

Sudhir K. Sinha,<sup>1</sup> Ph.D.; Bruce Budowle,<sup>2</sup> Ph.D.; Ranajit Chakraborty,<sup>3</sup> Ph.D.; Ana Paunovic,<sup>1</sup> B.S.; Robin DeVille Guidry,<sup>1</sup> B.S.; Chris Larsen,<sup>1</sup> M.S.; Amrita Lal,<sup>1</sup> M.S.F.S.; Megan Schaffer,<sup>1</sup> Ph.D.; Gina Pineda,<sup>1</sup> M.S.; Siddhartha K. Sinha,<sup>1</sup> B.S.; Elaine Schneida,<sup>1</sup> B.S.; Huma Nasir,<sup>1</sup> B.S.; and Jaiprakash G. Shewale,<sup>1</sup> Ph.D.

## Utility of the Y-STR Typing Systems Y-PLEX™ 6 and Y-PLEX™ 5 in Forensic Casework and 11 Y-STR Haplotype Database for Three Major Population Groups in the United States\*

**ABSTRACT:** The Y-PLEX™ 6 and Y-PLEX™ 5 systems enable analysis for 11 Y-STR loci. We present here the utility of these systems in forensic casework. A total of 188 samples, including 127 evidence samples, were analyzed using either or both of the systems. The evidence sample types included fingernail scrapings, sperm or seminal fluid, epithelial cells, blood and other tissues. The Y-STR typing systems provided useful probative results in difficult cases. A reference database for Caucasian ( $n = 517$ ), African American ( $n = 535$ ), and Hispanic ( $n = 245$ ) population groups within the United States was generated for estimating the haplotype frequency in forensic casework. Among the individuals profiled, 311 Caucasians, 412 African Americans, and 194 Hispanics provided unique profiles in their respective population datasets. This is the first report describing the haplotype database for the set of 11 Y-STR loci recommended by the Scientific Working Group on DNA Analysis Methods (SWGDM). Linkage analysis reveals that the frequencies from forensically important autosomal loci can be multiplied with the Y-STR haplotype frequency. The results from Y-PLEX™ systems have been accepted in courts in the United States.

**KEYWORDS:** forensic science, Y-chromosome, short tandem repeats, DNA typing, human identification, Y-STR, Y-PLEX, DYS393, DYS19, DYS390, DYS391, DYS385, DYS389I, DYS389II, DYS439, DYS438, DYS392, D2S1388, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D19S433, D21S11, Amelogenin, CSF1PO, DQA1, FGA, GC, GYPA, HBGG, LDLR, Penta D, Penta E, TH01, TPOX, vWA

Analysis for short tandem repeat loci (STRs) has become a preferred method in forensic DNA analysis because of their high discrimination power, good sensitivity (0.2 to 0.3 ng), ability to multiplex, ease of interpretation, and availability of databases (1–4). High sensitivity and advancements in PCR have made possible the routine analysis of samples containing low quantities of DNA, such as stains from body fluids, cigarette butts, and ancient or highly degraded samples. Analysis for samples containing mixtures of male and female DNA using autosomal STRs at times raises challenges; for example, evidence samples from sexual assault cases may contain large quantities of female DNA compared with male DNA such that the profile from female DNA masks the male DNA profile or there is just too little male DNA to be detected. The Y-chromosome short tandem repeat loci (Y-STRs) find application in resolving such cases. In recent years, Y-STRs have become useful markers in forensic DNA analysis (5–9).

Validation studies for the two Y-STR genotyping systems, Y-PLEX™ 6 and Y-PLEX™ 5, for forensic casework have been

published (10,11). Y-PLEX™ 6 enables amplification of seven loci, namely, DYS393, DYS19, DYS389II, DYS390, DYS391, and DYS385(a/b); and amplification of five loci, namely, DYS389I, DYS389II, DYS439, DYS438, and DYS392, is achieved with Y-PLEX™ 5. The DYS389II locus is common to both systems. In the present paper, utility of these systems in the forensic casework is demonstrated. A database for Caucasian, African American, and Hispanic population groups within the United States was also generated for use in the forensic community.

### Materials and Methods

AmpliQ Gold™, performance optimized polymer POP 4, matrix standards, GS500ROX, AmpFℓSTR® SGM Plus™, Profiler Plus™ and Cofiler™ amplification kits, formamide and other supplies for use of the 310 Genetic Analyzer and 377 DNA Sequencer were obtained from Applied Biosystems (Foster City, CA). The PowerPlex™ 16 kit was obtained from Promega Corporation (Madison, WI). TBE buffer (100×) was obtained from Life Technologies (Rockville, MD). Long Ranger® gel packs were from BioWhittaker Molecular Applications ApS (Denmark). All other chemicals used in this study were of analytical grade. The samples for the database studies were obtained from unrelated males belonging to different population groups residing in different geographical locations within the United States. The samples were anonymized prior to analysis.

<sup>1</sup> ReliaGene Technologies, Inc., 5525 Mounes St., Suite 101, New Orleans, LA 70123.

<sup>2</sup> FBI Laboratory, 2501 Investigation Parkway, Quantico, VA 22135.

<sup>3</sup> Center for Genome Information, Department of Environmental Health, University of Cincinnati, 3223 Eden Ave., Room 110 Kettering Lab., Cincinnati, OH.

\* Part of this work was presented at the Annual Meeting of AAFS, 2003.

Received 4 Aug. 2003; and in revised form 4 Feb. 2004; accepted 5 Feb. 2004; published 26 May 2004.

### *Extraction and Quantitation of DNA*

The DNA from reference and anonymous donor samples was extracted either by phenol-chloroform (12), Chelex<sup>®</sup> (12) or QIAamp<sup>®</sup> MiniKit (Qiagen, Valencia, CA) procedures. While using the QIAamp<sup>®</sup> MiniKit, the elution of DNA was performed using 10 mM Tris HCl buffer, pH 8.3 containing 5 mM KCl, and 0.1 mM EDTA. This buffer was used as a replacement of the AE-buffer (elution buffer, 10 mM Tris.Cl, 0.5 mM EDTA, pH 9.0) provided with the kit since AE-buffer contains 0.5 mM EDTA and such a high amount of EDTA may inhibit the PCR. The DNA from evidence samples was extracted by the phenol-chloroform method followed by concentration using Centricon 100 in 10 mM Tris HCl buffer, pH 8.3 containing 5 mM KCl, and 0.1 mM EDTA. The quantity of human DNA in casework samples was estimated by slot blot hybridization using the Quantiblot kit (Applied Biosystems, Foster City, CA). The amount of DNA in the extract for database samples was estimated by measuring the absorbance at 260 nm on a Gene Spec III (MiraiBio, Inc., Alameda, CA).

