J Forensic Sci, September 2003, Vol. 48, No. 5
Paper ID JFS2003043_485
Available online at: www.astm.org

John M. Butler,[1] *Ph.D.; Yin Shen,*[2,3] *Ph.D.; and Bruce R. McCord, Ph.D.*[2]

The Development of Reduced Size STR Amplicons as Tools for Analysis of Degraded DNA*

ABSTRACT: New multiplex PCR sets of commonly used short tandem repeat (STR) markers have been developed to produce PCR products that are reduced in size when compared to standard commercial STR kits. The reduction in size of these amplicons can facilitate the examination and analysis of degraded DNA evidence by improving amplification efficiency. This "miniSTR" approach will permit current forensic practitioners to use STR markers and instrumentation already present in their laboratories and to generate genotyping data that is directly comparable to reference samples and searchable through the FBI's Combined DNA Index System (CODIS) databases. This paper discusses the development of these new primer sets and presents some initial results in the analysis of degraded and aged DNA samples. A method for removal of problematic fluorescent dye artifacts is also described. Comparison studies in over 100 samples have verified that these miniSTR primers can provide fully concordant results to commercial STR kits and can provide improved signal from degraded DNA specimens. These miniplex sets should prove valuable in the analysis of samples where allele dropout and reduced sensitivity of larger STR alleles occurs.

KEYWORDS: forensic science, DNA typing, degraded DNA, STR, miniSTR, dye artifacts, CSF1PO, FGA, TH01, TPOX, vWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, Penta D, Penta E, D2S1338

Short tandem repeat (STR) markers have become the workhorse of forensic DNA typing (1–3). STRs are highly polymorphic and capable of generating typing results from very little material through multiplex amplification using the polymerase chain reaction (PCR). However, in many situations in forensic casework, the DNA samples are far from pristine. If DNA is exposed to the elements or to fire for any length of time, degradation can occur due to bacterial, biochemical, or oxidative processes (4). Additionally, there may be environmental contaminants commingled with the forensic evidence. In such specimens a loss of signal is typically observed with larger-sized STR products. This loss of signal may be the result of either PCR inhibitors present in the forensic evidence or a DNA template that has been fragmented to small sizes.

When either PCR inhibitors or degraded DNA are present, a partial genetic profile with allele and/or complete locus dropout often results (5). The problem is exacerbated when large multiplex PCR reactions are used due to the wide size range of PCR products produced. Commercial multiplex STR kits used in forensic DNA typing can generate amplicons in the size range of 100 to 450 bp (6). In situations where samples are so badly degraded that STR analysis is not possible, sequence analysis of hypervariable regions of mitochondrial DNA (mtDNA) is typically used (7). However, mtDNA testing is a time-consuming process, and, due to the haploid, non-Mendelian nature of mtDNA inheritance, the data are not as powerful for identification purposes as a full 13-locus STR match.

Another approach to trying to recover information from degraded DNA samples is to reduce the size of the PCR products by moving primers in as close as possible to the STR repeat region (8,9). The observation that smaller-sized PCR products from STR markers produced a higher success rate with degraded DNA samples was first reported in 1995 with analysis of victims of the Waco Branch Davidian fire (5). These results have been confirmed by a number of other publications in recent years (8–13).

A major advantage of these smaller STRs, or "miniSTRs," is that database compatibility could be maintained with convicted offender samples processed using commercial STR multiplexes. In addition, these smaller product sizes are more favorably analyzed by alternative technologies such as time-of-flight mass spectrometry (14,15) and rapid microchannel electrophoretic separations (16). Within the forensic community, a core set of STR markers have been developed for utilization in forensic casework, and large databases such as the Combined DNA Index System (CODIS) have been developed incorporating these markers (3,17). Due to their wide use, these STR markers would thus be ideal for candidates in designing new primers that could create smaller PCR products.

This paper describes new PCR primers capable of amplifying the 13 CODIS STRs plus the Penta D, Penta E, and D2S1338 markers

[1] Biotechnology Division, National Institute of Standards and Technology, Gaithersburg, MD.
[2] Department of Chemistry and Biochemistry, Ohio University, Athens, OH.
[3] Current address: OSI Specialties, Crompton Corporation, 3500 South State Route 2, Friendly, WV.

