J Forensic Sci, Mar. 2003, Vol. 48, No. 2
Paper ID JFS2001316_482
Available online at: www.astm.org

TECHNICAL NOTE

Edward L. Buse,[1] *M.A.; Jeanne C. Putinier,*[1] *M.S.; Mary M. Hong,*[1] *B.S.; Aimee E. Yap,*[1] *B.S.; and John M. Hartmann,*[1] *B.S., M.B.A*

Performance Evaluation of Two Multiplexes Used in Fluorescent Short Tandem Repeat DNA Analysis*

ABSTRACT: The performance of two commercial multiplex kits that together amplify the 13 core short tandem repeat (STR) loci currently in use by forensic laboratories and the U.S. national Combined DNA Indexing System (CODIS) were evaluated. The typing systems examined were AmpFℓSTR® Profiler Plus™ and AmpFℓSTR COfiler™ (PE Applied Biosystems, Foster City, CA). Electrophoretic separation and detection of the fluorescent PCR products was achieved by capillary electrophoresis (CE) using an ABI Prism® 310 Genetic Analyzer. The studies addressed the on-site validation of the instrument, the software, and each typing system. These studies included instrument sensitivity, resolution, precision, binning, peak height ratios, mixtures, stutter, and the amplification of non-probative and simulated forensic samples. Other additional developmental-type work is also reported herein, such as species specificity testing and amplification of environmentally insulted samples. Amplification conditions were found to be robust and the primer sets shown to be specific to human DNA. Stutter and peak height ratios fell within limits published by the manufacturer and other laboratories. The data demonstrate that the CE instrument can consistently resolve fragments differing in length by one base and that the ± 0.5 base bin used by the Genotyper® software is acceptable for making accurate allele calls. Correct typing results were obtained from non-probative and simulated case samples, as well as samples exposed to outdoor environmental conditions. The results support the conclusion that DNA extracted from biological samples routinely encountered in the forensic laboratory can be reliably analyzed with AmpFℓSTR Profiler Plus and COfiler using CE.

KEYWORDS: forensic science, DNA typing, PCR, capillary electrophoresis, STR, validation, Profiler Plus, COfiler

Developmental and in-house validation studies should be performed prior to use of a procedure or method on forensic DNA casework (1—3). Although the Applied Biosystems' (ABI, Foster City, CA) AmpFℓSTR® Profiler Plus™ and COfiler™ multiplex systems have been commercially available and widely adopted by the U.S. forensic community for several years, published validation work on these systems has appeared in the literature only more recently. This may be due in part to the fact that these systems simply represent a further evolutionary development of other well-established PCR-based technologies (4–10). The AmpFℓSTR systems represent another step in this technology evolution and combine to amplify the full set of thirteen CODIS core short tandem repeat (STR) loci: D3S1358, vWA, FGA, D16S539, D8S1179, D21S11, D18S51, TH01, TPOX, CSF1PO, D5S818, D13S317, and D7S820, plus Amelogenin (gender determination). Applied Biosystems has included validation and population studies in its AmpFℓSTR Profiler Plus and COfiler User's Manuals (11,12) and similar studies addressing validation, population databases, locus independence, and concordance specific to the AmpFℓSTR multiplex systems have appeared in the past few years (13–25).

Detection of PCR product alleles can be accomplished with either flat-bed or capillary electrophoresis (CE), well-established technologies in their own right. The literature on CE and forensic STR typing is too large to cite here, but there is an extensive on-line bibliography described by Ruitberg, et al. (26). The ABI Prism® 310 Genetic Analyzer, a single-capillary platform capable of unattended analysis of up to 96 samples, has demonstrated its utility for typing STR loci (27–30). Along with the Genotyper® software, the simultaneous separation and analysis of six to ten loci enable rapid and efficient profiling of reference and evidence samples. This study examines both AmpFℓSTR Profiler Plus and COfiler systems by altering and examining amplification conditions, CE injection parameters, resolution, sizing precision, allele designation accuracy, stutter, and heterozygote peak ratios. Mixtures, non-probative evidence, environmental degradation effects, and species specificity were also observed.

Materials and Methods

DNA Extraction, Quantitation and Sample Sources

Genomic DNA was prepared using an organic phenol/chloroform procedure (31), followed by quantitation with either a chemiluminescence-based, slot blot hybridization procedure employing a

[1] Forensic Science Services, DNA Section, Orange County Sheriff-Coroner Department, 320 N. Flower Street, Santa Ana, CA.

* Presented in part as a poster at the 12th International Symposium on Human Identification, Biloxi, MS, October 2001.

Received 14 Sept. 2001; and in revised form 9 April 2002; accepted 3 Nov. 2002; published 6 Feb. 2003.

primate-specific *Alu* probe (PS-1) developed in-house (32) or with the QuantiBlot® Human DNA Quantitation Kit (Perkin Elmer, Norwalk, CT) (33). Recovery of non-human DNA was estimated by yield-gel electrophoresis and visualized by ethidium bromide staining against known high molecular weight standards.

Most studies (except where indicated in the text) used known blood and buccal cell samples derived from laboratory personnel. Other sample sources included the following: (*a*) DNA from the commercial cell line GM9947A (34); (*b*) non-primate animal tissue (mammals—bear, cat, cow, coyote, dog, horse, opossum, pig, rabbit, raccoon, sheep, skunk, and squirrel; reptiles—snake and turtle; fish—cod and salmon; birds—chicken, crow, duck, goose, and turkey) donated by the U.S. Department of Fish and Wildlife laboratory (Ashland, OR); (*c*) primate tissues (chimpanzee, gorilla, lemur, macaque, orangutan, and red howler) donated by the San Diego Zoo (San Diego, CA); (*d*) bacterial DNA (*Escherichia coli*) purchased commercially (Boehringer Mannheim Inc., now Roche Diagnostics Corp., Indianapolis, IN) and baker's yeast (*Saccharomyces cerevisiae*) from a local market; (*e*) unpreserved, whole blood deposited on cotton cloth that was exposed to environmental conditions for various periods of time: indoors at room temperature for 6 years, outdoors in the shade for 6 months, and in the sun for 4, 6, 8, 10, 12, 16 weeks; and (*f*) sixty-five forensic samples derived from seven non-probative cases, five simulated sexual assault kits, and three proficiency test sets.

Amplification Conditions

Human DNA (125pg–20ng) and non-human DNA (2ng and 20ng) was amplified in 50 μL reaction volumes with the reagents provided in the AmpFℓSTR Profiler Plus and COfiler PCR Amplification Kits (Applied Biosystems Inc., Foster City, CA). Except as noted, the protocols suggested in the AmpFℓSTR User's Manuals (11,12) were followed. Amplification was performed in the GeneAmp PCR Systems 9600 thermalcycler (Perkin Elmer) using 0.2 mL MicroAmp® reaction tubes.

Sample Preparation, Electrophoresis and Data Analysis

All instrumentation, software and materials described below were obtained from PE Applied Biosystems, Inc. (Foster City, CA). Amplification products were separated and detected on an ABI Prism® 310 Genetic Analyzer. Generally, 2.0 μl of amplified product and 0.5 μL of GeneScan®-500 (ROX) Internal Lane Size Standard were added to 24 μl of HiDi™ formamide, denatured at 95°C for 3 min, and snap-cooled in an ice water bath for 3 min. Unless otherwise indicated, the PCR products were injected for 5 s and separated at 15 kV in Performance Optimized Polymer 4 (POP-4™) using a 47 cm capillary (50 μM i.d.), 1X Genetic Analyzer Buffer (with EDTA) and 60°C run temperature. Data were collected using the ABI Prism 310 Collection software (version 1.0.2) with virtual filter set F. Analysis of the data was performed with GeneScan® Analysis software (version 2.1) with peak detection threshold set at 50 rfu and without smoothing. Genotyper® software (version 2.0) was used to assign alleles to sample peaks.