### *Amplification and Fragment Analysis*

The PCRs using Y-PLEX<sup>™</sup> 6 and Y-PLEX<sup>™</sup> 5 amplification kits (ReliaGene Technologies, Inc., New Orleans, LA) were performed as described previously (10,11). Amplifications using AmpF<sub>ℓ</sub>STR<sup>®</sup> SGM Plus<sup>™</sup>, Profiler Plus<sup>™</sup>, COfiler<sup>™</sup>, and PowerPlex<sup>™</sup> 16 amplification kits were performed as recommended by the manufacturers (3,4). A total of 0.5–2.0 ng of DNA template (unless otherwise stated) was used for amplification. While performing the amplification for Y-STRs, the general amount of male DNA in the evidence sample containing large quantities of female DNA was judged by the ratio of X and Y allele peaks at the Amelogenin locus (results from autosomal loci). The amount of male DNA derived from the ratio of X and Y allele peaks at the Amelogenin locus was essentially an empirical estimate and does not provide accurate quantity. Therefore, 0.5–2.0 ng of DNA template contained about 0.2 to 0.5 ng of male DNA. Amplification reactions were performed in a GeneAmp<sup>®</sup> PCR system 9700 (Applied Biosystems, Foster City, CA).

The amplified products were separated by electrophoresis on either the 310 Genetic Analyzer or 377 DNA Sequencer as recommended by the manufacturers. Electrophoresis on the 310 Genetic Analyzer was performed using Performance Optimized Polymer 4 (POP 4) and the conditions were 5 s injection, 60°C, 7–8 μA, 15 kV and run time for 26 min. The electrophoresis conditions for 377 DNA Sequencer were 51°C, 25–30 mA, 3 kV and run time for 2.5 h.

Genotyper<sup>®</sup> software (Applied Biosystems, Foster City, CA) was used for DNA typing.

### *Statistical Analyses*

Allele frequencies for each marker were determined by the gene-count method (13). Haplotype diversity was calculated according to the method of Tajima (14) and haplotype random match probability was calculated according to the method of Stoneking et al. (15). The DNATYPE software was developed and kindly provided by Ranajit Chakraborty (University of Cincinnati, Cincinnati, OH).

## **Results and Discussion**

The Scientific Working Group on DNA Analysis Methods (SWGDM) has recommended a set of 11 Y-STR loci for forensic

analysis and database studies in the United States (16). These Y-STR loci are DYS19, DYS385a/b, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS439, and DYS438. Of these, DYS385a/b is a duplicated locus and provides a two-allele profile (17,18). Y-STR genotyping systems Y-PLEX<sup>™</sup> 6 and Y-PLEX<sup>™</sup> 5 together provide analysis for these 11 Y-STR loci. Typing of Y-STRs can offer certain advantages over the use of autosomal STRs for obtaining a male profile in the presence of a much larger quantity of female DNA, for bypassing the differential extraction of sperm and epithelial cells, analyzing of azoospermic semen samples, estimating the number of male contributors in a rape case with multiple perpetrators, rapidly excluding suspects, and simplifying interpretation due to single allele profile per individual for most loci (19). However, Y-STR analysis has limitations also, which include a limitation that the product rule cannot be applied across the loci since the larger part of the male-specific region on the Y-chromosome does not undergo recombination and the profiles among paternal relatives cannot be distinguished. This leads to a lower power of discrimination and increased population substructure than are observed for autosomal loci (16,19). In spite of the above limitations, Y-STRs have proven useful in resolving a number of forensic cases. Corach et al. (6) analyzed over 350 cases of rape, corpse identification, and homicide investigations using Y-STRs. Prinz et al. (8) observed a full or partial profile of the male contributor in 81% of the evidence samples analyzed with the Y-STR system YM1; the study involved more than 500 cases. In addition, 56 casework samples were analyzed in parallel for autosomal STRs and Y-STRs (8). Partial detection of nonvictim alleles were observed in four cases (4%) using autosomal STRs and in nine cases (16%) using Y-STRs. Similarly, predominant or clean male type was obtained in 19 cases (34%) using autosomal STRs and in 29 cases (52%) using Y-STRs. Thus, Y-STRs provide higher success rate. Sibille et al. (20) analyzed by Y-STRs 104 swabs, wherein spermatozoa were not detected by cytology, from 79 alleged sexually assaulted female victims. The results from 28.8% of the swabs were positive for the Y-STRs.

### *Forensic Casework Using Y-PLEX<sup>™</sup> Systems*

In the past three years, a total of 188 forensic samples, including 127 evidence samples, were analyzed using the Y-PLEX<sup>™</sup> 6 or Y-PLEX<sup>™</sup> 5 systems at ReliaGene. Distribution of the samples analyzed for Y-STRs was epithelial cells (32%); azoospermic seminal fluid, sweat, or saliva (26%); sperm (26%); fingernails (15%), blood (15%), and other tissues (12%). Analysis for Y-STRs was useful in assisting in the resolution of some difficult cases. The results of using the Y-PLEX<sup>™</sup> 6 and Y-PLEX<sup>™</sup> 5 kits have been accepted in several jurisdictions throughout the United States (Table 1). The seven cases discussed below describe scenarios where analysis of Y-STRs provided useful probative results that could not be achieved with analysis of autosomal STRs. Some of these cases have been presented and discussed briefly elsewhere (21,22).

### *Forensic Paternity Cases (Cases 1 and 2)*

Case 1 was a criminal paternity involving a mother, alleged father, and product of conception. The autosomal STR results (Table 2) from the product of conception exhibited a mixture of DNA in which the mother's DNA was the major component and a minor component was detected at four autosomal loci including weak Y allele peak at the Amelogenin locus indicating that the fetus was male. The fetus's profile was dropping out using Profiler Plus<sup>™</sup> and

TABLE 1—Y-STR cases using the Y-PLEX™ 6 and Y-PLEX™ 5 kits that have been accepted in U.S. courts.