* Contribution of the U.S. National Institute of Standards and Technology. Certain commercial equipment, instruments, and materials are identified in order to specify experimental procedures as completely as possible. In no case does such identification imply a recommendation or endorsement by the National Institute of Standards and Technology nor does it imply that any of the materials, instruments or equipment identified are necessarily the best available for the purpose. Sources of Support: The National Institute of Justice (NIJ) funded the work described here through an interagency agreement with the NIST Office of Law Enforcement Standards to John Butler and Research Grant 2002-IJ-CX-K007 to Bruce McCord.
Received 14 Feb. 2003; and in revised form 21 April 2003; accepted 22 April 2003; published 4 Aug. 2003.

Copyright © 2003 by ASTM International, 100 Barr Harbor Drive, PO Box C700, West Conshohocken, PA 19428-2959.

present in commercially available STR kits. Through redesign of the primer binding sites, each STR marker has been made as small as possible. The newly designed primers are combined into various multiplexed sets, referred to as "miniplexes," usually with a single locus in each dye color. We describe some important issues involved in the development of these new primer sequences and miniplex sets including some initial concordance studies between the miniplexes and commercial STR kits.

Materials and Methods

Reference Sequences and Allele Range Information

Reference sequences for the STR markers CSF1PO, FGA, TH01, TPOX, vWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, Penta D, Penta E, and D2S1338 were obtained from GenBank® (http://www.ncbi.nlm.nih.gov) through the accession numbers listed in Table 1. Approximately 200 to 300 bases on either side of the STR repeat are part of the reference sequences, which may be found at http://www.cstl.nist.gov/biotech/strbase/seq_ref.htm. The observed allele range of information for each locus listed in Table 1 is also available on the STRBase website (18).

Primer Design

PCR primers were designed against each reference sequence using web-based Primer3 (19). Typically, the default parameters were used, and the PCR product size was made as small as possible around the target region that consisted of the STR repeat. Resulting primers (Table 2) were tested for potential interactions with each other using an in-house program described elsewhere (20,21).

PCR Primers and Other Reagents

Fluorescently labeled primers, AmpliTaq Gold® DNA polymerase, and associated buffers were obtained from Applied Biosystems (Foster City, CA). All forward miniSTR primers were labeled with 6FAM™, VIC™, and NED™ dyes (Applied Biosystems), which enabled the use of available matrix standards and reliable color separations on the ABI 310 and 3100 Genetic Analyzers. Unlabeled primers were purchased from Qiagen Operon (Alameda, CA). Commercial multiplexed genotyping kits utilized in these studies included Identifiler™, COfiler™, SGM Plus™, Profiler™, and Profiler Plus™ from Applied Biosystems and Powerplex® 16, *GenePrint*® CTTv, and GammaSTR® Quadruplexes from Promega Corporation (Madison, WI).

Source of DNA Samples

Anonymous human blood samples were purchased from a commercial blood bank named Millennium Biotech, Inc. (Ft. Lauderdale, FL). These samples were approved for use through an institutional review board at the National Institute of Standards and Technology and were extracted using conventional organic extraction methods. A number of aged bloodstains stored at room temperature on untreated Schleicher and Schuell 903 paper were also examined after the DNA from these samples was extracted via a Chelex procedure described previously (22).

PCR Amplification

Amplifications were performed in reaction volumes of 25 μL using a master mix containing 1X GeneAmp® PCR Gold buffer (Applied Biosystems), 1.5 mmol/L $MgCl_2$, 200 μmol/L deoxynucleotide triphosphates (Life Technologies, dNTPs: dATP, dCTP, dGTP, dTTP), and two units of AmpliTaq Gold® DNA polymerase. Primer concentrations were adjusted empirically to balance peak heights within and between dye colors usually starting with 1 μmol/L. Thermal cycling was performed with the GeneAmp 9700 (Applied Biosystems) using the following conditions in 9600-emulation mode (i.e., ramp speeds of 1°C/s):

- 95°C for 10 min
- 28 cycles: 94°C for 1 min, 55°C for 1 min, 72°C for 1 min
- 60°C for 45 min
- 25°C forever