Amplification Performance Studies

For the studies below only one variable at a time was altered while holding the other conditions constant. Each study used lab staff samples (1.25 ng) typed with AmpFℓSTR Profiler Plus and/or 9947A (2.0 ng) typed with AmpFℓSTR COfiler: (*a*) *AmpliTaq Gold® DNA Polymerase*—Amplification was performed under standard conditions except that the *Taq* Gold was diluted in 1X GeneAmp PCR buffer to yield (x) units of enzyme in a fixed 10 μ L volume (where x = 0.5, 1.0, 2.5, 5.0, 7.5, 10.0, 20.0 units for AmpFℓSTR Profiler Plus and x = 1.0, 2.0, and 20.0 units for AmpFℓSTR COfiler). (*b*) *Amplification Components*—Samples were subjected to alterations in the proportions of the AmpFℓSTR PCR Reaction/Primer mixture. The reaction buffer or the primer mix was adjusted ±25% by volume, while holding the other component volumes constant. Total volume was not held constant. (*c*) *Final Reaction Volume (AmpFℓSTR COfiler only)*—The final PCR mixture consists of AmpFℓSTR PCR Reaction Mix, COfiler Primer Set, and *Taq* Gold. The normal total volume used per reaction is 30 μL. This was varied by ±25%, keeping the proportions of the components constant. (*d*) *Annealing/Denaturing Temperature*—Samples were subjected to incremental changes (±2°C in 1°C increments) in either annealing or denaturing temperature, while holding the remaining parameters unchanged. The sample block wells were temperature checked for accuracy with a NIST-traceable calibrated digital thermometer (model HH21; Omega Engineering Inc, Stamford, CT). (*e*) *Cycle Number (2.0 ng 9947A only)*—Amplifications were performed in which the cycle number was varied from 26–30 cycles. All other conditions were standard. (*f*) *Final Extension Time (AmpFℓSTR Profiler Plus only)*—A single lab staff sample was amplified under standard conditions, except that a 3 μL portion of the sample was removed at regular intervals (15, 30, 45, 60, 90, 120 min; 45 min is standard) during the final extension phase of the thermalcycler run. Total reaction volume was reduced during the course of the experiment from 50 μL (t = 0 min) to 35 μL (t = 120 min).

ABI Prism 310 Performance Studies

Sensitivity and Instrument Comparison—DNA from four blood samples (AmpFℓSTR Profiler Plus only) and 9947A (AmpFℓSTR COfiler only) were prepared at several different template concentrations (0.125–20 ng) in a fixed volume of 20 μL. Typing results were assessed for accuracy, PCR artifacts, and baseline anomalies. The performance between two different instruments (CE1 and CE2) was evaluated by comparing the relative peak heights of samples prepared at the template concentrations above and with varying volumes (4, 6, and 8 μL) of PCR product. The PCR product was prepared at double the normal volume, divided equally, and run in duplicate on both CE1 and CE2. Capillaries with approximately the same number of injections were used in each instrument.

Resolution—Data was collected from validation and casework runs over several months. For D21S11 alleles 28.2 and 29 which exhibit baseline separation, the Base Resolution (R_b) was defined as: $R_b = \Delta M/R$, where ΔM = actual base difference between two peaks (approximately 2.0) and $R = [2(d_1 - d_2)/(W_1 + W_2)]$, where d = distance between the two measured peaks in scan numbers, and W = peak base width in scan numbers (30,35,36). One-base resolution was examined between alleles 9.3 and 10 at locus TH01. These peaks, which are not resolved to baseline, were evaluated by measuring the Valley Value (V) defined as $V = H_v/H_h$, where H_v is the height (in rfu) of the "valley" between two peaks and H_h is the height (in rfu) of the higher of the two peaks (29,30).

Precision and Binning Studies—Measurement precision (inter- and intra-) was defined as the standard deviation of mean fragment size and was evaluated by plotting the standard deviation in bases against the mean fragment size for each allele studied. Within each run, the size deviation was defined as the absolute size difference in bases between the 9947A sample allele and the corresponding

allele in the allelic ladder (37). Some runs contained more than a single 9947A sample. Injection interval was defined as the number of intervening injections between the 9947A sample and the nearest allelic ladder. The size deviation in bases between the allelic ladder and the nearest 9947A sample was plotted against this injection interval.

Peak Height Ratio and Stutter—The mean peak height and the peak height ratio (peak height of the lower peak divided by the peak height of the higher peak) were calculated for each allelic pair. Both homozygote and heterozygote profiles were used to calculate stutter, except where primary peaks overlapped with neighboring stutter peaks. Stutter data from primary peaks exhibiting very low (<100 rfu) or very high (>5500 rfu) signal strength were not included. Stutter signals below 50 rfu were measured manually using the cursor provided by GeneScan. A second stutter study and an additional data set for locus D21S11 was obtained from database samples amplified with the AmpFℓSTR® Green II kit (PE Applied Biosystems) and typed with an ABI Prism 310 Genetic Analyzer.

Mixtures—Several two-component DNA mixtures were generated at three total DNA concentrations of approximately 5.0, 2.0, and 0.5 ng. The contribution of one component (A) in a two-component (A+B) mixture was varied (A = 0, 5, 10, 25, 50, 75, 90, 95, 100%), while holding the total DNA concentration and total amplification volume fixed. The 2.0 ng stock solution was internally calibrated by checking that a 50:50 mixture gave roughly identical peak heights for the non-overlapping alleles within a locus, and that the overall peak height was close to that for 2.0 ng of the 9947A DNA.

Results and Discussion

Amplification Performance Studies

Modifications to the manufacturer's recommended amplification protocol were made in order to ascertain whether non-standard operating conditions affect the accuracy of the AmpFℓSTR typing results. Such conditions can arise from the inadvertent use of poorly calibrated equipment (e.g., pipettors and thermalcyclers), the use of compromised reagents, and/or due to human error. Only a brief written summary of the results is presented (data not shown). Altering the amount of AmpliTaq Gold during amplification of both AmpFℓSTR systems had little effect on the resulting DNA profiles. As expected, a reduction in signal intensity was noted for amounts of enzyme below 2.5 units (5.0 units is recommended). This was especially apparent for larger DNA fragments. An increase of "minus-A" PCR product was noted primarily for the smaller DNA fragments when a limiting (<1 unit) amount of enzyme was used. Alterations to the reaction mix (±25% by volume) had little or no effect on the resulting DNA profiles. In contrast, a notable increase in peak height was observed when the primer component volume was increased +25%, while a reduction in this same component (−25%) had little effect on signal intensity relative to the control sample. Variation in the total amplification volume (±25%) while maintaining component proportions constant had little or no effect on the results, with peak heights relatively unchanged.

Modification of the cycling temperatures during PCR (annealing 59°C ± 2°C, denaturing 94°C ± 2°C) did not lead to allele dropout or false positive results. As expected, varying the cycle number caused the greatest change with a noticeable loss of signal (decrease in rfu) at lower cycle numbers (26 and 27 cycles) and an increase in signal with higher cycle numbers (29 and 30 cycles). No discernable difference was observed in the resulting DNA profiles for any of the modified extension times (15–120 min). However, under non-standard conditions (e.g., presence of inhibitors), other investigators have reported that longer final extension times proved beneficial, especially in increased efficiency of non-templated nucleotide addition (38).