Case	Date	Jurisdiction	Docket No.	Notes
State of LA vs. Samuel Williams	10/23/01	Orleans Parish	416-355	Criminal paternity case Sexual assault case-also had other STRs, Y-STR produced no result
State of MS vs. Leon Felder	6/26/01	Pike County	00-557-KA	
State of GA vs. Ali R. Shabazz	7/31/02	Dekalb County	01-CR-4002	Sexual assault case Sexual assault case Daubert Hearing
United States vs. Spc. Michael Kelly	10/16/02	Ft. Knox	...	
State of OH vs. Chuckie Unsworth	4/16/03	Lucas County	G-4801-CR-200301510	

TABLE 2—Examples of cases wherein weak alleles were detected in the evidence samples for the autosomal STR markers (cases 1, 3, and 4)†.

Locus	Case 1			Case 3			Case 4	
	Mother	Alleged Father	Product of Conception	Victim	Suspect	Blood Stain on Bedsheet	Suspect	Blood Stain on Cloth
AUTOSOMAL LOCI								
D3S1358	16	14, 15	INC	15, 16	<12, 18	<12, 15, 18, (16) <sup>W</sup> , (14, 17) <sup>*</sup>	...	...
FGA	18.2, 23	19, 25	18.2, 23, (17.2, 22) <sup>*</sup>	23, 30.2	19, 23	23, (26, 30.2) <sup>W</sup> , (22) <sup>*</sup>	...	...
AMEL	X	X, Y	X, (Y) <sup>W</sup>	X	X, Y	X, (Y) <sup>W</sup>	...	...
D8S1179	13, 14	12, 13	13, 14 (12) <sup>*</sup>	14	12, 16	12, 14, 16, (13, 15) <sup>*</sup>	...	...
D21S11	28, 29	28, 31	28, 29, (27) <sup>*</sup>	28	32.2, 33.2	28, (30.2, 32.2) <sup>W</sup> , (27) <sup>*</sup>	...	...
D18S51	18	14, 15	18, (17) <sup>*</sup>	15, 16	12, 14	12, 16, (14.2) <sup>W</sup> , (15) <sup>*</sup>	...	...
D5S818	11, 13	12, 14	11, 13, (10, 12) <sup>*</sup>	13	11, 12	13, (11, 12) <sup>W</sup>	...	...
D13S317	11, 12	11, 13	11,12, (13) <sup>W</sup>	...	...	...	...	...
D16S539	10, 12	8, 12	10, 12, (8) <sup>W</sup> , (9, 11) <sup>*</sup>	...	...	...	...	...
TPOX	8, 11	8, 9	8, 11, (9) <sup>W</sup> , (10) <sup>*</sup>	...	...	...	...	...
D7S820	10, 11	10, 11	10, 11, (9) <sup>*</sup>	11	9, 11	11, (8) <sup>W</sup>	...	...
Y-STR LOCI (Y-PLEX™ 6)								
DYS393	NR	15	15	NR	13	13	15	15
DYS19	NR	17	17	NR	15	INC	16	16
DYS389II	NR	29	29	NR	34	28	33	33
DYS390	NR	23	23	NR	21	22	23	23
DYS391	NR	10	10	NR	10	10	10	10
DYS385	NR	17	INC	NR	16, 17	12, 14	15	15

W = Weak allele.

\* = Stutter allele.

AMEL = Amelogenin.

NR = No results; as expected female samples did not produce a profile for Y-STRs.

... = Not tested.

INC = Inconclusive.

†Modified from Ref 22.

COfiler™ systems. Little determination could be made as to the inclusion or exclusion of the suspect/alleged father. The Y-PLEX™ 6 kit was used to determine if the tested man could be excluded as the biological father of the product of conception (Table 2). Results were obtained for five of the seven loci tested. The haplotype for the product of conception for the five Y-STR loci was consistent with the reference bloodstain sample of suspect/alleged father. Therefore, suspect/alleged father or a paternal lineage relative of the alleged father could not be excluded as the biological father of the product of conception. The occurrence of the haplotype for the five loci in African American population was 0 in 1169. Utilizing the counting method the haplotype frequency will be 1/1170, which is equal to 0.0008547. The paternity index (PI) was calculated using the formula  $PI = X/Y$  where X is the probability that a man with the phenotype of the tested man could produce one sperm which carries the obligatory paternal gene and Y is the probability that some other man could produce the sperm that carries the obligatory paternal gene. X is the gamete frequency, which is 1.0 for Y-chromosome markers. Y is the haplotype frequency of the relevant population. The PI in the present case will be  $X/Y = 1/0.0008547 =$

1170. The probability of paternity of the suspect was 99.91% considering 0.5 prior probability. The probability of the paternity was calculated using the formula  $W = (Pr) (CPI) / [(Pr) (CPI) + (1 - Pr)]$  where W is probability of paternity, Pr is prior probability, and CPI is combined paternity index (www.dna-view.com; *DNA—View 2003 User's Manual*). The paternity case report for such cases should include a statement that the suspect is not excluded for the tested loci and other individuals sharing the same paternal lineage of the suspect cannot be excluded for the tested loci.

In another paternity case (case 2), the alleged father could not be excluded as the biological father except at the autosomal locus D13S317 (Table 3). However, the difference may be due to a mutation and not nonpaternity. The mother's sample was not available for profiling. A paternity index of 29 was calculated using all STR loci amplified with Profiler Plus™, COfiler™, PowerPlex™ 16, and SGM Plus™. Since the child was a male, further analysis was carried out using the Y-PLEX™ 6 and Y-PLEX™ 5 systems (Table 3). The haplotype for 11 Y-STR loci was considered as a single locus and its value was combined with that of the autosomal loci. Using Y-STRs, the combined paternity index increased

TABLE 3—Utility of Y-PLEX™ 6 and Y-PLEX™ 5 systems in a motherless paternity case with lower paternity index (case 2).

Genetic Loci	Child	Alleged Father	Paternity Index
AUTOSOMAL LOCI			
CSF1PO	11, 12	10, 12	0.829
D13S317	12	11, 13	0.0017
D16S539	11, 13	11	1.69
D18S51	19, 23	14, 19	3.11
D19S433	12, 13	12, 14	2.33
D21S11	28, 32,2	28, 31	1.15
D2S1338	21, 25	21,22	1.81
D3S1358	15, 17	15	1.71
D5S818	11	11, 14	1.9
D7S820	9, 10	9, 10	2.34
D8S1179	11, 14	14, 15	0.0746
FGA	18, 23	23, 24	1.96
TH01	7, 9,3	7	1.13
TPOX	6, 8	6, 8	3.51
vWA	14, 16	16, 20	0.921
Penta E	8, 15	13, 15	4.68
Penta D	9, 13	9, 13	4.19
Y-STR LOCI			
DYS393	13	13	
DYS19	15	15	
DYS389II	32	32	
DYS390	21	21	
DYS391	10	10	
DYS385	16	16	
DYS389I	13	13	
DYS439	12	12	
DYS438	11	11	
DYS392	11	11	536.02

to 15,705 and the probability of paternity was 99.994% (0.5 prior probability). As mentioned earlier, the paternal lineage relatives of the alleged father could also not be excluded using Y-STRs. The possibility that any other paternal lineage relative was not the father was confirmed by interview with the involved parties. Thus, Y-STR analysis could resolve the case of mutation in autosomal locus.