The 72°C extension time is typically increased to 2 min during each cycle and the final extension is lengthened to 60 or 90 min

TABLE 1—*Information on 16 STRs examined in this study. The reference sequences used for primer design are available at http://www.cstl.nist.gov/biotech/strbase/seq_ref.htm. Allele ranges are from Appendix 1 of Forensic DNA Typing (3) or from the kit manuals in the cases of Penta D, Penta E, and D2S1338.*

STR Locus	GenBank Accession	GenBank Allele	Allele Range	Allele Spread	Product Size (STR Kit)	MiniSTR Size	Size Reduction
CSF1PO	X14720	12	6–16	40 bp	280–320 bp (Cofiler)	89–129 bp	191 bp
FGA	M64982	21	12.2–51.2	156 bp	196–352 bp (ProPlus)	125–281 bp	71 bp
TH01	D00269	9	3–14	44 bp	160–204 bp (Cofiler)	51–98 bp	105 bp
TPOX	M68651	11	5–14	36 bp	213–249 bp (Cofiler)	65–101 bp	148 bp
vWA	M25858	18	10–25	60 bp	152–212 bp (ProPlus)	88–148 bp	64 bp
D3S1358	NT_005997	18	8–20	48 bp	97–145 bp (ProPlus)	72–120 bp	25 bp
D5S818	AC008512	11	7–16	36 bp	134–170 bp (ProPlus)	81–117 bp	53 bp
D7S820	AC004848	13	5–15	40 bp	253–293 bp (ProPlus)	136–176 bp	117 bp
D8S1179	AF216671	13	7–19	48 bp	123–171 bp (ProPlus)	86–134 bp	37 bp
D13S317	AL353628	11	5–16	44 bp	193–237 bp (ProPlus)	88–132 bp	105 bp
D16S539	AC024591	11	5–15	40 bp	233–273 bp (Cofiler)	81–121 bp	152 bp
D18S51	AP001534	18	7–27	80 bp	264–344 bp (ProPlus)	113–193 bp	151 bp
D21S11	AP000433	29	24–38.2	58 bp	186–244 bp (ProPlus)	153–211 bp	33 bp
Penta D	AP001752	13	2.2–17	73 bp	376–449 bp (PP16)	94–167 bp	282 bp
Penta E	AC027004	5	5–24	95 bp	379–474 bp (PP16)	80–175 bp	299 bp
D2S1338	AC010136	20	15–28	52 bp	288–340 bp (SGM Plus)	90–142 bp	198 bp

TABLE 2—*MiniSTR primer used in this study. Calculated primer melting temperatures (Tm) are from the Primer3 program using the default values. PCR products containing the 5′ tail on the reverse primer will be 7 bp longer than the values listed in Table 1. The "From repeat" column refers to the distance of the 3′ end of the primer from the STR repeat region. A negative number indicates that the end of the primer is in the repeat region.*

