J Forensic Sci, November 2003, Vol. 48, No. 6
Paper ID JFS2003114_486
Available online at: www.astm.org

Cassie L. Johnson,[1] *M.S.; Joseph H. Warren,*[1] *B.A.; Robert C. Giles,*[1] *Ph.D.; and Rick W. Staub,*[1] *Ph.D.*

Validation and Uses of a Y-Chromosome STR 10-Plex for Forensic and Paternity Laboratories*

ABSTRACT: Y-chromosome short tandem repeats (Y-STRs) provide valuable information in cases of rape and questioned paternity, and they allow for the genetic identification of male lineages. The present study validated a Y-STR 10-plex on the ABI PRISM® 3100 Genetic Analyzer for use in forensic and paternity laboratories at Orchid Cellmark. Following optimization of the polymerase chain reaction, father-son pairs were analyzed to ensure that each pair generated identical haplotypes. This study demonstrated that the 10-plex is sensitive to 0.125 ng of input DNA and that female samples mixed with male samples did not interfere with Y-STR haplotyping. In a sample of 525 males, there were three instances of locus multiplication, two at DYS19 and one at DYS435. Overall, haplotype diversity was 0.996, suggesting that the 10-plex can effectively distinguish among male lineages.

KEYWORDS: forensic science, DNA typing, Y-chromosome, short tandem repeat, DYS436, DYS439, DYS435, DYS19, DYS460, Y-GATA-H4, DYS391, DYS392, DYS438, DYS437

In recent years, STRs located on the Y-chromosome (Y-STRs) have attracted the interest of the forensic community. These STRs may be valuable in sexual assault cases, paternity cases involving male offspring, and genealogical studies (1,2). Since the human Y-chromosome is male-specific, the use of Y-STRs in forensic evidence will allow for the genetic identification of the male component(s) of a sample. Each individual male is Y hemizygous, no recombination occurs (2,3), and all Y-STR alleles are inherited as a unit. Thus, the genetic information derived from a panel of Y-STR markers is referred to as a haplotype. With the exception of mutational events, the Y-chromosome is passed down in an unchanged form from father to son(s) (2,3).

The individual Y-STR loci do not assort independently, so the frequencies of alleles found at each locus cannot be combined into one probability statement using the product rule (2). Therefore, empirical examination of a number of Y-STR loci is necessary to provide sufficient discrimination of haplotypes among male lineages. Accurate estimations of empirically determined haplotype frequencies are limited by the sample size (4).

Y-STRs can provide useful information in the analysis of sexual assault evidence. The traditional method of extracting male DNA from sexual assault evidence involves performing a lengthy two-step differential extraction that occasionally produces a mixed autosomal STR profile (5,6). The female DNA can mask the DNA profile of the male suspect, making interpretation difficult. Therefore, the use of Y-STRs is advantageous because it allows the analyst to examine the male component of the sexual assault evidence.

Y-STR typing also may be applied to paternity or genealogical investigations involving male children in order to exclude or establish paternity or a paternal lineage. In paternity cases in which the father is unavailable for testing, one can generate a Y-STR profile from a patrilineal blood relative of the alleged father to determine if a paternal relationship could exist (6,7).

Y-STR testing may also be valuable in cases of an apparent (rare) male amelogenin deletion or mutation (4). A sequence alteration in the amelogenin gene on the Y-chromosome can cause failure of amplification of the Y-amelogenin fragment, causing a male specimen to be mistaken for female. In this situation, Y-STRs can help confirm the presence or absence of a Y-chromosome.

The present study validated a Y-STR multiplex originally developed by NIST (the National Institute of Standards and Technology). An internal validation of the Y-STR multiplex was performed in accordance with the DNA Advisory Board Quality Assurance Standard 8.1.3 (8). An internal validation verifies procedures that were established by another laboratory to ensure that they will be effective in one's own laboratory (2). The internal validation used control and reference DNA samples. The validation studies included optimization of the PCR reaction, analysis of father-son pairs, development of population databases, and mixture and sensitivity studies. The Y-STR 10-plex examines ten non-overlapping loci on the Y-chromosome. Table 1A provides a summary. The 10-plex uses forward and reverse primer sets (Table 1B) that reduce the lengths of the PCR amplicons, thereby making them more applicable in the analysis of the degraded samples encountered in forensic casework (9). The present study was undertaken to validate the Y-STR 10-plex for use in forensic and paternity cases.

Materials and Methods

DNA Samples

Control DNA samples included the male DNA extracts ATCC 45514 (American Type Culture Collection, Manassas, VA) and SRM 2395 (National Institute of Standards and Technology, Gaithersburg, MD), as well as the female extract 9947A (Promega, Madison, WI). "BS #1" refers to a DNA extract obtained from a male volunteer.

[1] Orchid Cellmark-Dallas, Dallas, TX.

* Poster presented at the 13th International Symposium on Human Identification, Phoenix, AZ, 2002.

Received 24 Mar. 2002; and in revised form 21 June 2003; accepted 10 July 2003; published 25 Sept. 2003.

TABLE 1—A. *Summary of the Y-STR 10-plex loci. Repeat numbers are those observed in the current study*. B. *Summary of the Y-STR 10-plex primers*.

(A)

Locus	Dye	Repeat Motif (9,24)	Approx. Base Pairs	Repeat Numbers
DYS436	FAM	GTT	78–92	10–14
DYS439	FAM	$[GATA]_2N_4[GATA]_3N_{14}$ $[GATA]_1N_3[GATA]_1N_7[GATA]_n$	107–136	15–22
DYS435	FAM	TGGA	142–151	8–12
DYS19	FAM	$[TAGA]_3N_4[TAGA]_n$	180–205	12–18
DYS460	VIC	ATAG	95–120	7–13
H4	VIC	TAGA	122–148	9–15
DYS391	VIC	TCTA	158–175	9–14
DYS392	NED	TAT	97–122	9–17
DYS438	NED	TTTTC	134–167	7–13
DYS437	NED	$[TCTA]_n[TCTG]_2[TCTA]_4$	173–198	12–18

(B)