ABI Prism 310 Performance Studies

Sensitivity and Instrument Comparison—Complete DNA profiles were obtained from all test samples (in both AmpFℓSTR kits) with input DNA between 500 pg and 5.0 ng using the Genotyper software (data not shown). Template concentrations at 125 pg yielded peak heights below 100 rfu and frequently irregular peak shape (no data smoothing; data not shown). As a result, at some larger loci, peaks became difficult to distinguish from background noise. Heterozygote peak height ratios were also found to be more variable and would sometimes drop below nominal levels (see Peak Height Ratio section below). At the lowest level of input DNA (125 pg) nearly complete profiles were obtained, although some alleles fell below the allele assignment threshold and the overall performance of the system was diminished at this level. Although some individual allele dropout was observed at the lowest concentration (125 pg) tested with either typing system, it did not appear to be associated with any particular locus. The phenomenon of allele dropout or "sister allele" imbalance at heterozygous loci can be due to stochastic amplification (5,39), an effect due to very low number of template DNA molecules. The establishment of an empirically determined stochastic threshold ("S-line") has been proposed as an aid to interpretation (17), such that patterns with peaks below this level should be interpreted with caution.

DNA template amounts greater than 5.0 ng resulted in peak heights exceeding the linear response limit of the detection apparatus (approx. > 4500 rfu) or reached the detector's absolute upper limit (>8191 rfu in the raw data) (11,12). As template concentration increased, many primary peaks exhibited varying amounts of "pull-up," artificially inflated stutter values, and incomplete 3′-adenylation ("minus-A peaks") (40). The cleanest and most reliable results were achieved in the range between 500 pg and 2.5 ng of input DNA with no anomalies or gross artifacts observed. A preferred DNA target for amplification lies between 1 and 2 ng; a level for which a two-fold variation in the estimate of input DNA should still yield reliable results.

Variability in the relative performance of CE instruments can lead to potentially different stochastic thresholds, since sensitivity depends upon such factors as laser output, optics, and the sensitivity and spectral response of the camera. The comparison of peak height data between CE1 and CE2 using duplicate samples and different PCR product volumes demonstrated that the relative difference in sensitivity on these two particular instruments was not significant (data not shown). Automated genotyping results obtained from both instruments were identical, with the exception of grossly overloaded samples (differing OLA calls) or in very weak samples (125 pg) where some individual peaks fell below the threshold (for making allele calls) of one, but not the other, instrument.

Resolution—The evaluation of resolution can serve as a quality control measure for an individual injection or run and as an indicator of system performance over an extended time period (30). In practice, changes in resolution may be used to monitor the run-to-run consistency of a system's materials and/or run conditions and may include the formamide, separation media, column, voltage and

TABLE 1a—*Resolution (R_b, V) comparison between CE1 and CE2 in casework.*

Alleles	Fragment Length (Bases)	CE	Mean R_b^*	S.D.†	Mean V‡	S.D.†
D21S11 28.2, 29	205, 207	1	1.24	0.09 (87)	…	…
		2	1.31	0.08 (38)	…	…
TH01 9.3, 10	188, 189	1	…	…	0.43	0.03 (6)
		2	…	…	0.49	0.05 (22)

* R_b = base resolution. See Method section.
† Standard deviation followed by number of individual runs examined in parentheses.
‡ V = height of the "valley" between two peaks relative to the height of the larger peak.

TABLE 1b—*Injection parameters and resolution (V)*.*

Injection Voltage	Injection Time (s)				
	3	4	5	6	7
10,000	0.56	0.57	0.56	0.57	0.57
12,000	0.55	0.56	0.56	0.58	0.59
14,000	0.56	0.56	0.56	0.61	0.67
15,000	0.55	0.59	0.62	0.62	0.68

* See footnote under Table 1a for definition of (V).

TABLE 2—*Precision study. Three types of samples were evaluated with both AmpF*ℓ*STR systems: an allelic ladder, 9947A cell line control, and a forensic mixed sample.*

Sample	No. of Injections	n*	Standard Deviation	
			Max.	Min.
Profiler Plus				
Allelic ladder	12†	1416	0.13	0.04
9947A	11	165	0.08	0.03
Mixed sample	11	319	0.09	0.01
COfiler				
Allelic ladder	20	1080	0.11	0.04
9947A	20	240	0.08	0.05
Mixed sample	20	380	0.13	0.05

* n = total alleles analyzed.
† An additional bracketing ladder was included at the end of the CE run.

temperature (41–43). Resolution on CE1 and CE2 was examined over a period of several months using the 9.3 and 10 alleles at locus TH01 (one-base separation) and the 28.2 and 29 alleles at locus D21S11 (two-base separation) (Table 1*a*). Base resolution (R_b) did not exceed a value of 1.44 on either instrument. A plot of R_b over time revealed no trend (data not shown). Single-base peak separations were monitored by measuring the height of the "valley" (V) between them relative to the higher peak. Maximum values of 0.47 and 0.57 were observed for CE1 and CE2 respectively, however, the mean V values did not exceed 0.50. A pair of theoretical peaks merged at half-height would have a V value of 0.50 and an approximate value of R_b of 1.7 (30). In practice, however, peak overlap and the summing of this overlap would cause the equivalent measured V to be larger. For example, using TH01 data and the same electrophoretic system, Buel, et al. (see Table 1; Ref. 30) have reported an R_b value of 1.13 and equated this to a measured V of 0.51. Changes in electrokinetic injection parameters can also alter the resolution (Table 1*b*). Raising the injection time and voltage simultaneously will generally lead to a reduction in resolution. With short injection times alone, resolution was less sensitive to changes in injection voltage. Lengthening the duration at lower voltages also had little effect on resolution. Signal intensity increased with larger voltages and time, but in a non-linear manner (data not shown).

Precision and Binning—DNA fragment-size data were tabulated for three different types of samples on CE1 and CE2 (Table 2). The maximum observed standard deviation for intra-run data across all alleles and samples was 0.13 bases, however, the majority of values were well below 0.10. A regression analysis of the precision against mean fragment size of the ladder alleles sized in AmpFℓSTR Profiler Plus revealed that absolute measurement imprecision increased with increasing fragment size, but the coefficient of variation fell slightly (Fig. 1). Similar results were obtained with AmpFℓSTR COfiler (data not shown). Inter-run precision data was also collected for 9947A alleles and their corresponding sizing ladder alleles over a $1\frac{1}{2}$-year period. The maximum inter-run standard deviation was 0.16 bases (data not shown).

DNA fragment size data was used to record the absolute size differences in bases between the 9947A sample alleles and the corresponding alleles in the sizing ladder. This yielded a total of 3102 pair-wise comparisons (Fig. 2). In no instance did the observed comparisons fall outside the established ± 0.50 base bin boundary. An apparently negative (i.e., the 9947A allele is smaller in size than the corresponding ladder allele) or positive sizing error was observed for a number of alleles. For example, allele 11 ($\sim$217 bases; $n = 146$) at locus D13S317 appears to consistently size slightly smaller than the allelic ladder, possibly due to differences in base sequence between the two measured DNA fragments; however, sampling error cannot be ruled out. Size deviation was also assessed as a function of the injection interval, defined as the number of injections occurring between the sizing ladder and the test sample (9947A). The degree of size deviation never exceeded the established bin boundary even when the injection interval was as great as 16 (approx. 9 h; data not shown). The precision and size deviation results demonstrate that a one-base bin is justified in making accurate allele assignments.