Locus		miniSTR Primer Sequence (5′ to 3′)	Tm (°C)	5′ Tail Added	From Repeat (bp)	Comments
CSF1PO	F	VIC-ACAGTAACTGCCTTCATAGATAG	52.4		14	Partial repeat just 5′ of core repeat
	R	GTGTCAGACCCTGTTCTAAGTA	53.6		6	
FGA	F	6FAM-AAATAAAATTAGGCATATTTACAAGC	55.9		3	
	R	GCTGAGTGATTTGTCTGTAATTG	56.6		23	Partial repeat just 3′ of core repeat
TH01	F	6FAM-CCTGTTCCTCCCTTATTTCCC	61.0		0	
	R	GGGAACACAGACTCCATGGTG	62.8	GTTTCTT	1	
TPOX	F	NED-CTTAGGGAACCCTCACTGAATG	60.0		−4	
	R	GTCCTTGTCAGCGTTTATTTGC	61.0	GTTTCTT	5	
vWA	F	6FAM-AATAATCAGTATGTGACTTGGATTGA	58.1		0	
	R	ATAGGATGGATGGATAGATGGA	57.3		0	
D3S1358	F	NED-CAGAGCAAGACCCTGTCTCAT	59.5		−1	
	R	TCAACAGAGGCTTGCATGTAT	58.4		−1	
D5S818	F	6FAM-GGGTGATTTTCCTCTTTGGT	58.0		4	
	R	AACATTTGTATCTTTATCTGTATCCTTATTTAT	58.3		−5	
D7S820	F	NED-GAACACTTGTCATAGTTTAGAACGAAC	58.9		4	
	R	TCATTGACAGAATTGCACCA	58.6	GTTTCTT	65	PolyT stretch 3′ of core repeat
D8S1179	F	VIC-TTTGTATTTCATGTGTACATTCGTATC	58.5		−4	
	R	ACCTATCCTGTAGATTATTTTCACTGTG	59.4		5	
D13S317	F	NED-TCTGACCCATCTAACGCCTA	58.3		19	
	R	CAGACAGAAAGATAGATAGATGATTGA	57.4	GTTTCTT	2	
D16S539	F	NED-ATACAGACAGACAGACAGGTG	52.5		0	
	R	GCATGTATCTATCATCCATCTCT	55.0		16	
D18S51	F	VIC-TGAGTGACAAATTGAGACCTT	54.8		5	
	R	GTCTTACAATAACAGTTGCTACTATT	52.7		33	Partial repeat just 3′ of core repeat
D21S11	F	VIC-ATTCCCCAAGTGAATTGC	55.8		2	
	R	GGTAGATAGACTGGATAGATAGACGA	56.5		0	
Penta D	F	6FAM-GAGCAAGACACCATCTCAAGAA	59.5		11	
	R	GAAATTTTACATTTATGTTTATGATTCTCT	57.3		19	
Penta E	F	VIC-GGCGACTGAGCAAGACTC	57.1		6	
	R	GGTTATTAATTGAGAAAACTCCTTACA	57.6		4	
D2S1338	F	NED-TGGAAACAGAAATGGCTTGG	61.0		3	
	R	GATTGCAGGAGGGAAGGAAG	61.1		3	

when reamplifying allelic ladders because there are more PCR products to create.

Removal of Residual Dyes Following PCR

Removal of residual dye molecules that generated the so-called "dye blobs" in capillary electrophoresis electropherograms was accomplished with Edge Performa™ DTR Gel Filtration Cartridges from Edge BioSystems (Gaithersburg, MD) following the manufacturer's protocol. Briefly, the Gel Filtration Cartridge was centrifuged for 2 min at 750 × *g* to remove the storage liquid in the cartridge. Then a 15-μL aliquot of the 25-μL PCR reaction was placed in the center of the slanted gel bed surface and centrifuged for 2 min at 750 × *g*. The resulting eluate contained the purified PCR product from which 1 μL may be used for ABI 310 or 3100 sample preparation.

Analysis on ABI 310 (Single Capillary) Genetic Analyzer

The ABI Prism® 310 Genetic Analyzer (Applied Biosystems) was used with filter set F to process the data from the four dyes 6FAM, VIC, NED, and ROX after an appropriate matrix had been created using materials from matrix standard sets DS-32 and DS-33 (Applied Biosystems). Each sample for analysis on the 310 was prepared by adding 1 μL of PCR product to 19 μL of Hi-Di™ formamide (Applied Biosystems) containing 0.75 μL of GS500 ROX size standard (Applied Biosystems). Samples were then placed immediately into the instrument for analysis without heat denaturing or snap cooling the samples prior to running them. Samples were injected for 5 s at 15 000 V and separated at 15 000 V for 24 min with a run temperature of 60°C. Standard electrophoretic conditions were used including 310 Genetic Analyzer POP™-4, 1X Genetic Analyzer Buffer with EDTA, and a 47-cm × 50-μm capillary (Applied Biosystems).