Locus	Primer Sequences (5′ to 3′) (9)
DYS436	Forward: CCAGGAGAGCACACACAAAA
	Reverse: FAM-ACGAGCTGCGTTAGAGGTGA
DYS439	Forward: FAM-ACATAGGTGGAGACAGATAGATGAT
	Reverse: GCCTGGCTTGGAATTCTTTT
DYS435	Forward: GGGTTGTCCAGAGAAACAGC
	Reverse: FAM-CCCCCTCCTCTCGTCTATCT
DYS19	Forward: FAM-CTACTGAGTTTCTGTTATAGT
	Reverse: ATGGCATGTAGTGAGGACA
DYS460	Forward: GAGGAATCTGACACCTCTGACA
	Reverse: VIC-TCCATATCATCTATCCTCTGCCTA
H4	Forward: VIC-ATGCTGAGGAGAATTTCCAA
	Reverse: CTATTCATCCATCTAATCTATCCATT
DYS391	Forward: VIC-CTATTCATTCAATCATACACCCATAT
	Reverse: ACATAGCCAAATATCTCCTGGG
DYS392	Forward: NED-AAAAGCCAAGAAGGAAAACAAA
	Reverse: AAACCTACCAATCCCATTCCTT
DYS438	Forward: NED-TGGGGAATAGTTGAACGGTAA
	Reverse: GGAGGTTGTGGTGAGTCGAG
DYS437	Forward: GACTATGGGCGTGAGTGCAT
	Reverse: NED-AGACCCTGTCATTCACAGATGA

The paternity laboratory at Orchid GeneScreen-Dallas provided male samples from previously characterized parentage analyses for the haplotyping of father-son pairs and development of a database. The father-son pairs were typed with AmpFℓSTR® SGM Plus® (Applied Biosystems, Foster City, CA) and produced a probability of paternity of ≥99.33%. The population database consists of alleged father samples (one sample per case) from males of Caucasian, African American, and Southeast (SE) Hispanic origin that were selected via convenience sampling to ensure randomness (10). Autosomal STR DNA profiles were examined when a given haplotype appeared in more than one sample in order to ascertain specimen uniqueness. The DNA was isolated by FasTract® extraction (Orchid Cellmark, Dallas, TX). The paternity laboratory also provided 25 randomly selected female samples.

Optimization of the Polymerase Chain Reaction (PCR)

The PCR conditions established by NIST for the Y-STR 10-plex were modified to optimize its performance. Variables such as the magnesium concentration and annealing temperature were altered. Other modifications to the NIST protocol included slight alterations in primer concentrations and the addition of BSA (bovine serum albumin) to the reaction mix.

PCR

Ten Y-chromosome microsatellites were amplified in one PCR multiplex reaction. The multiplex examined the following loci: DYS436, DYS439, DYS435, DYS19, DYS460, Y-GATA-H4 (H4), DYS391, DYS392, DYS438, and DYS437.

The samples were amplified on a PE Applied Biosystems 9700 thermocycler (Applied Biosystems, Foster City, CA). The PCR thermocycling parameters were as follows:

- 95°C for 10 min
- 28 cycles of 94°C for 1 min, 60°C for 1 min, and 72°C for 1 min
- 60°C for 55 min

The total reaction volume for PCR was 20 μL. The PCR reaction mix contained 2.0 m*M* $MgCl_2$, 1x GeneAmp® PCR Gold buffer, and 2.5 units AmpliTaq Gold® DNA polymerase (Applied Biosystems, Foster City, CA). The master mix also contained 250 μ*M* of each GeneAmp® dNTP (Applied Biosystems, Foster City, CA), 0.2 μg/μL BSA, and primers ranging in concentration from 0.2 to 1.6 μ*M*. The primers were synthesized by Applied Biosystems, and one primer from each primer set was tagged with one of three dyes:

FAM™, VIC™, or NED™. The loci DYS436, DYS439, DYS435, and DYS19 were labeled with FAM™; DYS460, H4, and DYS391 with VIC™; and DYS392, DYS438, and DYS437 with NED™. GeneScan™-500 LIZ™ (Applied Biosystems, Foster City, CA) served as the internal size standard. The amount of genomic DNA added to each reaction was 1.25 ng unless otherwise stated or 2.5 μL of unquantified DNA solutions in the case of the father-son pairs and database samples. Two microliters of female DNA (0.5 to 2.5 ng/μL) were amplified to determine the specificity of the 10-plex.

Capillary Electrophoresis

PCR amplification products were subjected to capillary electrophoresis on the ABI PRISM® 3100 Genetic Analyzer (Applied Biosystems, Foster City, CA). The amplified products were first diluted 1:2 with sterile, molecular biology-grade water. Sample preparation for the ABI PRISM® 3100 Genetic Analyzer was as follows (per sample): 8.7 μL Hi-Di™ formamide (Applied Biosystems, Foster City, CA), 0.3 μL GeneScan™-500 LIZ™ Size Standard (Applied Biosystems, Foster City, CA), and 1 μL PCR product. The samples were denatured for 3 to 5 min at 95°C and subsequently snap-cooled on ice for 3 to 5 min. They were electrophoresed on the ABI PRISM® 3100 Genetic Analyzer using dye set G5 and the GeneScan36_POP4 Default Module. Following capillary electrophoresis, the samples were analyzed using GeneScan® Analysis (a minimum of 100 RFU [relative fluorescence units] peak heights were called) and Genotyper® version 2.5.2 software (Applied Biosystems, Foster City, CA).

Sensitivity Studies

Varying amounts of ATCC 45514 ranging from 0.025 to 2.5 ng were amplified to determine the minimum amount of input DNA that could be used to obtain a full Y-STR 10-plex profile. Each quantity of DNA was tested three times, with the exception of 0.5 ng, which was tested twice.

Mixture Studies

Male (ATCC 45514) and female (9947A) DNA was mixed in the following ratios: 1:1, 1:5, 1:10, 1:25, 1:50, and 1:100. In each mixture the amount of male DNA was kept constant at 0.5 ng. Male:male mixtures (ATCC 45514:BS #1) were performed in the following ratios using a total of 1.5 ng DNA: 100:0, 90:10, 75:25, 50:50, 25:75, 10:90, and 0:100.