AmpFℓSTR System Performance Studies

Peak Height Ratios—Preferential amplification or an imbalance among heterozygote allele pairs (sister alleles) has been observed for alleles differing greatly in size (39). A slight imbalance in peak intensity among closely spaced alleles may also occur even when input DNA concentration is above stochastic-effect levels. In this study, a total of 981 allelic pairs were evaluated for evidence of peak imbalance at varying concentrations of input DNA. The majority of allele pairs differed in size by eight bases, with the great-

est observed difference of 32 bases occurring at locus D18S51. Generally, peak heights were well balanced when input DNA was 1.25 ng or greater. The median peak height ratio was found to be about 0.90 across all loci (Table 3). However, a single value as low as 0.49 was observed when mean peak heights fell below 500 rfu. Conversely, peak height ratios below 0.72 (95th percentile) were rarely observed when peak heights occurred above 1000 rfu. Nominal amplifications presumed to contain 1.25 ng of template could actually contain less than this amount as a result of error in quantitation estimates. PCR amplification at stochastic levels and/or under low copy number conditions (44) is observed to increase heterozygote peak imbalance.

In Fig. 3 a subset of samples (filled triangles) had either very low peak height ratios or low peak height ratios accompanied by relatively high rfu values. These samples were examined further by re-amplification using the same or higher DNA concentration. The peak height ratios upon re-amplification returned to nominal levels (i.e., >0.70; data not shown). This suggests that the original peak

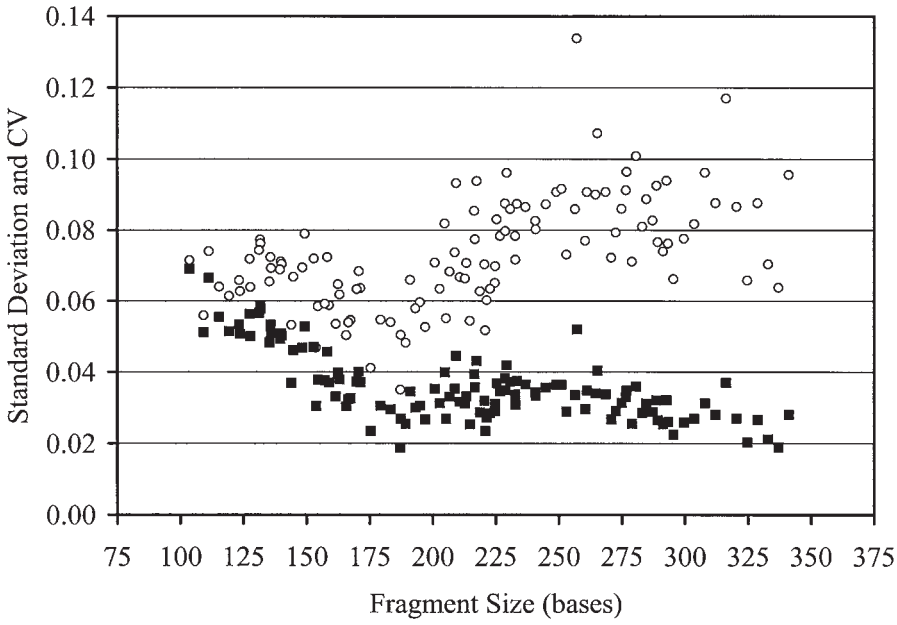

FIG. 1—*Sizing precision on CE1 using the AmpFℓSTR Profiler Plus allelic ladder. The standard deviation (○) and the coefficient of variation (CV; ■) were determined for each ladder allele. AmpFℓSTR Profiler Plus ladder injected 12 times.*

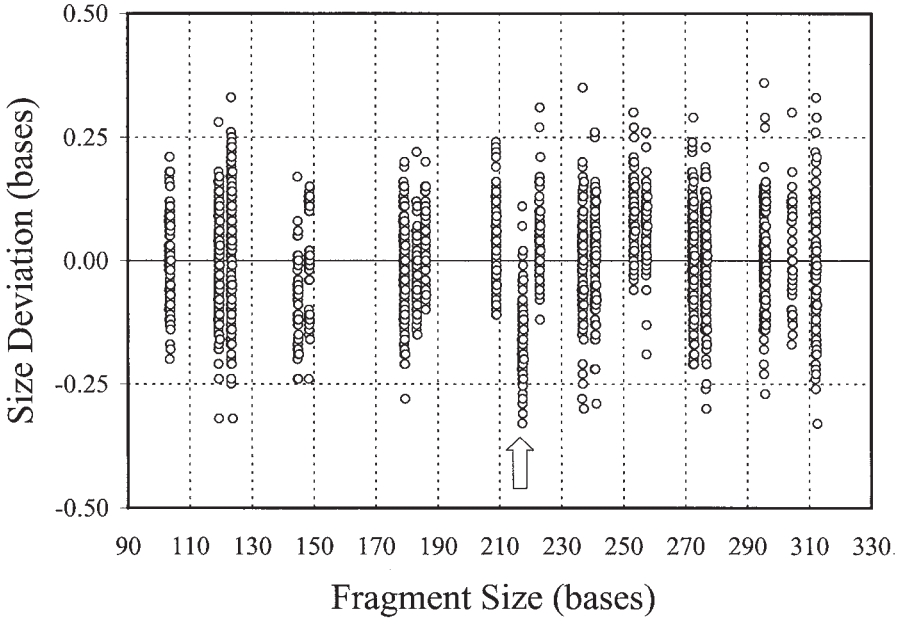

FIG. 2—*Size variation as a function of allele fragment size. A total of 22 alleles (9947A sample) were examined covering a broad range of fragment sizes and collected over numerous CE runs (Jan. 1998–Aug. 1999). Pair-wise comparisons (n = 3,102) were used to measure the size difference in bases between the sample allele and the corresponding allele in the reference ladder. Overlapping data sets occur at ~181 bases (vWA allele 17, TH01 allele 8) and at ~295 bases (D18S51 allele 15 and CSF1PO allele 10).*

height imbalances could have been caused by stochastic fluctuations during amplification. Other observed peak imbalances could possibly be due to primer binding site mutations (19,45–48,60). One particular sample (arrow; Fig. 3) upon re-amplification (at 2.0 and 2.5 ng) consistently maintained a low relative peak ratio at D16S539 when compared to other loci within the same profile. Further studies are planned examining the effects of reduced stringency conditions on primer binding.

Slight variation in relative peak heights has been reported for replicate amplifications of the same DNA sample (see Fig. 6; Ref. 17). Typically, the smaller of two heterozygous DNA fragments will exhibit on average slightly greater peak height (i.e., positive asymmetry). AmpFℓSTR Profiler Plus allelic pairs (n = 266) having a ≥ 8 base separation were examined for positive or negative asymmetry. Across all loci, the majority of allelic pairs (64%) displayed positive asymmetry (data not shown). Individually, loci exhibited differing levels of positive asymmetry (55–76%), except locus D8S1179, where virtually no bias (48%) in asymmetry was observed. The prevalence of positive asymmetry may be attributed to preferential amplification, and/or preferential electrokinetic injection of smaller DNA fragments.