Analysis on ABI 3100 (16-Capillary) Genetic Analyzer

Prior to running any miniplex samples on the ABI 3100, a four-dye matrix was established under the "Z filter" with the dyes 6FAM (blue), VIC (green), NED (yellow), and ROX (red) using matrix standard set DS-30 (Applied Biosystems) and substituting a VIC matrix standard for HEX. Samples were typically prepared with 9 μL of Hi-Di™ formamide, 0.6 μL of GS500 ROX, and with 1 μL of PCR product (either filtered or unfiltered). Again, neither heat denaturation nor snap cooling were performed prior to injection on the ABI 3100. The miniSTR and commercial STR kit samples were run using the default module GeneScan36_POP4 DefaultModule, which performs an electrokinetic injection onto the 16-capillary array for 10 s at 3000 volts. The STR alleles were then separated at 15 000 volts for approximately 30 min with a run temperature of 60°C using the 3100 POP™-4 sieving polymer, 1X Genetic Analyzer Buffer with EDTA, and a 36-cm array (Applied Biosystems). Data from both the ABI 310 and ABI 3100 were analyzed using GeneScan 3.7 and Genotyper 3.7 programs (Applied Biosystems) for a Windows NT platform. In earlier work, Macintosh versions of GeneScan and Genotyper were used.

Generation of Allelic Ladders and Genotyper Macros

Allelic ladders were created with the miniSTRs using a dilution of allelic ladders from the Identifiler™ (Applied Biosystems) or PowerPlex® 16 (Promega) kits. Briefly, a 1:1000 dilution of the kit allelic ladders was prepared, and then 2 μL of these diluted ladders were amplified, either individually or as a multiplex set, using the thermal cycling parameters outlined for the PCR above, except amplified for 15 cycles instead of the standard 28 cycles. Genotyper macros were constructed for each of the miniplex combinations (Table 3) to work with the new allelic ladders.

Results and Discussion

MiniSTR Primer Design for Common STR Markers

The web-based Primer3 program (19) was used to design the smallest possible PCR products for each STR marker of interest. The reference sequence for each STR marker found at http://www.cstl.nist.gov/biotech/strbase/seq_ref.htm was pasted into the DNA sequence window of Primer3, and a target region was defined that

TABLE 3—*Combinations of STR loci examined by dye color. The letters B, G, Y, and R indicate the nominal color of the dye label for each locus: blue, green, yellow, or red. Note that D3S1358, D19S433, and amelogenin, which are already fairly small in their kit PCR product sizes, are not included in the various miniplex combinations. However, new D3S1358 primers are listed in Table 2. The "Big Mini" is a combination of miniplexes 1 and 3 and permits analysis of 6 STRs that do not overlap in size.*

Locus	Identifiler	PP16	Miniplex 1	Miniplex 2	Miniplex 3	Miniplex 4	Miniplex 5	Big Mini
CSF1PO	B	G	G					G
FGA	R	Y			B			B
TH01	G	B	B					B
TPOX	Y	Y	Y					Y
vWA	Y	Y				B		
D5S818	R	G		B				
D7S820	B	G			Y			Y
D8S1179	B	Y		G				
D13S317	G	G				Y		
D16S539	G	G		Y				
D18S51	Y	B				G		
D21S11	B	B			G			G
Penta D		G					B	
Penta E		B					G	
D2S1338	G						Y	
D3S1358	G	B						
D19S433	Y							
Amelogenin	R	Y						

included the STR repeat region and any flanking regions where partial repeats or mononucleotide repeat stretches were observed. The Primer3 software allows a user to easily manipulate parameters such as minimum and maximum primer length, melting temperature, and PCR product sizes during the design process. Typically 80 to 100 bp was used as a starting point for the desired PCR product size, and the other parameters were all left on their default values. Thus, the calculated melting temperature of each primer should be in the 57 to 63°C range and work well with a PCR standard annealing temperature of 55°C. However, several primer pairs are below the 57°C design criteria, such as CSF1PO and D16S539 (see Table 2), because these primers had been shown to work well empirically during previous time-of-flight mass spectrometry work (15).