Statistical Analysis

All alleles were reported in repeat numbers as they compare to SRM 2395, a sequenced standard obtained from NIST. Allele frequencies were calculated for each locus using the formula p_i = # occurrences/n, where p_i is the allele frequency and n represents the sample size. Gene diversity, D, of each locus was computed using the formula $D = (n/n - 1)(1 - \sum p_i^2)$ (11). Haplotype diversity (*HD*) was calculated with the same equation using haplotype frequencies rather than allele frequencies. For the Y-chromosome, haplotype diversity is equal to the power of discrimination (11). The number of haplotype occurrences and their relative frequencies were calculated using Arlequin software (12). Paternity calculations were performed with an in-house laboratory information management system using formulae from the American Association of Blood Banks Guidance for Standards for Parentage Testing Laboratories (13).

To determine the likelihood of patrilineal relationship, the following equation was used: $P_RI = X/Y = X/[(d + 1)/n + 1)]$, where d is the number of times the profile was found in the Y-STR haplotype database, and n is the number of individuals in the database. P_RI is the patrilineal relationship index. X is the likelihood of the individuals tested sharing the same haplotype. X is equal to 1 if the two Y-STR profiles are identical or 0 if they are significantly different. The probability of patrilineal relationship was calculated as $P_RI/(P_RI + 1)$ using a prior probability of 0.5.

Results and Discussion

Optimization of the PCR Reaction

The PCR conditions were optimized as described in Materials and Methods. Figure 1 provides an example of Y-STR results both before and after optimization of the PCR reaction. PCR optimization led to a decrease in the number of non-specific peaks, decreased baseline levels, and increased RFUs.

Analysis of Father-Son Pairs

Father-son pairs from previously characterized parentage analyses were examined to ensure that each pair generated identical Y-chromosome haplotypes. Of the 42 father-son pairs examined, 41 pairs produced identical haplotypes. In one father-son pair, a difference was observed at the H4 locus in which the father had a 13-repeat allele and the son had a 12-repeat allele. The difference is likely caused by a mutational event, which DNA sequencing could confirm. Calculation of a mutation rate in the present study would be inaccurate due to the small sample size. Based on 4999 male germ line transmissions, Kayser and Sajantila have previously estimated the locus-specific mutation rate at up to eight mutations per 1000 father-son pairs (14). The frequency of mutations is dependent on the Y-STR locus, with an average of three per 1000 father-son pairs (14). This is comparable to the mutation rates reported for autosomal STR loci (14–17). It is possible that the H4 locus may be more prone to mutational processes than other Y-STR loci. A mutation study that analyzes additional father-son pairs would facilitate either confirming or rejecting this hypothesis. Knowledge of the mutation rates at the H4 and other Y-STR loci is an important consideration in developing interpretational guidelines for the laboratory.

Development of a Population Database

In order to determine the distribution of Y-chromosome haplotypes across populations, samples of Caucasian, African American, and SE Hispanic males were analyzed using the Y-STR 10-plex.

As shown in Table 2, overall gene diversity (*D*) ranged from 0.071 (DYS436) to 0.698 (DYS19). Three hundred fifty seven different haplotypes were detected among the three populations (n_t = 525). The most frequent haplotype across all populations was found in 20 individuals (3.8%) and consisted of: DYS436 (12), DYS439 (19), DYS435 (11), DYS19 (14), DYS460 (11), H4 (12), DYS391 (11), DYS392 (13), DYS438 (12), DYS437 (15). The cumulative haplotype diversity (*HD*) was 0.996, and the probability of finding the identical haplotype in a pair of random, unrelated males is 0.37% (calculated as $[1 - HD] \times 100$).

There were 158 different haplotypes in the Caucasian population (n = 225), producing a haplotype diversity of 0.991. DYS439 was the most diverse locus (D = 0.630), while DYS436 was the least diverse (D = 0.044) in the Caucasian population. The average per locus gene diversity over all ten loci was 0.452290 ± 0.247141.