TABLE 3—*Peak height ratios for the AmpFℓSTR loci.*

Mean Peak Height (RFU)	Count*	Peak Height Ratio		
		Median	Minimum	S.D.†
<500	75	0.87	0.49	0.13
500–1000	161	0.89	0.57	0.09
1000–1500	176	0.91	0.66	0.08
1500–2000	136	0.90	0.68	0.07
2000–2500	124	0.91	0.66	0.07
2500–3500	110	0.93	0.77	0.06
>3500	120	0.94	0.72	0.06
All data	902	0.91	0.49	0.09

Percentile (%)	Peak Height Ratio ≥
25	0.96
50	0.91
75	0.85
95	0.72
99	0.61

* Count = Number of peak pairs examined at a given RFU range.
† S.D. = Standard deviation.

Stutter—All thirteen CODIS core loci are tetranucleotide repeats and known to possess moderate levels of stutter, usually less than 10% (9,11,12,17,18,48,49,50). Stutter product is typically found to be four bases shorter (n−4) than the actual allele, although $n-8$, $n-12$, and $n+4$ stutter have been observed, but usually only under non-standard PCR conditions (e.g., excessive template concentration) (17). The determination and evaluation of stutter from known single source samples may be used as an aid for interpreting mixed samples (49). Mean allele-specific stutter for both AmpFℓSTR systems ranged from 2.6% (TH01,TPOX) to 7.5% (D18S51), with an observed maximum at 13.8% (Table 4). A large a number of primary peaks (n = 133) had no associated stutter and 32% of these were TH01 alleles. The lack of measurable stutter, however, may be due to measurement limitations and/or limited data. In no instance did the stutter values exceed those published in the AmpFℓSTR Users Manuals (11,12).

The amount of observed stutter increased with increasing DNA fragment size, although a few exceptions can be found for specific alleles within a locus, presumably due to allele-specific sequence heterogeneity (48,49,51–53). Interrupting long stretches of a core repeat sequence with other unique sequences can result in a reduction in overall stutter, as has been reported for vWA (Fig. 3; Ref. 49) and FGA (48). A series of alleles with reduced stutter can be observed at locus D21S11 (Fig. 4). The D21S11 locus has a complex sequence motif in which a two base pair TA insertion gives

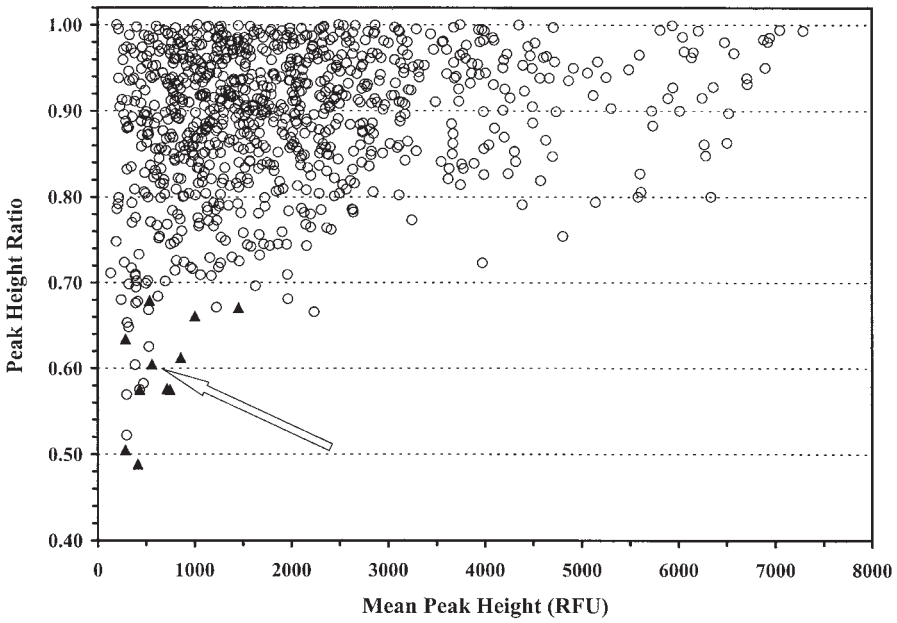

FIG. 3—*Peak height ratio study. A total of 902 allelic pairs were examined from samples amplified using 1.25, 2.0, and 5.0 ng of template DNA. The mean peak height in rfu was determined for each pair of sister alleles and plotted against the associated peak height ratio. Some samples with unusual peak ratios (▲) were subsequently re-amplified at the same or higher template concentration. All, but one (see arrow), peak ratio examined in this manner then yielded values >0.70.*

TABLE 4—*Examination of stutter in the AmpFℓSTR loci.*

Locus	n*	Mean	S.D.†	Upper‡ Limit	Observed High Value	Observed Low Value
Profiler Plus						
D3S1358	63	6.8	1.3	10.7	9.5	4.3
vWA	78	7.4	1.9	13.0	11.4	2.8
FGA	122	7.2	1.5	11.9	10.2	2.7
D8S1179	87	6.1	1.5	10.5	9.3	2.9
D21S11	104	6.5	1.2	10.3	9.5	3.7
D18S51	116	7.5	1.9	13.3	13.8	4.1
D5S818	69	5.1	1.3	9.1	7.9	1.7
D13S317	65	4.0	1.5	8.5	7.2	1.5
D7S820	75	4.2	1.3	8.0	7.5	1.8
Cofiler						
D3S1358	84	6.8	1.6	11.6	12.9	3.5
D16S539	45	5.2	1.7	10.2	9.0	1.9
TH01	60	2.6	1.0	5.5	4.8	1.2
TPOX	101	2.6	1.0	5.5	6.5	1.2
CSF1P0	98	5.0	1.1	8.4	8.7	2.7
D7S820	91	4.7	1.3	8.7	7.8	2.1

* n = number of alleles examined.
† S.D. = Standard Deviation.
‡ Upper limit is defined as the mean value + 3 S.D.

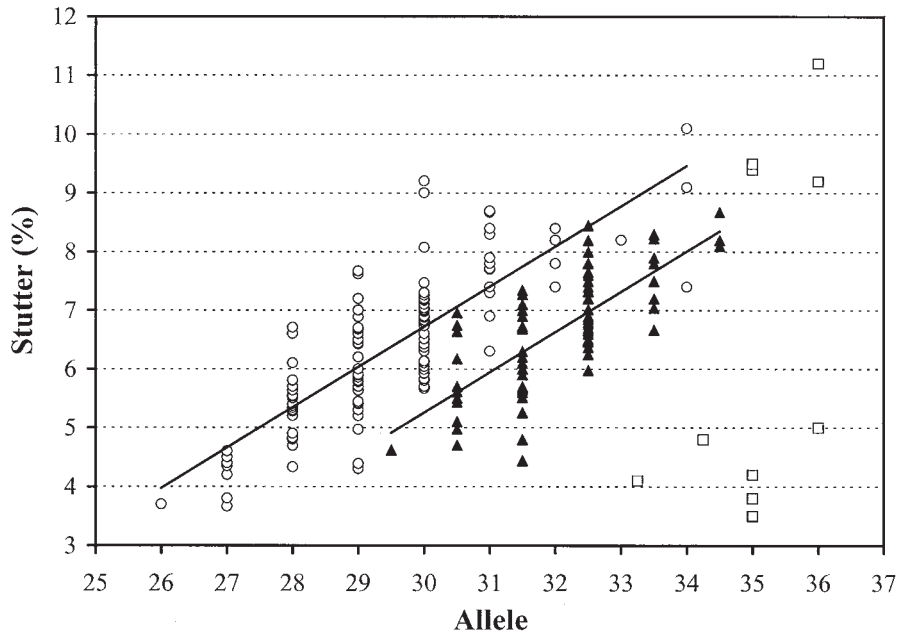

FIG. 4—*Stutter examined at locus D21S11. Stutter values (n = 222) were assigned to three allele categories: whole repeats (○), 2-base microvariants (▲), and other (□) for alleles 33.1, 34.1, 35, and 36. A linear regression was applied to the data in the first two categories. All stutter values were obtained from data that excluded 4-base peak separations.*

rise to a subclass of repeats (51–53). These two-base microvariants tend to have significantly less stutter than the consensus alleles having no insertion. Although FGA and other loci are known to have a number of microvariant alleles, it was not possible to examine the occurrence of reduced stutter at these loci due to the relative rarity of these microvariant alleles and to the limited size of our database.