For the 16 STRs examined, attempts were made to bring primers as close as possible to the STR repeat region. However, the flanking sequences for some of these STRs contain polymorphic nucleotides, partial repeats, mononucleotide repeat stretches, or insertions/deletions that could prevent stable primer annealing. Thus, not all primers could be designed to bind within a few nucleotides of the STR region (see Table 2). For example, the STR marker FGA has a partial repeat and mononucleotide repeat stretch "TTTC TTCC TTTC TTTTTT" immediately downstream of the core STR repeat that impacts the position of the reverse primer. It would be unwise to place a primer here because the GC content would be low and variations may occur next to the FGA compound repeat that would cause disruption of primer annealing. Thus, the 3′ end of the miniSTR FGA-reverse primer is located 23 nucleotides away from the end of the designated core repeat (see Table 2). The primer that is located furthest away from its marker's repeat region is the D7S820-reverse primer. Closer primers have been used in other studies but with instances of allele dropout for some eight and ten alleles (15). A poly-T stretch is found 13 nucleotides downstream of the D7S820 core GATA repeat that has been shown to contain eight, nine, or ten T's (23). Indeed, this polymorphic stretch of T's likely gives rise to the abundance of x.1 and x.3 variant alleles reported for the D7S820 locus in STRBase: http://www.cstl.nist.gov/biotech/strbase/var_d782.htm. Thus, a reverse primer for the D7S820 locus will need to be outside the poly-T stretch in order to capture this variation and permit full concordance with commercial STR kit results.

Every attempt was made to avoid known polymorphic nucleotides in the regions immediately adjacent to the STR repeat that have been discovered during concordance studies with commercial STR kits (2,24,25). For example, the D8S1179 flanking region has a polymorphic nucleotide 55 bases downstream from the repeat (26) that impacts the reverse primer in the Profiler Plus kit and causes allele dropout in some Asian samples (25,26). The miniSTR D8S1179-reverse primer is internal to this sequence polymorphism and thus will not be affected. Likewise, the polymorphic nucleotide 52 bases upstream of the vWA locus that has produced some previously reported allele dropouts (24) is outside of the miniSTR vWA-forward primer binding region and will thus not impact PCR amplification.

However, there is a danger with placing primers close to the repeat region if insertion/deletions occur in the flanking regions of the STR marker but *outside* of the miniSTR primer binding sites. In this situation, it is possible to have full amplification with either primer set but with different allele calls. For example, the D13S317 locus has a four-base deletion of TGTC that is 24 bases downstream from the TATC core (27). The miniSTR D13S317-reverse primer listed in Table 2 is located between the repeat region and the potential deletion sequence, whereas commercial STR kit primers are located outside of the deletion region (27). Thus, in a sample containing an allele with this 4-bp deletion, a heterozygous sample could occur in both the standard and miniSTR amplifications, yet exhibit a difference in size of 4 bp between the two samples due to the presence of this deletion. We have observed this phenomenon with an African American sample that was called a 9,12 at D13S317 using the Profiler Plus kit but a 10,12 with the miniSTR primers. Although we have not sequenced this sample, the "9" allele amplified using the Profiler Plus kit likely contains 10 TATC repeats with a TGTC deletion as described above.

We did not design new primers for the D19S433 locus or the sex-typing marker amelogenin as these amplicons are already small in the widely used kits. In addition, there is minimal size reduction with the miniSTR primers for D3S1358, D21S11, D8S1179, and D5S818 relative to currently available commercial kits. However, these primers may find value in checking for potential allele dropout due to primer binding site mutations, as the new primer positions do not overlap the commercial kit primer locations.

It is interesting to note that it is possible to put the 3′ end of a primer into the repeat region up to two full repeats and still obtain successful amplifications. This fact was discovered while working with new primers for time-of-flight mass spectrometry (14,28). Examples in the current set of primers include the TPOX-forward primer and D8S1179 primers, which are both four bases into the repeat region (one full repeat).

All forward primers were dye-labeled for consistency. The reverse primers were originally obtained as indicated in Table 2. However, several loci were not fully adenylated when the allelic ladders were generated (Fig. 1, top panel). To aid nontemplate addition, a 5′ tail consisting of the 7-base sequence GTTTCTT (29) was added to the reverse primer, enabling more complete adenylation when miniSTR PCR products were amplified (Fig. 1, bottom panel).