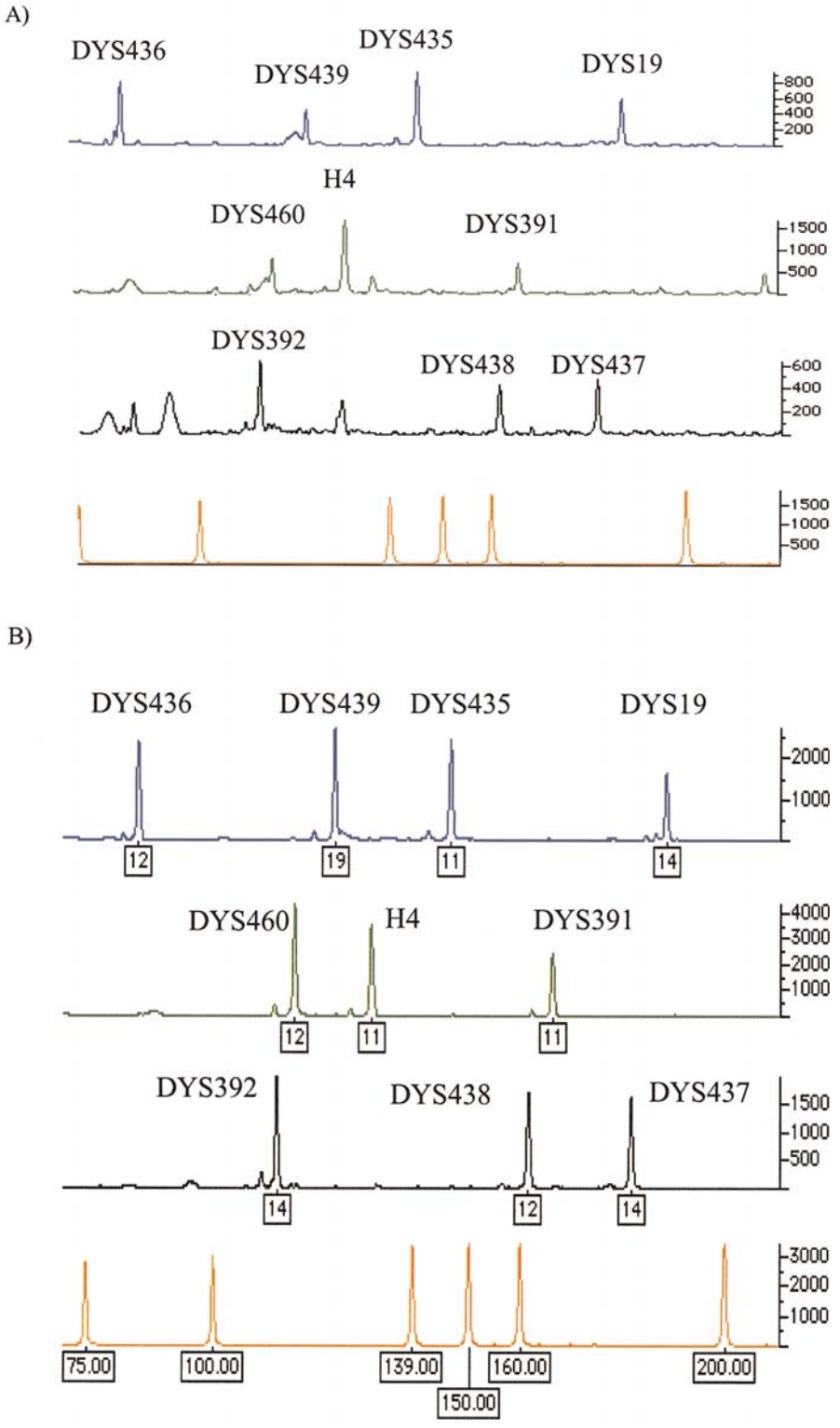

FIG. 1—*Optimization of the Y-STR 10-plex PCR reaction; (A) Y-STR 10-plex before optimization of the PCR reaction; (B) Y-STR 10-plex following optimization of the PCR reaction.*

In a sample size of 206 African American males, there were 149 different haplotypes, and the haplotype diversity was 0.995. The most diverse locus in the African American population was DYS19 ($D = 0.747$), and DYS435 was the least diverse ($D = 0.075$). Average gene diversity per locus for ten loci was calculated to be 0.434037 ± 0.238505.

A total of 87 different haplotypes were detected in the SE Hispanic population ($n = 94$). The haplotype diversity was calculated as 0.998. DYS19 exhibited the most diversity in the Hispanic population ($D = 0.709$), whereas DYS436 showed the least diversity ($D = 0.063$). The average per locus gene diversity across all loci was 0.523084 ± 0.282743.

TABLE 2—*Allele frequencies and gene diversity across the Caucasian (n = 225), African American (n = 206), and SE Hispanic (n = 94) populations (n_t = 525). All sizes are in repeat numbers as they compare to a sequenced standard, SRM 2395. Gene diversity (D) for each locus was calculated as described in Materials and Methods.*

DYS436		DYS439		DYS435		DYS19	
Size	Freq	Size	Freq	Size	Freq	Size	Freq
10	0.006	15	0.002	6	0.002	12	0.004
11	0.019	16	0.006	8	0.002	13	0.057
12	0.964	17	0.021	10	0.011	14	0.459
13	0.01	18	0.267	11	0.954	15	0.252
14	0.002	19	0.512	12	0.03	16	0.118
		20	0.17			17	0.106
		21	0.021			18	0.004
		22	0.002				
D = 0.071		*D* = 0.638		*D* = 0.088		*D* = 0.698	

DYS460		H4		DYS391	
Size	Freq	Size	Freq	Size	Freq
7	0.002	9	0.01	9	0.017
9	0.03	10	0.051	10	0.617
10	0.335	11	0.451	11	0.345
11	0.568	12	0.436	12	0.019
12	0.059	13	0.048	13	0.002
13	0.006	14	0.002		
		15	0.002		
D = 0.562		*D* = 0.602		*D* = 0.501	

DYS392		DYS438		DYS437	
Size	Freq	Size	Freq	Size	Freq
9	0.004	7	0.002	12	0.002
10	0.008	8	0.013	13	0.032
11	0.501	9	0.046	14	0.469
12	0.053	10	0.179	15	0.39
13	0.35	11	0.328	16	0.095
13.2	0.004	12	0.408	17	0.01
14	0.065	13	0.025	18	0.002
15	0.008				
16	0.004				
17	0.004				
D = 0.620		*D* = 0.693		*D* = 0.619	

Locus multiplication has previously been reported at several Y-STR loci, including DYS19, DYS385, DYS390, and DYS436 (6,11,14,18). Three instances of locus duplication were observed in the present study, two at DYS19 and one at DYS435 (Fig. 2). In each of these three cases, two peaks of approximately even heights were observed, thus exceeding the acceptable stutter percentages, and in each instance of locus duplication the alleles were separated by only one repeat. The appearance of two peaks cannot be due to the males having Supermale Syndrome (XYY), as the X and Y amelogenin peak heights from AmpFℓSTR® SGM Plus® were approximately equivalent. Previous studies have estimated the frequency of allele duplication at DYS19 to be 0.12%, based on analysis of 7,772 individuals (14). Calculating the frequency of locus duplication in the current study would be inaccurate due to the limited sample size.

Locus duplication observed in the present study may be explained by one of two mechanisms. Although implausible, the appearance of two peaks may occur if the primer is complementary to the DNA in a manner in which it is able to bind in two places that are four base pairs apart. However, if this mechanism were to occur, one would expect the appearance of two peaks to occur more frequently than it did in the present study. Rather, locus multiplication likely occurs due to replicative transposition, a well-documented event in many eukaryotes, including humans (19,20). Replicative transposition causes the target area of DNA to be duplicated so that the target area remains unchanged and the transposable element is inserted into a second site (21). Replicative transposition results in an increased number of copies of the target DNA sequence (21). In order to be detectable by standard electrophoretic analysis, the transposition must be followed by a mutational event, thereby producing two alleles of different sizes.