#### Mixture Case (Case 3)

The victim was beaten to death. The police suspected her live-in boyfriend was the assailant. Bloodstains were present on several areas of the bedsheet. Analysis with the Profiler Plus™ kit resulted in an uninterpretable mixture profile. The bloodstain sample from the bedsheet was not a victim's intimate sample (e.g., vaginal swab) and hence the assumption of victim's contribution to evidence sample was not made. Using Y-STRs, a profile was obtained for the male contributor (Table 2). The boyfriend was excluded at four loci.

#### DNA from Perspiration (Case 4)

The victim was stabbed to death and was found in her home with her arms and legs bound. Her blood was found throughout the house. Police theorized the suspect had wiped his hands on a white cloth that was slightly bloodstained. This cloth was found near cleaning solutions. Analysis for autosomal loci was not performed since the evidence samples were soaked with the victim's blood. Y-STR testing was performed on this cloth and on the ligatures that were used to bind the victim and cuttings from the ends of a blood-soaked bra. The areas used for testing on the ligature/binding were the areas

TABLE 4—Additional discrimination provided by Y- PLEX™ 5 in case 5\*.

Genetic Loci	Suspect 1	Suspect 2	Fingernail Scrapings from Female Victim
Y- PLEX™ 6			
DYS393	12	12	12
DYS19	14	14	14
DYS389II	29	29	29
DYS390	24	24	24
DYS391	11	11	11
DYS385	11, 14	11, 14	11, 14
Y- PLEX™ 5			
DYS389I	13	13	13
DYS439	11	12	12
DYS438	12	12	12
DYS392	13	14	14

\* Reproduced with permission from Ref 22.

that the assailant would use to tighten the ligature/binding. A complete profile was obtained from the cloth, which was consistent with the profile from the suspect (Table 2). The cuttings near the largest knot on the ligature yielded allele 15 for DYS393 and allele 23 for DYS390 loci. The cuttings from both ends of the blood-soaked bra yielded allele 23 for DYS390 and allele 15 for DYS385 loci. Though complete profiles were not obtained from these samples, the allele calls matched the bloodstain from the cloth for the respective loci.

#### Fingernail Scrapings (Case 5)

The victim in this case was stabbed to death. Her body was found at the bottom of a stairwell. The fingernail scrapings from the victim were tested to determine if a male profile could be obtained. The reference samples from the two suspects and fingernail scrapings from victim were analyzed using Y-PLEX™ 6. Both suspects and the fingernail cuttings had an identical profile for seven loci. The police stated that the two suspects were not related. Additional testing using Y-PLEX™ 5 kit was then performed on all three samples. The profile for suspect #2 was consistent with that of the fingernail cuttings at all 11 Y-STR loci tested, and suspect #1 was excluded at the DYS439 and DYS392 loci (Table 4). Additional analysis using the Y-PLEX™ 5 system provided the discrimination between the two suspects.

#### Semen Positive, but No Sperm Cells (Cases 6 and 7)

In Case 6, the evidence was a semen stain found on a bathrobe. Profiler Plus™ results were consistent with that of the female victim, and there was no indication of a male contributor since the profile did not exhibit a Y-allele peak at the Amelogenin locus (Fig. 1). However, the evidence sample was positive for P30. Therefore, a large amount of DNA (total approximately 100 ng) was placed in the PCR for Y-STRs. When the evidence sample was analyzed using the Y-PLEX™ 6, a conclusive male profile was obtained (Fig. 2). The Y-STR results were consistent with the profile of the male suspect. An "off-ladder" allele of 172 bases in length was observed in the evidence semen stain profile. To attempt to determine the source of the "off-ladder" peak, the victim's reference sample was analyzed as a control with a template DNA level approximating that of the evidence. The victim's reference sample exhibited two peaks of 172 and 255 bases, which are possibly products from cross-reactivity with the X-chromosome (Fig. 2). The Y-STR profile of the suspect

