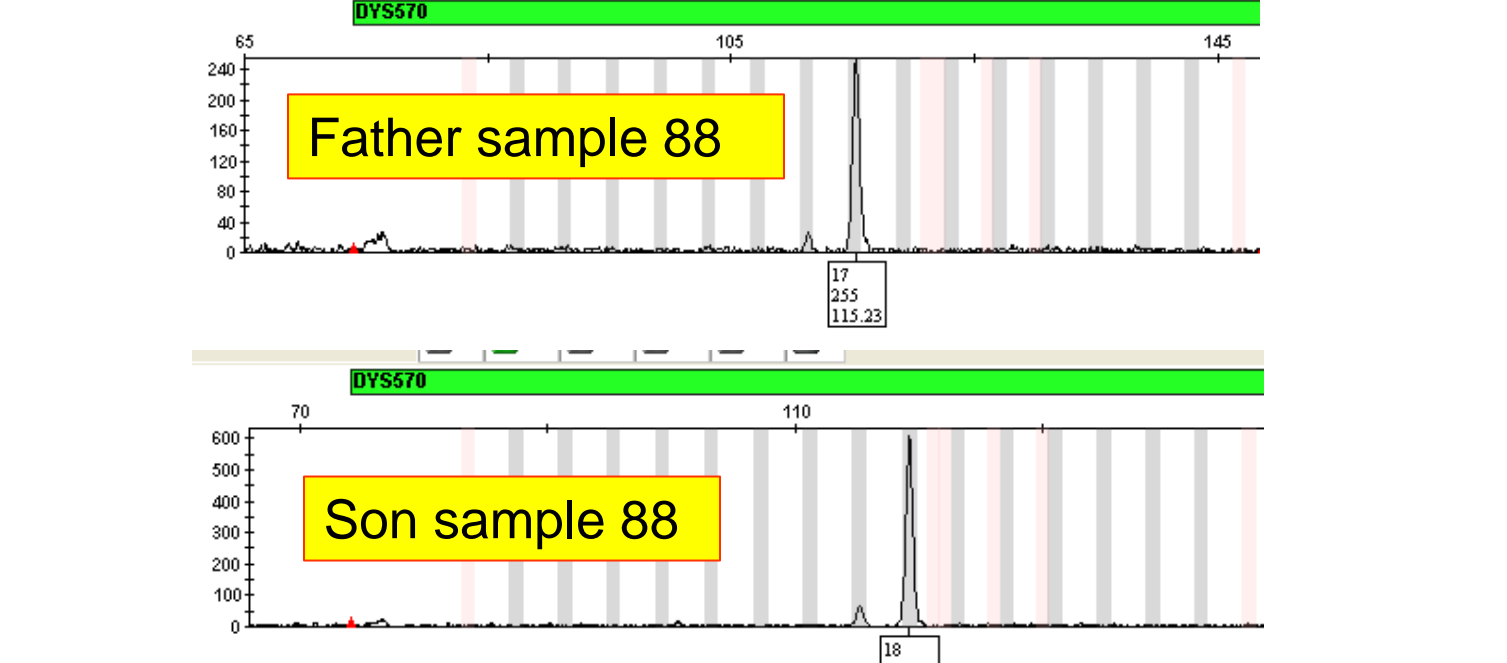
Sample Info	CSF1PO 2	CSF1PO 2	D13S317 1	D13S317 2	D16S539 1	D16S539 2	D18S51 1	D18S51 2
ZT79994	9	11	12	14	13	13	13	14
ZT79995	11	12	11	14	11	13	13	14

Autosomal STR testing of 24 markers showed at least one allele in common at all but two loci which suggests that these samples are closely related.

## Mutation Rate Analysis of Father-Son Pairs

386 Father-Son pairs were analyzed with PP Y23 to determine the mutation rates of the six new markers (right). Nearly all of the mutations observed were one-step changes (+/- one repeat) with two instances of a two-step mutation (below). The six new markers in PP Y23 are noted in red font.

	Father to Son		Father to Son	
	+1 repeat	-1 repeat	+2 repeat	-2 repeat
DYS439	2	4		
DYS389II	3	2		
<b>DYS481</b>	3	2		
<b>DYS570</b>	3	2		
<b>DYS576</b>	2	2		
DYS389I	2	2		
Y-GATA-H4	1	1	1	1
DYS458	2	1		
<b>DYS549</b>	1	1		
DYS635	1	1		
DYS19	1	0		
DYS390	0	1		
<b>DYS643</b>	1	0		
DYS385a	1	0		
DYS448	0	0		
DYS391	0	0		
<b>DYS533</b>	0	0		
DYS438	0	0		
DYS437	0	0		
DYS392	0	0		
DYS393	0	0		
DYS385b	0	0		
DYS456	0	0		
(sum)	23	23	1	1



Overall, three of the new markers ranked among the top five in the list of mutations observed. Two of the markers in the PP Y23 kit (red arrows, left) are also part of the "Rapidly Mutating" Y-STR markers recently published (9). These markers may be useful in helping to distinguish close paternal relatives.

Marker	Mutations
DYS439	6
DYS389II	5
<b>DYS481</b>	5
<b>DYS570</b>	5
<b>DYS576</b>	4
DYS389I	4
Y-GATA-H4	4
DYS458	3
<b>DYS549</b>	2
DYS635	2
DYS19	1
DYS390	1
<b>DYS643</b>	1
DYS385a	1
DYS448	0
DYS391	0
<b>DYS533</b>	0
DYS438	0
DYS437	0
DYS392	0
DYS393	0
DYS385b	0
DYS456	0

## Conclusions

- The additional 6 Y-STRs with PowerPlex Y23 improved haplotype resolution. In our NIST 1032 males, haplotype discrimination capacity went from 98.2% (Y-File) to 99.7% (PP Y23)
- Excellent sensitivity was observed with the PP Y23 kit with full profiles at 125pg input DNA.
- The expanded allelic ladder will be useful for decreasing the number of calls outside of the allele range.
- More information can be found in references 5 and 7.
- Haplotype data will be released to YHRD, the US-YSTR database, and STRBase upon final publication of (7).