Mixtures—Detection of a minor component depends substantially on the amount of input DNA. In general, minor components not overlapped by other peaks or stutter could be detected at levels greater than 25% of the major component at the 0.5 ng DNA input level, at 10–25% (locus dependant) with the 2.0 ng input level, and at 5% at the highest DNA input level of 5.0 ng (data not shown). Minor peak detection also varied by locus with detection generally greater for the smaller loci. Minor components differing by only one base from the major peak were difficult to resolve and detection required that the minor component comprise 25–50% of the sample. All mixed samples with peaks over 100 rfu, including many peaks differing from each other by a single base were cor-

rectly typed (data not shown). The interpretation and assessment of mixed profiles and the application of any associated inclusion probabilities is beyond the scope of this paper. Guidelines and rules for mixture interpretation have been proposed by several authors (17,54–56).

Specificity—Nonhuman species were examined for the possible appearance of peaks as a result of primer binding with the nonhuman DNA. The appearance of nonhuman DNA fragments has been reported by other laboratories (9,11–13,15,48,57–59), with most observations limited to single peaks or peaks occurring outside the normal locus category. In this study, we report either no signal or only a single yellow (Y) or green (G) peak from the following species: bear (G), cat, chicken, cod, cow (G), coyote (G), crow, dog (G), duck, *E. coli*, goose, horse (G), opossum, pig (G), rabbit, raccoon (G), salmon, sheep (G), skunk (Y), squirrel, snake, turkey (Y), turtle (Y), and yeast (data not shown). The single green peaks observed are smaller in molecular weight than the human X amelogenin allele (59). The skunk sample gave an unusual result whereby it reproducibly yielded a 132-base NED-labeled (yellow) peak when amplified with AmpFℓSTR COfiler, but not with AmpFℓSTR Profiler Plus (data not shown). The primer for locus D7S820 is the only NED-labeled primer in the AmpFℓSTR COfiler multiplex. This same locus (and primer pair) is also present in the AmpFℓSTR Profiler Plus multiplex. Apparently, the NED-labeled D7S820 primer in AmpFℓSTR COfiler is pairing with an unlabeled primer from one of the other loci in the multiplex.

The lower primates possessed an amelogenin profile identified by Genotyper as either X or X,Y. The lemur DNA, however, produced no other signals for any of the thirteen loci tested. The macaque sample gave weak results at FGA and CSF1PO, but these were all classified as off-ladder. Interestingly, red howler DNA amplified with AmpFℓSTR Profiler Plus yielded an unusual NED-labeled DNA pattern in the 260 base size range (normally associated with locus D7S820) that consisted of nine peaks of varying height each separated by two bases (Fig. 5). AmpFℓSTR COfiler, however, did not reproduce these results, suggesting that this nine-peak pattern is the product of a NED-labeled primer paired with an unlabeled primer from another locus found in AmpFℓSTR Profiler Plus but not in AmpFℓSTR COfiler.

The apes (Chimpanzee, Gorilla, and Orangutan) all exhibited some peaks that Genotyper assigned to alleles at a particular locus. All the multi-locus profiles were distinct from human profiles in that many of the peaks were either weak, off ladder, unbalanced, or out of the normal reported size range for the loci involved. There is no possibility that these primate profiles could be confused with that of a human. A recent study comparing human and non-human higher primate DNA sequences at six loci (FGA, D3S1358, vWA, CSF1PO, TPOX, TH01) reported significant differences in sequence homology (48).

Environmental Exposure—Two unpreserved blood samples were exposed to Southern California outdoor conditions with fluctuations in temperature, sun, humidity, air pollutants, and airborne microorganisms. Complete profiles were obtained with both AmpFℓSTR systems for all controls and a subset samples tested (data not shown). Locus dropout was first noted at eight weeks of sun exposure (D7S820). As expected, the larger DNA fragments were more susceptible to degradation and were preferentially lost with increased exposure. In contrast, D3S1358 and amelogenin consistently yielded results up to 16 weeks of exposure.

Degradation did not lead to an increase in stutter (data not shown). Peak height ratios (n = 154) were not grossly affected, but values below 0.70 became more prevalent (9% of total) with increased exposure, primarily as a result of reduced signal strength (data not shown). Although the data support the observa-

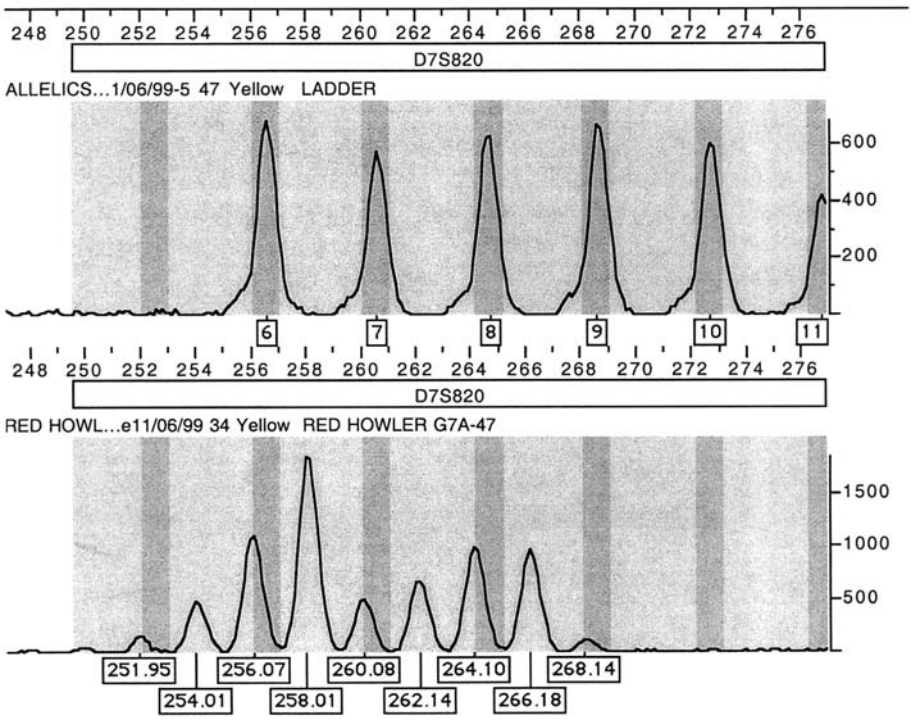

FIG. 5—*NED-dye labeled DNA fragments from Red Howler upon amplification with AmpFℓSTR Profiler Plus. A multi-peak profile (2-base separation) is observed between 250 and 270 bases falling in the analytical range of locus D7S820. No profile was obtained with AmpFℓSTR COfiler (D7S820 only), suggesting that these amplicons are the product of a NED-labeled primer paired with an unlabeled primer not found in the AmpFℓSTR COfiler multiplex.*

tion that DNA damage and locus dropout increase with the length of exposure, no mistyping occurred as a result of the environmental degradation.