Stutter Values

Stutter is a known PCR artifact that refers to the appearance of a minor peak one repeat unit smaller (or, more occasionally, larger) than the major STR product. The average stutter percentages observed at the Y-STR 10-plex loci are presented in Table 3 (n = 35). The average stutter values ranged from 3.6% (±0.23, 95% confidence interval) at DYS438 to 11.6% (±1.102, 95% confidence interval) at DYS392. The low percentage of stutter observed at DYS438 is likely due to its pentameric repeating unit as compared to the other 10-plex loci, which are tetra- and trinucleotide repeats.

Sensitivity Studies

The quantity of amplifiable DNA is often limited in forensic cases. It is advantageous to have a method of Y-STR testing that requires very little template DNA for the amplification process. The minimum amount of input DNA that could be used to obtain a full Y-STR profile and the corresponding peak heights are shown in Table 4.

Allele calls were obtained for each sample at all Y-STR 10-plex loci when the amount of input DNA was ≥0.125 ng. The optimal amount of DNA to incorporate into each reaction is 1.0 to 1.5 ng. At 2.5 ng DNA, some pull-up from other dyes was observed. Pull-up occurs when one dye color bleeds through to another dye channel, typically because of excessive amounts of DNA that saturate the dye channel (2). At 0.125 ng input DNA, the average peak height ranged from 256 to 850 RFUs. At 0.1, 0.075, and 0.05 ng input DNA, an average of 8, 5, and 3 loci produced results, respectively. At the lower limit of DNA that was tested, 0.025 ng, no alleles were observed, with the exception of a single amplification at the H4 locus.

Effect of Female DNA

Twenty-five females were tested to determine the specificity of the 10-plex. Many of the electropherograms from female samples had humps or peaks with heights of ~200 RFUs. The peaks may be a result of incorporating an excess amount of DNA into the amplification reaction. The peaks tend to be off-ladder and appear most commonly at ~136 and ~164 base pairs in the VIC™-labeled loci and at ~168 base pairs in the NED™-labeled loci. The nonspecific humps are caused by free dye artifacts, which can readily be removed through additional primer purification, rather than resulting from amplified DNA (data not shown). The artifacts observed in the female samples did not interfere with interpretation of

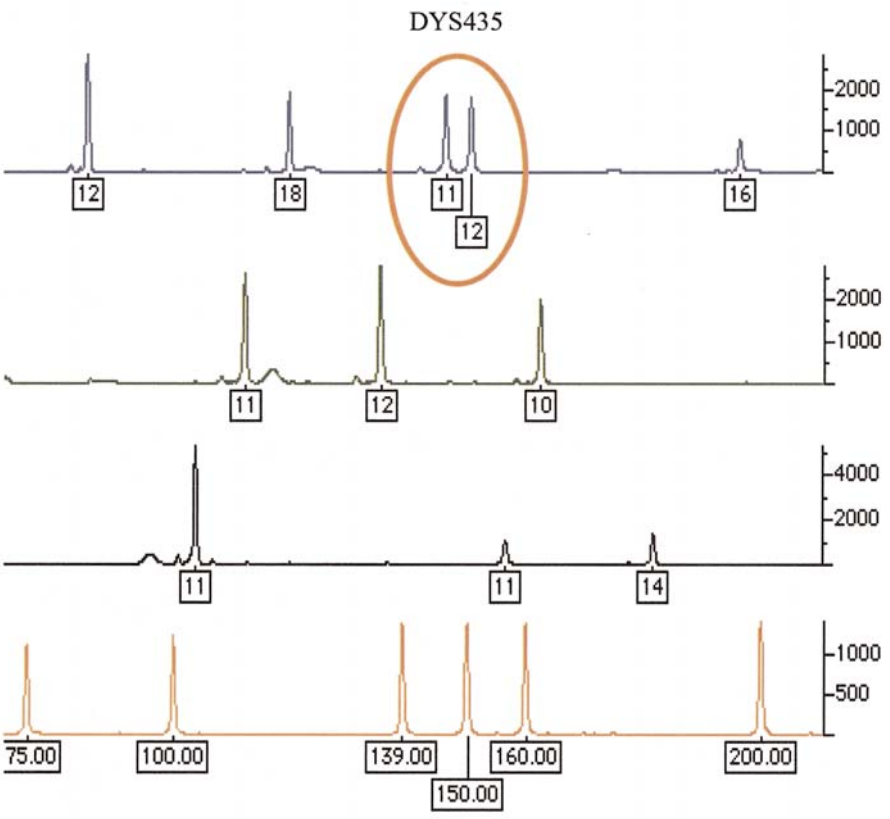

FIG. 2—*Locus duplication at DYS435.*

TABLE 3—*Stutter values for the Y-STR 10-plex. The average percent stutter (n = 35), standard deviation, and 95% confidence interval are listed for each Y-STR locus.*

	DYS436	DYS439	DYS435	DYS19	DYS460	H4	DYS391	DYS392	DYS438	DYS437
% Stutter	6.7	9.2	5.9	7.5	6.8	7.1	7.0	11.6	3.6	6.6
Std. Dev.	0.5	4.0	0.5	1.3	1.0	0.7	0.9	3.3	0.7	1.2
95% CI	0.170	1.324	0.177	0.418	0.335	0.224	0.300	1.102	0.230	0.401

TABLE 4—*Sensitivity of the Y-STR 10-plex using ATCC 45514. The values shown are the average peak heights of each locus in RFUs. NR (no result) indicates no alleles were observed.*

ng DNA	DYS436	DYS439	DYS435	DYS19	DYS460	H4	DYS391	DYS392	DYS438	DYS437
0.025	NR	NR	NR	NR	NR	116†	NR	NR	NR	NR
0.050	119*	112†	157†	NR	138*	172*	NR	117†	NR	NR
0.075	155	143*	225†	195†	229	135*	117†	174*	113†	NR
0.1	161*	197	190	165*	170	197*	166	292	140†*	136*
0.125	476	527	624	376	490	850	409	750	366	256
0.25	671	734	860	591	771	1041	525	960	440	388
0.5	1031	1293	1412	992	1322	1670	1169	1902	761	626
0.75	1812	2125	2396	1572	1905	2987	1719	2434	1217	984
1.0	2136	2494	2986	1783	2857	3052	1957	2542	1192	1204
1.5	5751	5965	6334	3562	7190	7745	5757	5426	3187	3838
2.0	5509	6354	6482	3606	6665	6832	5963	5844	3232	3370
2.5	5931	6151	6016	4234	6797	6999	5880	6233	3621	3796

* Only 2 of the 3 amplifications produced the allele.
† Only 1 of the 3 amplifications produced the allele.

the data, and in none of the cases would the female sample erroneously have been reported as originating from a male.