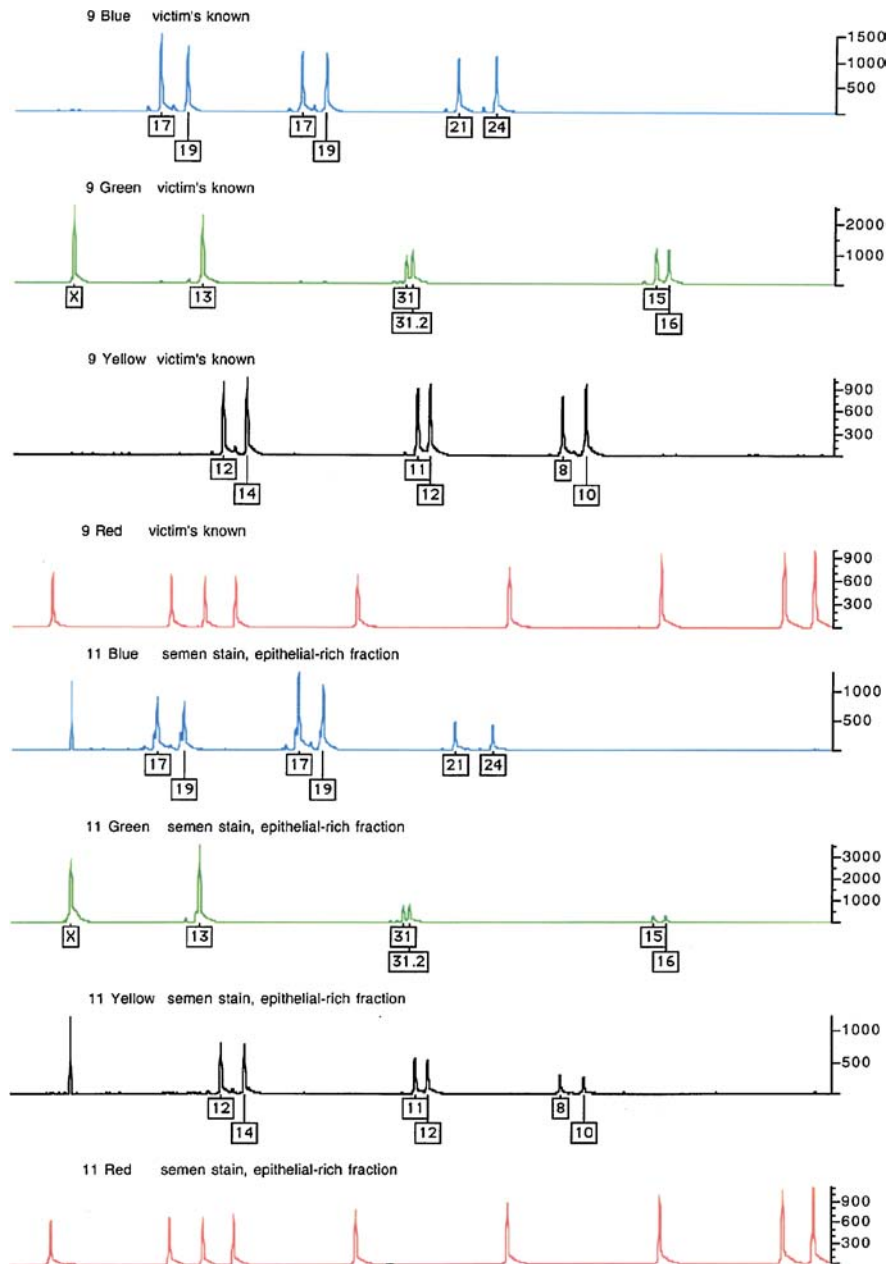


FIG. 1—Autosomal STR profile for the victim's reference sample (top) and P30 positive evidence sample (bottom) from case no. 6 using Profiler Plus™.

was interpreted as consistent with the profile of the semen stain since the source of the 172 base peak could be attributed to the victim's DNA when analyzed at such high levels of template DNA.

In another case (Case 7), the epithelial cell fraction of the victim's panties yielded no results with the PM/DQA1 loci analysis. This sample was then profiled using the Y-PLEX™ 6 and Y-PLEX™ 5 systems. The results for all 11 Y-STR loci tested were consistent with the profile of the suspect (Table 5).

#### Population Studies

A database plays an important role in estimating the rarity of a Y-STR haplotype. There are a number of available population databases for sets of Y-STR loci. These include Albanian immigrants in Italy (23), Baltic States (24), Belgium (25),

Catalonia (26), Central Spain (27), China (28), Chinese Han (29), Filipino (30), German (31), Iberian Peninsula (32), India (33,34), Japan (35,36), Pakistan (37,38), Poland (39), and Spain (40). The largest compilation of Y-STR population data has been described by Roewer et al. (41). Ten STR loci are typed in this dataset. To the best of our knowledge, our study is the first report documenting the database for a set 11 Y-STR loci recommended by the SWGDAM. Three population groups, Caucasian ( $n = 517$ ), African American ( $n = 535$ ), and Hispanic ( $n = 245$ ), were profiled for the 11 Y-STR loci using the Y-PLEX™ 6 and Y-PLEX™ 5 systems. About 90% of the samples were buccal swabs and the other 10% were whole blood. The reinjection rate was about 5%. Allele frequencies for each locus for the three population groups studied are presented in Table 6. Genetic diversity was greatest for the DYS385a/b locus in all three population groups, which is probably because this is a duplicated

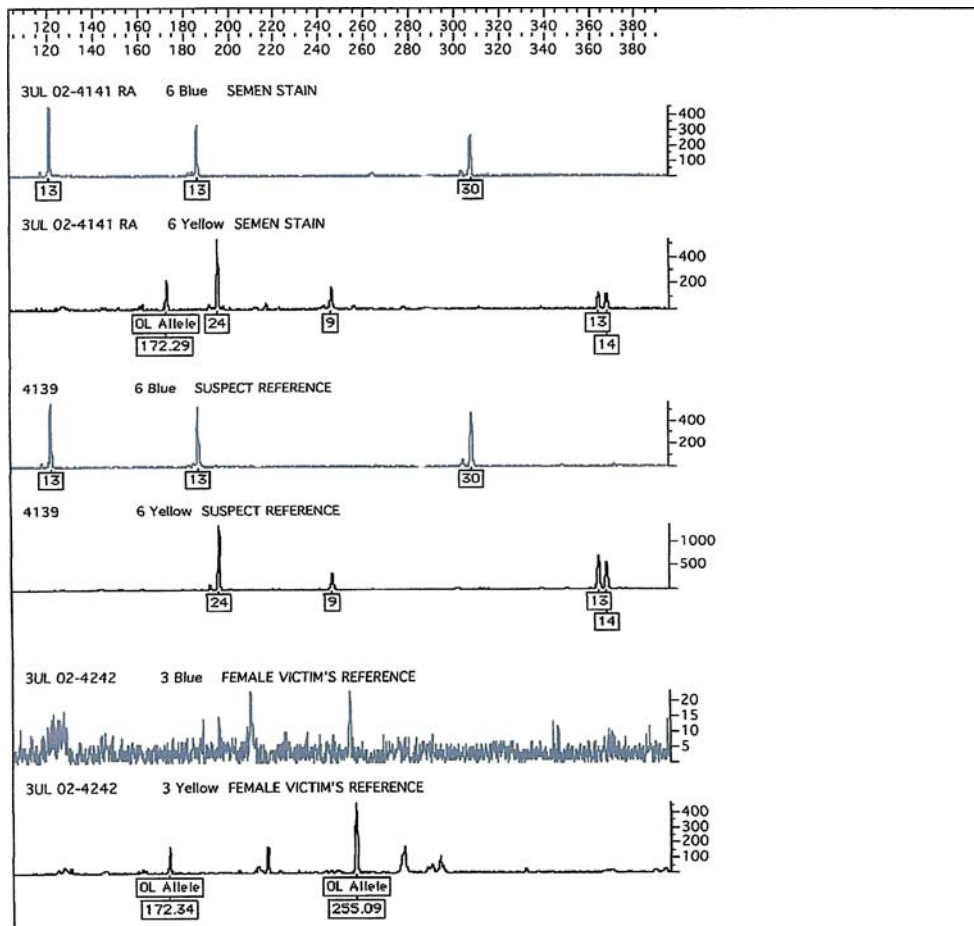


FIG. 2—Y-STR profile for the P30 positive evidence sample (top) and victim's reference sample (bottom) from case no. 6 using Y-PLEX™ 6. Reproduced with permission from Ref 22.