Nonprobative Evidence—A total of 65 forensic samples were typed in one or both AmpF*ℓ*STR systems. Typing results from known sources were compared to evidentiary samples as well as simulated questioned samples. All samples known to have a common origin gave matching results (data not shown). A subset of these samples (n = 55) typed in both systems gave matching results for the loci shared in common by the two systems (D3S1358, Amelogenin, and D7S820).

Conclusion

The studies presented support the conclusion that DNA extracted from biological samples routinely encountered in the forensic laboratory can be reliably amplified and typed with the AmpF*ℓ*STR Profiler Plus and COfiler systems followed by capillary electrophoresis. No instances of false results, even under adverse conditions, were observed. In addition, the data demonstrated that the ABI Prism 310 can consistently resolve fragments differing in length by one base and that the ±0.5 base bin used by the Genotyper software is appropriate for making accurate allele calls. Multiplex amplification and capillary electrophoresis have proven to be sensitive, specific, and precise. Both technologies are capable of processing a wide variety of samples (environmentally insulted, non-probative, mixed, and degraded) making them invaluable as a forensic tool.

Acknowledgment

The authors wish to acknowledge the contributions of Susan Beal and Richard Gustilo for their technical assistance in sample processing. We would also like to thank the following agencies for donating samples: U.S. Fish and Wildlife, San Diego Zoo, San Diego Police Department, and Los Angeles County Sheriff's Office.

References

1. Budowle, B. Technical Working Group on DNA Analytical Methods. Guidelines for a quality assurance program for DNA analysis. Crime Lab Digest 1995;22:21–43.
2. DNA Advisory Board (DAB). Federal Bureau of Investigation, U.S. Department of Justice, Washington D.C., 1998.
3. Gill P, Sparkes R, Fereday L, Werrett DJ. Report of the European Network of Forensic Science Institutes (ENSFI): formulation and testing of principles to evaluate STR multiplexes. Forensic Sci Int 2000;108:1–29.
4. Blake E, Mihalovich J, Higuchi R, Walsh PS, Erlich HA. Polymerase chain reaction (PCR) amplification and human leukocyte antigen (HLA)-DQα oligonucleotide typing on biological evidence samples: casework experience. J Forensic Sci 1993;37:700–26.
5. Budowle B, Lindsey JA, DeCou JA, Koons BW, Giusti AM, Comey CT. Validation and population studies of the loci LDLR, GYPA, HBGG, D7S8, and Gc (PM loci) and HLA-DQα using a multiplex amplification and typing procedure. J Forensic Sci 1995;40:45–54.
6. Cosso S, Reynolds R. Validation of the AmpliFLP™ D1S80 PCR amplification kit for forensic casework analysis according to TWGDAM guidelines. J Forensic Sci 1995;40:424–34.
7. Lins AM, Sprecher CJ, Puers C, Schumm JW. Multiplex sets for the amplification of polymorphic short tandem repeat loci—slver stain and fluorescent detection. Biotechniques 1996;20:882–9.
8. Budowle B, Moretti TR, Keys KM, Koons BW, Smerick JB. Validation studies of the CTT STR multiplex system. J Forensic Sci 1997;42:701–7.
9. Wallin JM, Buoncristiani MR, Lazaruk KD, Fildes N, Holt CL, Walsh PS. TWGDAM validation of the AmpF*ℓ*STR Blue PCR amplification kit for forensic casework analysis. J Forensic Sci 1998;43:117–33.
10. Micka KA, Amiott EA, Hockenberry Tl, Sprecher CJ, lins AM, Rabbach DR, et al. TWGDAM validation of a nine-locus and a four-locus fluorescent STR multiplex system. J Forensic Sci 1999;44:1243–57.
11. PE Applied Biosystems. AmpF*ℓ*STR COfiler User Bulletin. Foster City, CA, 1998.
12. PE Applied Biosystems. AmpF*ℓ*STR Profiler Plus User's Manual. Version A. Foster City, CA, 1997.
13. Cotton EA, Allsop RF, Guest JL, Frazier RRE, Koumi P, Callow IP, et al. Validation of the AMPF*ℓ*STR SGM Plus system for use in forensic casework. Forensic Sci Int 2000; 112:151–61.
14. Holt Cl, Buoncristiani M, Wallin JM, Nguyen T, Lazaruk KD, Walsh PS. TWGDAM validation of AmpF*ℓ*STR™ PCR amplification kits for forensic DNA casework. J Forensic Sci 2002;47:66–96.
15. Wallin JM, Holt Cl, Lazaruk KD, Nguyen TH, Walsh PS. Constructing universal multiplex PCR systems for comparative genotyping. J Forensic Sci 2002;47:52–65.
16. Pawlowski R, Maciejewska A. Forensic validation of a multiplex contained in nine STRs–population genetics in Northern Poland. Int J Legal Med 2000;114:45–49.
17. Moretti TR, Baumstark Al, Defenbaugh DA, Keys KM, Smerick JB, Budowle B. Validation of short tandem repeats (STRs) for forensic usage: performance testing of fluorescent multiplex STR systems and analysis of authentic and simulated forensic samples. J Forensic Sci 2001;46: 647–60.
18. Frank WE, Llewellyn BE, Fish PE, Riech AK, Marcacci Tl, Gandor DW, et al. Validation of the AmpF*ℓ*STR Profiler Plus PCR amplification kit for use in forensic casework. J Forensic Sci 2001;46:642–6.
19. Budowle B, Sprecher CJ. Concordance study on population database samples using the PowerPlex 16 kit and AmpF*ℓ*STR Profiler Plus kit and AmpF*ℓ*STR COfiler kit. J Forensic Sci 2001;46:637–41.
20. Budowle B, Moretti TR, Baumstark Al, Defenbaugh DA, Keys KM. Population data on the thirteen CODIS core short tandem repeat loci in African Americans, U.S. Caucasians, Hispanics, Bahamians, Jamaicans, and Trinidadians. J Forensic Sci 1999;44:1277–86.
21. Holt Cl, Stauffer C, Wallin JM, Lazaruk KD, Nguyen T, Budowle B, et al. Practical applications of genotypic surveys for forensic STR testing. Forensic Sci Int 2000;112:91–109.
22. Sacchetti L, Calcagno G, Coto I, Tinto N, Vuttariello E, Salvatore F. Efficiency of two different nine-loci short tandem repeat systems for DNA typing purposes. Clinical Chem 1999;45(2):178–83.
23. Tomsey CS, Kurtz M, Kist F, Hockensmith M, Call P. Comparison of PowerPlex 16, PowerPlex 1.1/2.1, and ABI AmpF*ℓ*STR Profiler Plus/COfiler for forensic use. Croat Med J 2001;42(3):239–43.
24. Budowle B, Shea B, Niezgoda S, Chakraborty R. CODIS STR loci data from 41 sample populations. J Forensic Sci 2001;46:453–89.
25. LaFountain MJ, Schwartz MB, Svete PA, Walkinshaw MA, Buel E. TWGDAM validation of the AmpF*ℓ*STR Profiler Plus and AmpF*ℓ*STR COfiler multiplex systems using capillary electrophoresis. J Forensic Sci 2001;46:1191–8.
26. Ruitberg CM, Reeder D, Butler JM. STRBase: a short tandem repeat DNA database for the human identity testing community. Nucleic Acid Res 2001;29(1):320–2.
27. Buel E, Schwartz MB, LaFountain MJ. Capillary electrophoresis STR analysis: comparison to gel based systems. J Forensic Sci 1998;43: 164–70.
28. Lazaruk K, Walsh PS, Oaks F, Gilbert D, Rosenblum BD, Menchen S, et al. Genotyping of forensic short tandem repeat (STR) systems based on sizing precision in a capillary electrophoresis instrument. Electrophoresis Sci 1998;19:86–93.
29. Moretti TR, Baumstark Al, Defenbaugh DA, Keys KM, Brown Al, Budowle B. Validation of STR typing by capillary electrophoresis. J Forensic Sci 2001;46:661–76.
30. Buel E, LaFountain M, Schwartz M, Walkinshaw M. Evaluation of capillary electrophoresis performance through resolution measurements. J Forensic Sci 2001;46:341–45.
31. Hartmann JM, Keister RS, Luber S, Baldwin RA, White JM. A DNA extraction and purification method used in the Orange County Sheriff-Coroner Department Crime laboratory. American Academy of Forensic Science, Anaheim, CA, 1991.
32. Buse E, Goodhart S. Development of an economical and sensitive alkaline-phosphatase probe for human DNA quantitation. Proceedings from the fifth international symposium on human identification, Scottsdale, AZ 1994;163–4.
33. Walsh PS, Valaro J, Reynolds R. A rapid chemiluminescent method for quantitation of human DNA. Nucleic Acids Res 1992;20:5061–65.
34. Fregeau CJ, Aubin RA, Elliott JC, Gill SS, Fourney RM. Characterization of human lymphoid cell lines GM9947 and GM9948 as intra- and