Mixture Studies

Male:female mixture studies were performed as described in Materials and Methods. The amount of female DNA incorporated into the reaction ranged from 0 to 100 times greater than the amount of male DNA. The female DNA did not cause any interference with interpretation of the male Y-STR profile at mixture ratios of 1:1, 1:5, or 1:10. A small peak, which could be accounted for by large amounts of female DNA, appeared in the NED™-labeled loci at a 1:25+ mixture. Although this peak exceeded the minimum RFU allele calling criterion (100 RFUs), it was larger than any DYS438 alleles observed in the current study and, therefore, did not interfere with interpretation of the male Y-STR profile at 1:25, 1:50, or 1:100 mixtures. Each laboratory must develop criteria for interpretation of male:female mixtures. At Orchid Cellmark, if an unexplained, extraneous peak appears in a male:female mixed sample, reference samples should be typed. If reference samples are unavailable for testing, we consider the locus in question to be inconclusive.

Mixtures of male:male DNA (ATCC 45514:BS #1) are shown in Fig. 3. Forensic samples, such as those from sexual assaults, sometimes have the DNA of more than one male present. In most of these cases, the male contribution is limited to two males (for example, a consensual partner and the sexual assault offender). For this reason, mixtures of two male DNA samples (ATCC 45514:BS #1) were analyzed (Fig. 3). Male:male mixtures could be readily detected when contaminating male genomic DNA was present at a relative proportion of at least 10%. When working with mixtures involving 10% contamination, two Y-STR haplotypes could readily be distinguished. At a 25:75 ratio, two haplotypes could be distinguished, yet in one instance the DYS439 locus produced ambiguous results.

When examining male:male mixtures, the ability to differentiate the major Y-STR haplotype from the minor component of the mixture depends on the peak positions of the alleles. If an allele falls in a stutter position, it may be difficult to determine whether the peak is a minor component of the mixture or if it is truly stutter. Clearly, this may impair the ability of an analyst to unambiguously differentiate between Y-STR haplotypes in the absence of reference samples. Knowledge of the acceptable stutter percentage for each locus is necessary for an analyst to be able to distinguish stutter peaks from true allele peaks (see Table 3). If a peak is above the acceptable stutter percentage, it should be considered an allele. Furthermore, when allele peak heights at a given locus are equivalent,

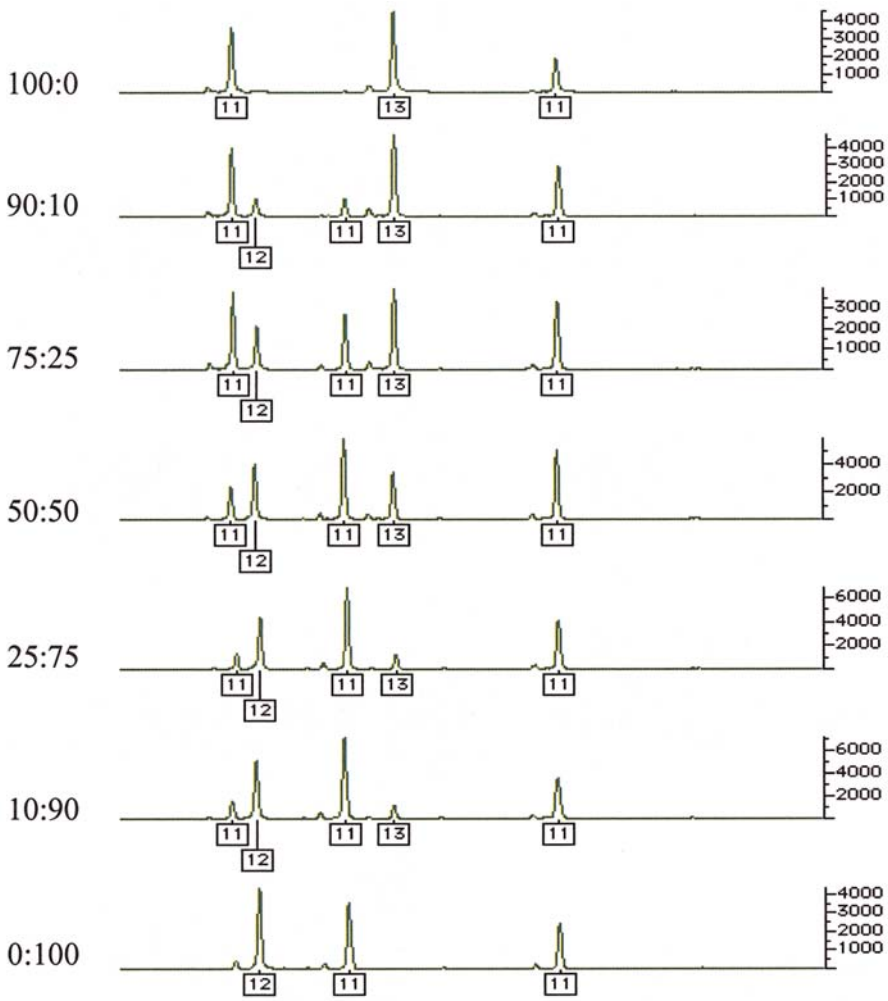

FIG. 3—*Male:male mixtures. Mixtures of the male DNA extracts ATCC 45514 and BS #1 were amplified in various ratios using a total of 1.5-ng input DNA. Only the VIC™-labeled loci are shown.*

two haplotypes cannot be distinguished. Because of the potential for locus multiplication to occur, analysts must take care when interpreting Y-STR haplotypes so that a single source sample is not erroneously interpreted as a mixed profile. Haplotyping a reference sample is necessary to confirm an instance(s) of locus duplication in an evidentiary sample. If a reference sample is unavailable and the only indication of a mixture is a possible locus duplication, we consider the locus in question to be inconclusive.