Creating Combinations of miniSTR Markers

In order to keep all of the miniSTR loci as short as possible, multiplexes were designed with one STR locus in each dye color in what we refer to as "miniplexes." This is in contrast with commercial kits, which place four or more loci in a single dye color by adjusting the lengths of each set of amplified products to avoid overlap. Table 3 lists the dye colors and locus combinations in the five miniplex sets created with the primers described in Table 2. Each miniplex consists of three different loci, with each locus labeled using a different colored dye. Two loci, D3S1358 and amelogenin, that are included in CODIS were not utilized in the current miniplex systems. These loci are already small in current systems, but could be implemented at a later date along with the non-CODIS locus D19S443 in a sixth miniplex set. Of course, other locus combinations are possible and could be put together with the primers sets described here by changing the dye label or by moving one or both of the primers away from the STR repeat region in order to fit a desired set of loci together. All primers were crosschecked against one another in order to permit all possible combinations in terms of primer compatibility.

As mentioned above, the principle goal in producing these miniplexes is to keep only one STR locus in each dye lane. This permits the location of the new primers to be moved inward without interference from other loci. However, a drawback to this miniSTR approach is that four to five amplifications are required in order to type all of the core loci, and an examiner may not have the luxury

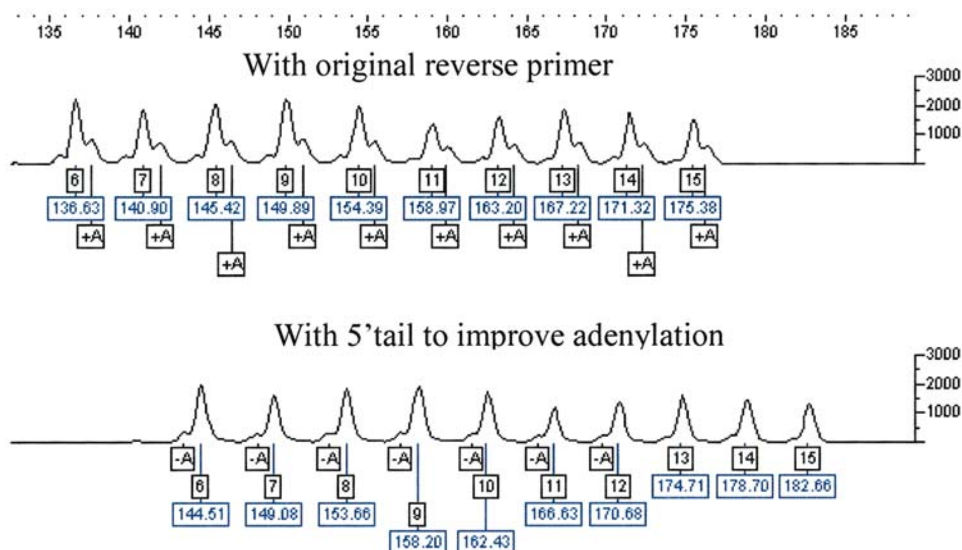


FIG. 1—*D7S820 miniSTR allelic ladder produced by re-amplifying the Identifiler™ ladder with original reverse primer listed in Table 2 (top panel) and with a 5′ tail GTTTCTT on the reverse primer (bottom panel). Peaks are shifted 8 bp in size with the addition of the tail and converted from a primarily "−A" form to a predominant "+A" form with the tail sequence. Samples were amplified side-by-side using a 60-min final extension at 60°C and run under the same electrophoretic conditions on an ABI 310 as indicated in the materials and methods section.*

of obtaining an abundance of degraded DNA from a limited forensic specimen.

Interestingly, it was possible to combine two of the miniplex sets because it is not currently possible to reduce the sizes of miniplex 3 PCR products below 140 bp. Thus we have designated "Big-Mini," a combination of Sets 1 and 3, as an alternative to running these two sets separately (see Table 3). The "Big Mini" alleles from different loci in the same color are more than 20 bp apart so they can easily be distinguished from one another. For the blue loci, the largest TH01 allele (14 repeats) is 98 bp, while the smallest FGA allele reported (12.2 repeats) is 125 bp. In the green loci, the largest CSF1PO allele (16 repeats) is 129 bp, while the smallest D21S11 allele (24 repeats) is 153 bp. For the yellow loci, the largest TPOX allele (14 repeats) is 101 bp, while the smallest D7S820 allele (five repeats) is 136 bp.