TABLE 5—Analysis of samples in case 7 for autosomal and Y-STR loci.

Genetic Loci	Reference Blood Stain of Victim	Reference Blood Stain of Suspect	Epithelial Cell Fraction of Victim's Panties with Semen Stain
AUTOSOMAL LOCI			
LDLR	B	B	NR
GYPA	A	AB	NR
HBGG	B	AC	NR
D7S8	A	A	NR
GC	C	B	NR
DQA1	1,2, 4,1	2, 3	NR
Y-STR LOCI			
DYS393	NR	13	13
DYS19	NR	15	15
DYS389II	NR	31	31
DYS390	NR	21	21
DYS391	NR	10	10
DYS385	NR	16, 18	16, 18
DYS389I	NR	13	13
DYS439	NR	11	11
DYS438	NR	9	9
DYS392	NR	11	11

NR = No result.

locus and provides genetic diversity values equivalent to two Y-STR loci. The haplotype diversity, and haplotype match probability values are summarized in Table 7. The haplotype diversity values

computed for the Caucasian and African American populations were 0.9946 and 0.9991, respectively. In the previous studies the haplotype diversity values for seven loci amplified with the Y-PLEX™ 6 for these population groups were 0.9921 and 0.9979, respectively (10). Thus, analysis for an additional four loci using the Y-PLEX™ 5 increased the power of discrimination (the DYS389II locus is common to two systems). The observed values for allele frequencies, haplotype diversity, and haplotype match probability values were consistent with data published for populations of similar anthropological affinity (10,11,24,37,41–44). Unique haplotypes for the 11 Y-STR loci were observed for 311 of 517 Caucasians, 412 of 535 African Americans, and 194 of 245 Hispanics. The occurrence of a Y-haplotype in more than one individual for a set of Y-STR loci has been commonly observed in sample population studies and is expected because of the lack of recombination in the Y-chromosome. For example, in a database (41) of 4688 individuals for nine Y-STR loci, 139 individuals share the same profile, and each of the next 30 most frequent haplotypes are shared by at least 14 individuals. The most frequent haplotypes for the 11 Y-STR loci investigated in the present study are presented in Table 8.

#### Linkage with the Autosomal Loci

One hundred and eleven Caucasian and 115 African American males were typed with Y-PLEX™ 6 and AmpFℓSTR® SGM Plus genotyping systems. Allele frequencies for each marker were determined by the gene-count method (13). Tests for deviations from

TABLE 6—Allele frequencies of the loci amplified with the Y-PLEX™ 6 and Y-PLEX™ 5 kits in Caucasian, African American, and Hispanic population groups.

Locus	Allele	Caucasian (n = 517)		African American (n = 535)		Hispanic (n = 245)	
		Count	%	Count	%	Count	%
DYS393	11	1	0.19	1	0.19	3	1.22
	12	45	8.70	24	4.49	33	13.47
	13	421	81.43	295	55.14	177	72.25
	14	38	7.35	151	28.22	26	10.61
	15	11	2.13	63	11.78	5	2.04
	16	1	0.19	1	0.19	1	0.41
DYS19	12	...	...	...	...	1	0.41
	13	18	3.48	12	2.24	49	20.00
	14	382	73.89	134	25.05	138	56.33
	15	80	15.47	208	38.88	39	15.92
	16	33	6.38	92	17.20	12	4.90
	17	4	0.77	89	16.64	6	2.45
DYS389II	27	6	1.61	5	0.94	2	0.82
	28	81	15.67	41	7.66	25	10.20
	29	274	53.00	116	21.68	102	41.63
	30	110	21.28	204	38.13	78	31.84
	31	39	7.54	132	24.67	30	12.25
	32	6	1.16	31	5.79	8	3.27
	33	1	0.19	6	1.12	...	...
DYS390	20	...	...	7	1.31	...	...
	21	10	1.93	318	59.44	9	3.67
	22	58	11.22	43	8.04	15	6.12
	23	136	26.31	53	9.91	60	24.49
	24	245	47.39	87	16.26	137	55.92
	25	68	13.15	27	5.05	22	8.98
	26	...	...	...	...	2	0.82
	DYS391	9	18	3.48	6	1.12	18
10		225	43.52	382	71.40	122	49.80
11		263	50.87	138	25.79	99	40.41
12		11	2.13	9	1.68	6	2.45
DYS385a/b	10	11	1.06	2	0.19	4	0.82
	11	341	32.98	124	11.59	129	26.33
	12	50	4.84	20	1.87	19	3.88
	13	105	10.16	32	2.99	50	10.20
	14	355	34.33	132	12.34	130	26.53
	15	111	10.74	135	12.62	59	12.04
	16	34	3.29	217	20.28	37	7.55
	17	12	1.16	215	20.09	23	4.69
	18	13	1.26	128	11.96	24	4.90
	19	2	0.19	65	6.08	14	2.86
	20	...	...	...	...	1	0.20
DYS389I	11	1	0.19	5	0.94	5	2.04
	12	100	19.34	7	14.39	28	11.43
	13	342	66.15	362	67.66	150	61.22
	14	71	13.73	90	16.82	61	24.90
	15	3	0.58	1	0.19	1	0.41
DYS439	10	35	6.77	7	1.31	19	7.76
	11	174	33.66	150	28.04	61	24.90
	12	243	47.00	281	52.52	125	51.02
	13	60	11.61	87	16.26	40	16.33
	14	5	0.97	10	1.87	...	...
DYS438	8	1	0.19	5	0.94	...	...
	9	16	3.10	6	1.12	23	9.39
	10	126	24.37	63	11.78	51	20.82
	11	40	7.74	342	63.93	36	14.69
	12	321	62.09	119	22.24	133	54.29
	13	13	2.52	...	...	2	0.82
DYS392	10	1	0.19	3	0.56	1	0.41
	11	145	28.05	399	74.58	74	30.20
	12	24	4.64	24	4.49	16	6.53
	13	300	58.03	91	17.01	127	51.84
	14	41	7.93	17	3.18	19	7.76
	15	6	1.16	1	0.19	8	3.27

TABLE 7—Observed haplotype diversity and haplotype match probability for 11 Y-STR loci typed with the Y-PLEX™ 6 and Y-PLEX™ 5 kits.