Allelic Ladder

Figure 4 illustrates the allelic ladder that has been developed for the Y-STR 10-plex. The ladder was developed using samples that had been characterized for the previously described population studies. The repeat numbers for each locus were determined by comparison to SRM 2395, a sequenced standard from NIST. Approximately 30 samples from the population data basing were selected so that many of the common alleles from each Y-STR locus were represented in the ladder.

Examples of Casework Using the Y-STR 10-Plex

Family Study—A family study was performed that involved two African American male children, their mother, and their alleged paternal grandfather. Each individual was analyzed at 19 autosomal STR loci and molecular HLA to give a combined paternity index (CPI) of 2.52, making it difficult to establish or exclude paternity. Y-STR analysis revealed that the two children matched at all Y-STR 10-plex loci. However, their Y-STR profile differed from the alleged paternal grandfather at 5 loci. Therefore, the two children cannot be excluded as being patrilineal relatives, but they are excluded from being biological grandchildren of the alleged paternal grandfather. Based on Y-STR analysis, the probability of patrilineal relationship of the two children is 99.52% as compared to an untested, randomly chosen man of the African American population (Prior Probability = 0.5).

Paternity—A recent case involved using Y-STRs to determine the paternity of a male child in which the alleged father (AF) is deceased. The persons available for testing in this case included the child, the child's mother, the AF's mother, and the AF's brother. Each individual involved in the case was typed using all available autosomal STR systems. As shown in Table 5, the paternity index obtained when testing all available parties is 0.5311, suggesting that the AF's mother and brother are not biological relatives of the child. When performing calculations based on different combinations of persons involved in the case (see Table 5), the paternity in-

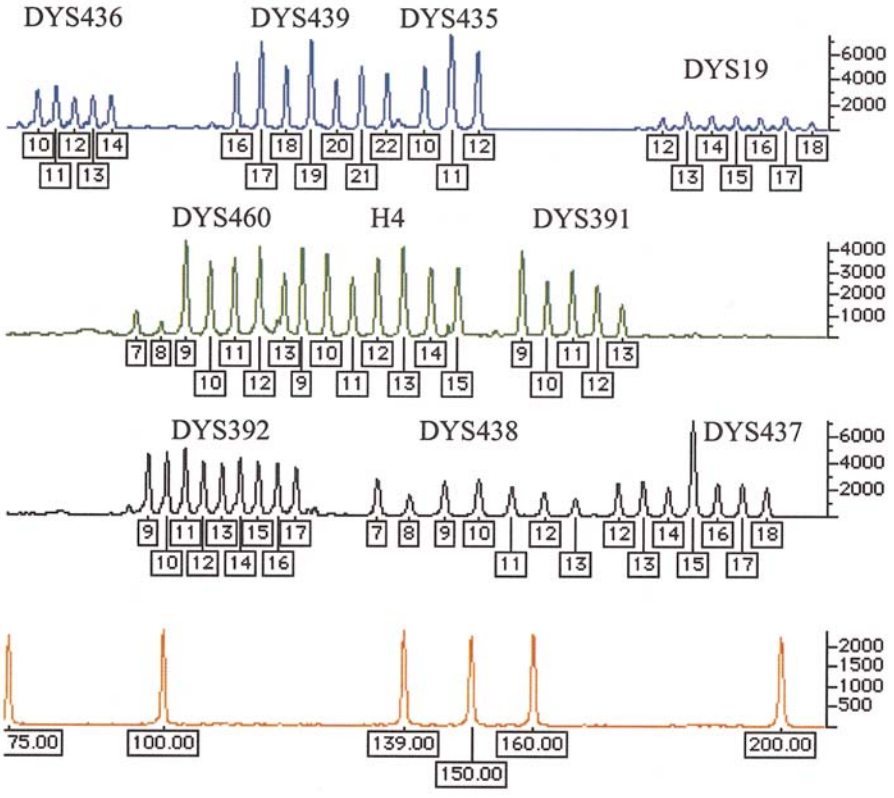

FIG. 4—*Y-STR 10-plex allelic ladder.*

TABLE 5—*Example case involving the questioned paternity of a male child. The paternity indices shown were calculated following autosomal STR testing.*

Individuals Tested	Male Child Mother AF's Mother AF's Brother	Male Child Mother AF's Brother	Male Child Mother AF's Mother	Male Child AF's Brother
Paternity Index	0.5311	0.5631	3.0086	0.04969

dex varied from 0.04969 to 3.0086, an inconclusive result.

DNA extracts from the child and AF's brother were submitted for Y-STR 10-plex analysis. The child and AF's brother did not match at five of the ten loci that were examined. Therefore, the two tested males were excluded as being patrilineal relatives. These results were based on the assumption that the AF's brother was paternally related to the alleged father.

In each of these scenarios, traditional methods of DNA testing failed to provide the statistical information needed to give convincing results in the case. These cases would likely have been reported as inconclusive, yet Y-STR analysis was able to demonstrate whether or not a patrilineal relationship existed between the males involved in each case.

Conclusions

Y-chromosome analysis gained public attention when the Y-chromosome haplotypes of two well-known leaders, Thomas Jefferson and Genghis Khan, were analyzed. It was believed that United States President Thomas Jefferson fathered Eston, the youngest son of Sally Hemings, one of his slaves (22). Although none of Thomas Jefferson's direct male descendents were available for testing, Y-chromosome data were collected on descendents of his paternal uncle, Field Jefferson (22). Five of Field Jefferson's male descendents were tested, and four had the same haplotype as Eston's male descendent (22). The fifth individual varied at one STR locus, which may be due to a mutation (22). Similar testing was performed in order to exclude other alleged fathers as being Eston's biological father (22).