inter-laboratory reference standards for DNA typing. Genomics 1995; 28:184–97.
35. Luckey JA, Norris TB, Smith LM. Analysis of resolution in DNA sequencing by capillary gel electrophoresis. J Phys Chem 1993;97: 3067–75.
36. Biega LA, Duceman BW. Substitution of H_2O for formamide in the sample preparation protocol for STR analysis using the capillary electrophoresis system: the effects on precision, resolution, and capillary life. J Forensic Sci 1999;44:1029–31.
37. Gill P, Urquhart A, Millican E, Oldroyd N, Watson S, Sparkes R,et al. A new method of STR interpretation using inferential logic—evelopment of a criminal intelligence database. Int J Leg Med 1996;109: 14–22.
38. Fregeau CJ, Germain O, Fourney RM. Fingerprint enhancement revisited and the effects of blood enhancement chemicals on subsequent Profiler Plus fluorescent short tandem repeat DNA analysis of fresh and aged bloody fingerprints. J Forensic Sci 2000; 45(2):354–80.
39. Walsh PS, Erlich HA, Higuchi R. Preferential PCR amplification of alleles: mechanisms and solutions. PCR Methods Appl 1992;1:241–50.
40. Magnuson Vl, Ally D, Nylund SJ, Karanjawala ZE, Rayman JB, Knapp JI, et al. Substrate nucleotide-determined non-templated addition of adenine by Taq DNA polymerase: implications for PCR based genotyping and cloning. Biotechniques 1996;21:700–9.
41. Guttman A, Schwartz HE. Artifacts related to sample introduction in capillary electrophoresis affecting separation performance and quantitation. Anal Chem 1995;67:2279–83.
42. Rosenblum BB, Oaks F, Menchen S, Johnson B. Improved single-strand DNA sizing accuracy in capillary electrophoresis. Nucleic Acids Res 1997;25(19):3925–9.
43. Wenz HM, Robertson JM, Menchen S, Oaks F, Demorest DM, Scheibler D, et al. High-precision genotyping by denaturing capillary electrophoresis. Genome Res 1998;8(1):69–80.
44. Whitaker JP, Cotton EA, Gill P. A comparison of the characteristics of profiles produced with AmpFℓSTR SGM Plus multiplex system for both standard and low copy number (LCN) STR DNA analysis. Forensic Sci Int 2001;123:215–23.
45. Kline MC, Jenkins B, Rodgers S. Non-amplification of a vWA allele. J Forensic Sci 1998;43:250.
46. Walsh S. Commentary on Kline MC, Jenkins B, Rodgers S. Non-amplification of a vWA allele. J Forensic Sci 1998;43:1103–4.
47. Alves C, Amorim A, Gusmao L, Pereira L. VWA STR genotyping: further inconsistencies between Perkin-Elmer and Promega kits. Int J Leg Med 2001;115(2):97–9.
48. Lazaruk K, Wallin J, Holt C, Nguyen T, Walsh PS. Sequence variation in humans and other primates at six short tandem repeat loci used in forensic identity testing. Forensic Sci Int 2001 Jun 1;119(1):1–10.
49. Walsh PS, Fildes NJ, Reynolds R. Sequence analysis and characterization of stutter products at the tetranucleotide repeat locus vWA. Nucleic Acid Res 1996;24(14):2807–12.
50. Schlotterer C, Tautz D. Slippage synthesis of simple sequence DNA. Nucleic Acids Res 1992;20(2):211–5.
51. Brinkmann B, Meyer E, Junge A. Complex mutational events at the HumD21S11 locus. Hum Genet 1996;98(1):60–64.
52. Xiao FX, Gilissen A, Cassiman JJ, Decorte R. Quadraplex fluorescent STR typing system (HumVWA, HumTH01, D21S11 and HPRT) with sequence-defined allelic ladders. Identification of a new allele at D21S11. Forensic Sci Int 1998;94:39–46.
53. Moller A, Meyer E, Brinkmann B. Different types of structural variation in STRs: HumFES/FPS, HumVWA and HumD21S11. Int J Legal Med 1994;106(6):319–23.
54. Tomsey CS, Kurtz M, Flowers B, Fumea J, Giles B, Kucherer S. Case work guidelines and interpretation of short tandem repeat complex mixture analysis. Croat Med J 2001 Jun; 42(3):276–80.
55. Clayton TM, Whitaker JP, Sparkes R, Gill PD. Analysis and interpretation of mixed forensic stains using DNA STR profiling. Forensic Sci Int 1998;91:55–70.
56. Gill P, Sparkes R, Pinchin R, Clayton T, Whittaker J, Buckleton JS. Interpreting simple STR mixtures using allele peak areas. Forensic Sci Int 1998;91:41–53.
57. Fregeau CJ, Bowen Kl, Fourney RM. Validation of highly polymorphic fluorescent multiplex short tandem repeat systems using two generations of DNA sequencers. J Forensic Sci 1999;44:133–66.
58. Crouse CA, Schumm JW. Investigation of species specificity using nine PCR-based-human STR systems. J Forensic Sci 1995;40:952–6.
59. Buel E, Wang G, Schwartz M. PCR amplification of animal DNA with human X-Y amelogenin primers used in gender determination. J Forensic Sci 1995;40:641–4.
60. Nelson MS, Levedakou EN, Matthews JR, Early BE, Freeman DA, Kuhn CA, et al. Detection of a primer-binding site polymorphism for the STR locus D16S539 using the Powerplex 1.1 system and validation of a degenerate primer to correct for the polymorphism. J Forensic Sci 2002;47:345–49.

Additional information and reprint requests:
Edward L. Buse, M.A.
Forensic Scientist III
Orange County Sheriff-Coroner Department
Forensic Science Services–DNA Section
P.O. Box 449
320 N. Flower Street
Santa Ana, CA 92703
E-mail: elb@fss.co.orange.ca.us