Population	Haplotype Diversity (h)	Haplotype Match Probability
Caucasian (n = 517)	0.9946	0.0073
African American (n = 535)	0.9991	0.0028
Hispanic (n = 245)	0.9973	0.0067

Hardy-Weinberg expectation of genotype frequencies at the autosomal loci were conducted by three test procedures (chi-square test for overall deficiency of heterozygosity at the locus, likelihood ratio test, and the exact test). The autosomal loci simultaneously analyzed were D3S1358, vWA, D16S539, D2S1338, D8S1179, D21S11, D18S51, D19S433, TH01, and FGA. For this purpose, we used the DNATYPE software, in which the levels of significance were determined by 10,000 shufflings of alleles across individuals. Independence of alleles between the Y-STRs and that between Y- and autosomal STRs was tested through contingency tables analysis, in which again the levels of significance were determined by permutation-based computer simulations. Finally, mutual independence of the ten autosomal loci was checked by computing the distribution of allele and genotype sharing between all possible pairs of individuals sampled.

Tables 9–11 show levels of significance of tests of independence of alleles between loci. The allele frequency distributions for each marker in African-Americans and European-Americans are consistent with those previously published for the same loci in populations of similar anthropological affinity.

Tests for HWE for the ten autosomal loci showed four departures (at the 5% level of significance) from allelic independence within loci for the European-American sample. These include one detected by the chi-square test at the D19S433 locus, one detected by the likelihood ratio test at the D3S1358 locus, and two detected by the exact test at the D3S1358 and D18S51 loci). In contrast, the African-American sample deviated from HWE (at 5% level of significance) only at the D18S51 locus, detected by the chi-square test. Examination of the causes of these significant departures indicated that in all cases significant deviation from HWE involved observing genotypes consisting of rare alleles (i.e., alleles that are found less than 2 or 3 counts in the entire sample). Thus, we conclude that when allele frequencies are augmented by a minimum threshold allele frequency, the observed significant departures from HWE is of no consequence for obtaining a conservative frequency estimate of single-locus genotype frequencies by employing the strict product rule, i.e., by assuming HWE (45–49).

While the allele frequencies for the autosomal STRs can be used directly for forensic casework analysis, employing the recommendations of the NRC report (45), the Y-markers should be used together in the form of haplotype frequencies. The  $r \times c$  contingency analysis tests the independence of STR alleles between the pairs of seven Y-chromosome loci. In spite of the fact that these seven loci (DYS 385 is subdivided in to *DYS 385-a* and *DYS 385-b* for statistical analysis) are tightly linked on the nonrecombining region of the Y-chromosome, only eight of the 21 pairs of loci show significant nonrandom association of alleles in the European-American sample (upper diagonal entries), while nine of the 21 test results were significant (at 5% level) for the African-American sample. The analysis indicates that, as expected, while the Y-STR loci are linked among themselves, the Y-STR loci are not linked with autosomal loci. The current analyses support the conclusion that frequencies

TABLE 8—Observed most frequent haplotypes for 11 Y-STR loci typed with the Y-PLEX<sup>TM</sup>6 and Y-PLEX<sup>TM</sup>5 kits.

Caucasian (n = 517)		African American (n = 535)		Hispanic (n = 248)	
Haplotype*	n	Haplotype*	n	Haplotype*	n
13, 14, 29, 24, 11, 11-14, 13, 12, 12, 13	22	13, 14, 29, 24, 11, 11-14, 13, 12, 12, 13	9	13, 14, 29, 24, 11, 11-14, 13, 12, 12, 13	7
13, 14, 29, 24, 10, 11-14, 13, 12, 12, 13	18	13, 14, 29, 24, 10, 11-14, 13, 12, 12, 13	5	13, 14, 29, 24, 11, 11-14, 13, 11, 12, 13	7
13, 14, 29, 23, 11, 11-14, 13, 12, 12, 13	14	15, 17, 31, 21, 10, 17-19, 14, 12, 11, 11	5	13, 13, 30, 24, 9, 13-14, 14, 10, 10, 11	7
13, 14, 29, 24, 11, 11-14, 13, 11, 12, 13	10	13, 15, 31, 21, 10, 16-17, 13, 11, 11, 11	4	13, 14, 29, 24, 11, 11-14, 13, 13, 12, 13	4
13, 14, 29, 24, 11, 11-14, 13, 13, 12, 13	7	14, 15, 30, 21, 10, 15-16, 13, 12, 11, 11	4	13, 14, 30, 24, 11, 11-14, 13, 12, 12, 13	4
13, 14, 29, 24, 11, 11-15, 13, 12, 12, 13	7				
13, 14, 30, 24, 10, 11-14, 13, 12, 12, 13	7				
13, 14, 29, 25, 11, 11-13, 13, 12, 12, 14	6				
13, 14, 28, 22, 10, 13-14, 12, 11, 10, 11	6				
13, 14, 29, 24, 10, 11-15, 13, 12, 12, 13	6				
13, 14, 29, 23, 11, 11-14, 13, 13, 12, 13	5				
13, 14, 29, 25, 11, 11-14, 13, 12, 12, 13	5				
13, 14, 30, 24, 11, 11-14, 13, 12, 12, 13	5				

\* The sequence of the loci in the haplotype is DYS393, DYS19, DYS389II, DYS390, DYS391, DYS385a/b, DYS389I, DYS432, DYS438, and DYS392.

TABLE 9—Levels of significance of tests of independence of alleles between seven Y-chromosome STR loci.

Loci	DYS393	DYS19	DYS389	DYS390	DYS391	DYS385-a	DYS385-b
DYS393	...	0.038*	0.149	0.180	0.904	0.014*	0.074
DYS19	0.012*	...	0.043*	0.086	<0.001**	0.009**	0.076
DYS389	0.597	0.505	...	0.046*	0.192	0.025*	0.119
DYS390	0.082	0.003**	0.035*	...	0.158	0.008**	0.110
DYS391	0.121	<0.001**	0.189	0.058	...	0.109	0.527
DYS385-a	0.171	<0.001**	0.068	0.005**	0.059	...	0.112
DYS385-b	0.447	0.031*	0.063	0.028*	0.150	0.028*	...

The levels of significance were determined by 10,000 replications of permutations of alleles across individuals. \* P < 0.01, \*\* P < 0.001. The upper diagonal entries are for the Caucasian sample and the lower diagonal entries are for the African-American sample.

TABLE 10—Levels of significance of tests of independence of alleles between seven Y-STRs and ten autosomal STRs in the Louisiana Caucasians.