Due to the mode of inheritance of the Y-chromosome, having identical Y-STR haplotypes includes an alleged father, or any of his patrilineal male relatives, as being a possible biological father of the child in question. Therefore, although Jefferson may in fact be Eston's biological father, the possibility cannot be ruled out that Eston was fathered by another patrilineal relative of Thomas and Field Jefferson.

DNA analysis of a second ruler, Genghis Khan, has also raised interest in Y-chromosome testing. Y-chromosome analysis of more than 2,000 males was performed in order to look at genetic variation within Asia (23). Interestingly, a cluster of closely related lineages (the "star cluster") was found in ~8% of the males throughout the portion of Asia that stretches from the Pacific to the Caspian Sea (23). The time to the most recent common ancestor (TMRCA) for this cluster was calculated to be ~1,000 years (23). The Hazaras have a Mongolian origin and believe that they are direct descendents of Genghis Khan (23). The Y-chromosome profiles of the Hazaras fell within the cluster, thereby suggesting that Genghis Khan did have the star cluster haplotype (23). Genghis Khan established history's largest land empire within ~1,000 years ago (23). Khan, along with his patrilineal relatives, conquered numerous populations and fathered many children (23). Although the Mongolian empire eventually disintegrated, Mongolians (including Khan's patrilineal descendents) continued to rule the region for several more centuries (23), thus spreading the Khan lineage. It is possible that the Khan lineage accounts for ~0.5% of Y-chromosome lineages worldwide (23).

As the Jefferson and Khan cases demonstrate, Y-chromosome testing can be used in instances of questioned paternity or population studies involving historical figures. Y-STRs continue to play a valuable role in modern-day forensics, such as in cases involving sexual assaults and forensic paternity. For this reason, we have validated a Y-STR 10-plex, as described in this paper, which has already been implemented in forensic and paternity casework. The results of this study suggest that the Y-STR 10-plex is highly discriminating and can effectively distinguish male lineages. The Y-STR 10-plex would, therefore, be an excellent addition to the autosomal STRs already employed in many forensic and paternity laboratories.

References

1. Ruitberg CM, Reeder DJ, Butler JM. STRBase: a short tandem repeat DNA database for the human identity testing community. Nuc Acids Res 2001;29:320–2.
2. Butler JM. Forensic DNA typing—biology and technology behind STR markers. San Diego: Academic Press, 2001.
3. Butler JM, Schoske R, Vallone PM, Kline MC, Redd AJ, Hammer MF. A novel multiplex for simultaneous amplification of 20 Y chromosome STR markers. Forensic Sci Int 2002;129:10–24.
4. Prinz M, Sansone M. Y chromosome-specific short tandem repeats in forensic casework. Croat Med J 2001;42:288–91.
5. Honda K, Roewer L, de Kniff P. Male DNA typing from 25-year-old vaginal swabs using Y chromosomal STR polymorphisms in a retrial request case. J Forensic Sci 1999;44:868–72.
6. Kayser M, Caglia A, Corach D, Fretwell N, Gehrig C, Graziosi G, et al. Evaluation of Y-chromosomal STRs: a multicenter study. Int J Legal Med 1997;110:125–33.
7. Roewer, L. Fundamental strategies in forensic and genealogical casework using the human Y-chromosome. "Y-chromosome analysis and its application to forensic casework" workshop; 2002 Feb. 11–16; Atlanta, Georgia. American Academy of Forensic Sciences, 2002.
8. Federal Bureau of Investigation. Quality Assurance Standards for Forensic DNA Testing Laboratories, 1998.
9. Schoske R, Vallone PM, Ruitberg CM, Butler JM. Multiplex PCR design strategy used for the simultaneous amplification of 10 Y chromosome short tandem repeat (STR) loci. Anal Bioanal Chem 2003;375:333–43.
10. National Research Council. The evaluation of forensic DNA evidence. Washington, D.C.: National Academy Press, 1996.
11. 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. Foren Sci Int 2002;125:42–51.
12. Schneider S, Roessli D, Excoffier L. Arlequin version 2.000: a software for population genetics data analysis. Genetics and Biometry Laboratory, University of Geneva, Switzerland, 2000.
13. American Association of Blood Banks. Guidance for standards for parentage testing laboratories. Bethesda: American Association of Blood Banks, 2000.
14. Kayser M, Sajantila A. Mutations at Y-STR loci: implications for paternity testing and forensic analysis. Forensic Sci Int 2001;118:116–21.
15. Brinkmann B, Klintschar M, Neuhuber F, Hühne J, Rolf B. Mutation rate in human microsatellites: influence of the structure and length of tandem repeat. Am J Hum Genet 1998;62:1408–15.
16. Henke L, Henke J. Mutation rate in human microsatellites. Am J Hum Genet 1999;64:1473.
17. Sajantila A, Lukka M, Syvänen A.-C. Experimentally observed germline mutations at human micro- and minisatellite loci. Eur J Hum Genet 1999;7:263–6.
18. http://www.promega.com/geneticidproc/eusymp2proc/03.pdf
19. Carvalho-Silva DR, Pena SD. Molecular characterization and population study of an X chromosome homolog of the Y-linked microsatellite DYS391. Gene 2000;247:233–40.
20. Sargent CA, Boucher CA, Blanco P, Chalmers IJ, Highet L, Hall N, et al. Characterization of the human Xq21.3/Yp11 homology block and conservation of organization in primates. Genomics 2001;73:77–85.
21. Lewin B. Genes VI. 6th rev. ed. Oxford: Oxford University Press, 1997.
22. Foster EA, Jobling MA, Taylor PG, Donnelly P, de Knijff P, Mieremet R, et al. Jefferson fathered slave's last child. Nature 1998;396:27–8.
23. Zerjal T, Xue Y, Bertorelle G, Wells RS, Bao W, Zhu S, et al. The genetic legacy of the Mongols. Am J Hum Genet 2003;72:717–21.
24. http://www.cstl.nist.gov/biotech/strbase

Additional information and reprint requests:
Cassie L. Johnson, M.S.
Forensic DNA Analyst
Orchid Cellmark-Dallas
2600 Stemmons Fwy, #133
Dallas, TX 75207
E-mail: cjohnson@orchid.com