Loci	DYS393	DYS19	DYS389	DYS390	DYS391	DYS385-a	DYS385-b
D3S1358	0.033*	0.685	0.342	0.637	0.651	0.028*	0.139
vWA	0.052	0.976	0.571	0.843	0.862	0.034*	0.857
D16S539	0.040*	0.224	0.712	0.532	0.097	0.422	0.007**
D2S1338	0.060	0.432	0.112	0.055	0.883	0.789	0.021*
D8S1179	0.041*	0.542	0.533	0.316	0.712	0.577	0.772
D21S11	0.921	0.273	0.658	0.503	0.801	0.028*	0.866
D18S51	0.611	0.187	0.784	0.674	0.715	0.908	0.437
D19S433	0.663	0.510	0.810	0.671	0.278	0.746	0.304
TH01	0.251	0.577	0.196	0.118	0.747	0.046*	0.568
FGA	0.101	0.582	0.168	0.781	0.817	0.204	0.774

The levels of significance were determined by 10,000 replications of permutations of alleles across individuals.

\* P < 0.01, \*\* P < 0.001.

TABLE 11—Levels of significance of tests of independence of alleles between seven Y-STRs and ten autosomal STRs in the Louisiana African Americans.

Loci	DYS393	DYS19	DYS389	DYS390	DYS391	DYS385-a	DYS385-b
D3S1358	0.060	0.854	0.648	0.840	0.108	0.913	0.617
vWA	0.609	0.712	0.441	0.167	0.287	0.537	0.963
D16S539	0.247	0.733	0.214	0.818	0.523	0.130	0.552
D2S1338	0.896	0.406	0.443	0.531	0.735	0.034*	0.204
D8S1179	0.557	0.522	0.019*	0.539	0.454	0.210	0.122
D21S11	0.621	0.411	0.458	0.778	0.671	0.995	0.881
D18S51	0.745	0.612	0.957	0.513	0.582	0.604	0.562
D19S433	0.280	0.158	0.797	0.936	0.414	0.082	0.462
TH01	0.493	0.744	0.995	0.266	0.234	0.978	0.529
FGA	0.136	0.168	0.266	0.162	0.073	0.246	0.044*

The levels of significance were determined by 10,000 replications of permutations of alleles across individuals.

\* P < 0.01.



32. Bosch E, Lee AC, Calafell F, Arroyo E, Henneman P, de Knijff P, et al. High resolution Y-chromosome typing: 19 STRs amplified in three multiplex reactions. *Forensic Sci Int* 2002;125:42–51.
33. Das B, Chauhan PS, Seshadri M. Y-chromosomal STR haplotypes in two population groups of Kerala in South India. *J Forensic Sci* 2002;47:690–1.
34. Das B, Chauhan PS, Seshadri M. Population database of Y-chromosomal haplotypes at five microsatellites among two distinct ethnic groups of Western India. *J Forensic Sci* 2002;47:692–3.
35. Deva K, Rand S, Hohoff C, Naitol E, Yamanouchi H, Brinkmann B. Haplotype analysis of 13 Y-STRs, YAP, SRY465 and M9 in a Japanese population. III. International Forensic Y-User Workshop; 7–9 Nov. 2002, Porto, Portugal: International Forensic Y-STR User Group.
36. Uchihi R, Yamamoto T, Usuda K, Yoshimoto T, Tanaka M, Tokunaga S, et al. Haplotype analysis with 14 Y-STR loci using 2 multiplex amplification and typing systems in 2 regional populations in Japan. *Int J Legal Medicine* 2003;117:34–8.
37. Ayub Q, Mohyuddin A, Qamar R, Mazhar K, Zeral T, Mehdi SQ, et al. [Identification and characterization of novel human Y-chromosomal microsatellites from sequence database information](#). *Nucleic Acid Res* 2000;28(2):e8.
38. Mohyuddin A, Ayub Q, Qamar R, Zerjal T, Helgason A, Mehdi SQ, et al. [Y-chromosomal STR haplotypes in Pakistani populations](#). *Forensic Sci Int* 2001;118:141–6.
39. Pawjowski R, Dettlaff-Kakol A. [Population data of nine Y-chromosomal STR loci in Northern Poland](#). *Forensic Sci Int* 2003;131:209–13.
40. Quintans B, Beza S, Brion M, Sanchez-Diz P, Lareu M, Carracedo A. [Population data of Galicia \(NW Spain\) on the new Y-STRs DYS437, DYS438, DYS439, GATA A10, GATA A7.1, GATA A 7.2, GATAC4 and GATA H4](#). *Forensic Sci Int* 2003;131:220–4.
41. Roewer L, Krawczak M, Willuweit S, Nagy M, Alves C, Amorim A, et al. On line reference database of European Y-chromosomal short tandem repeat (STR) haplotypes. *Forensic Sci Int* 2001;118:6–13.
42. Redd AJ, Clifford SL, Stoneking M. Multiplex DNA typing of short tandem repeat loci on the Y-chromosome. *Biol Chem* 1997;378:923–7.
43. Coyle HM, Budowle B, Bourke MT, Carita E, Hintz JL, Ladd C, et al. Population data for seven Y-chromosomal STR loci from three different population groups in Connecticut. *J Forensic Sci* 2003;48:435–7.
44. Garofalo MG, Gangitano DA, Juvenal GJ, Budowle B, Lorente JA, Padula RA. Six Y-chromosome frequencies in a population from Argentina. *J Forensic Sci* 2003;48:455–6.
45. National Research Council. *The evaluation of forensic DNA evidence*. Washington DC: National Academic Press, 1996.
46. Chakraborty R, Stivers DN, Su B, Zhong Y, Budowle B. [The utility of STR loci beyond human identification: implications for the development of new DNA typing system](#). *Electrophoresis* 1999;20:1682–96.
47. Budowle B, Shea B, Niezgoda S, Chakraborty R. CODIS STR loci data from 41 sample populations. *J Forensic Sci* 2001;46:453–89.
48. Budowle B, Chidambaram A, Strickland L, Beheim CW, Taft GM, Chakraborty R. [Population data at 13 STR loci for three Native Alaska population groups](#). *Forensic Sci Int* 2002;129:51–7.
49. Budowle B, Monson KL, Chakraborty R. Estimating minimum allele frequencies for DNA profile frequency estimates for PCR-based loci. *Int J Leg Med* 1996;108:173–6.

Additional information and reprint requests:  
 Sudhir K. Sinha, Ph.D.  
 ReliaGene Technologies, Inc.  
 5525 Mounes St., Suite 101  
 New Orleans, LA 70123  
 E-mail: sinha@reliagene.com