Verifying New Primers Yield Concordant Results

Even with all of the efforts taken to avoid known polymorphisms, experimental data are really the only way to determine if allele dropout can occur with a new PCR primer set. Previously undetected polymorphisms may be present in a population dataset and could produce allele dropout if these mutations are present in a primer-binding region.

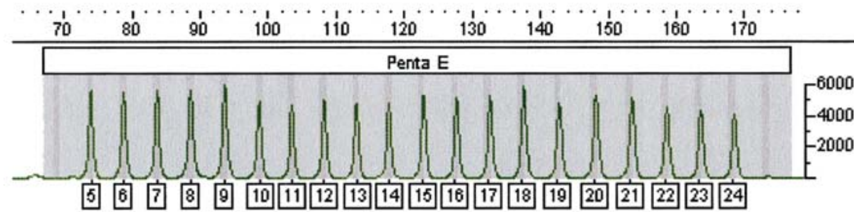
We used two methods for examining the potential of allele dropout with each primer set: re-amplification of STR kit allelic ladders and concordance studies with population samples. First, commercially available allelic ladders containing common alleles for each locus were re-amplified, and the balance of allele peak heights was examined. Figure 2 compares the allelic ladders from the Penta E locus generated with both the new miniSTR primers and the PowerPlex® 16 kit. Although the miniSTR alleles are almost 300 bp smaller than the kit PCR products, allele peak height ratios in the re-amplified miniSTR allelic ladder appear equivalent to those in the kit ladder, demonstrating that the new primers are accurately amplifying each common allele. For example, Penta E alleles 5, 9, and 18 are the highest in signal for both the kit and re-amplified ladders (see Fig. 2).

If a sequence polymorphism occurs within a primer-binding site for a common allele that is present in the allelic ladder, that variant allele will not be re-amplified as well as other alleles in the ladder. The primer-binding site polymorphism would cause an imbalance of peak heights in the re-amplified ladder relative to the original allelic ladder. The fact that re-amplification of commercial allelic ladders with a new set of primers produces similar peak balance helps to verify that no mutations exist under the new primer binding sites in the common alleles found within an allelic ladder. This rapid ladder re-amplification check is possible because the miniSTR primers bind internal to the PCR products produced from commercial kits.

We also conducted concordance studies to look for allelic dropout with our new primer sets by comparing allele calls from kits versus the miniSTR primers. For these studies, more than 50 high-quality DNA samples were examined including some CEPH family samples (3) to demonstrate expected allele inheritance patterns. Our miniSTR results were compared to Profiler™, SGM Plus™, Profiler Plus™, COfiler™, and PowerPlex® 16 kit results, although not all kits were run on all samples. All results were found to be concordant between the commercial kits and the miniSTR markers (data not shown) with the exception of the D13S317 sample noted previously and several D5S818 allele calls noted below.

Figure 3 displays the Genotyper results for allele calls and peak heights of several population samples using the miniSTR D5S818 primers. Note that the expected heterozygote peak ratio of $>70\%$ (3) is not seen in all of these samples. The heterozygote peak height imbalance observed is likely due to a polymorphic nucleotide occurring within a primer-binding site (30). Due to the fact that our new D5S818 primers only give a 25-bp size reduction relative to the Profiler Plus kit (see Table 1), these primers would probably not add much value in testing highly degraded DNA samples.

(A) MiniSTR Re-amplified Ladder



(B) PowerPlex 16 Kit Allelic Ladder

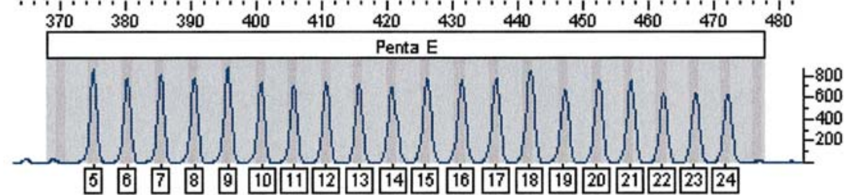


FIG. 2—*Comparison of allelic ladders from Penta E STR locus generated with miniSTR primers* (A) *or from the PowerPlex® 16 kit* (B) *The miniSTR PCR products for Penta E are 300 bp smaller than the kit amplicons. Conditions as described in the materials and methods.*

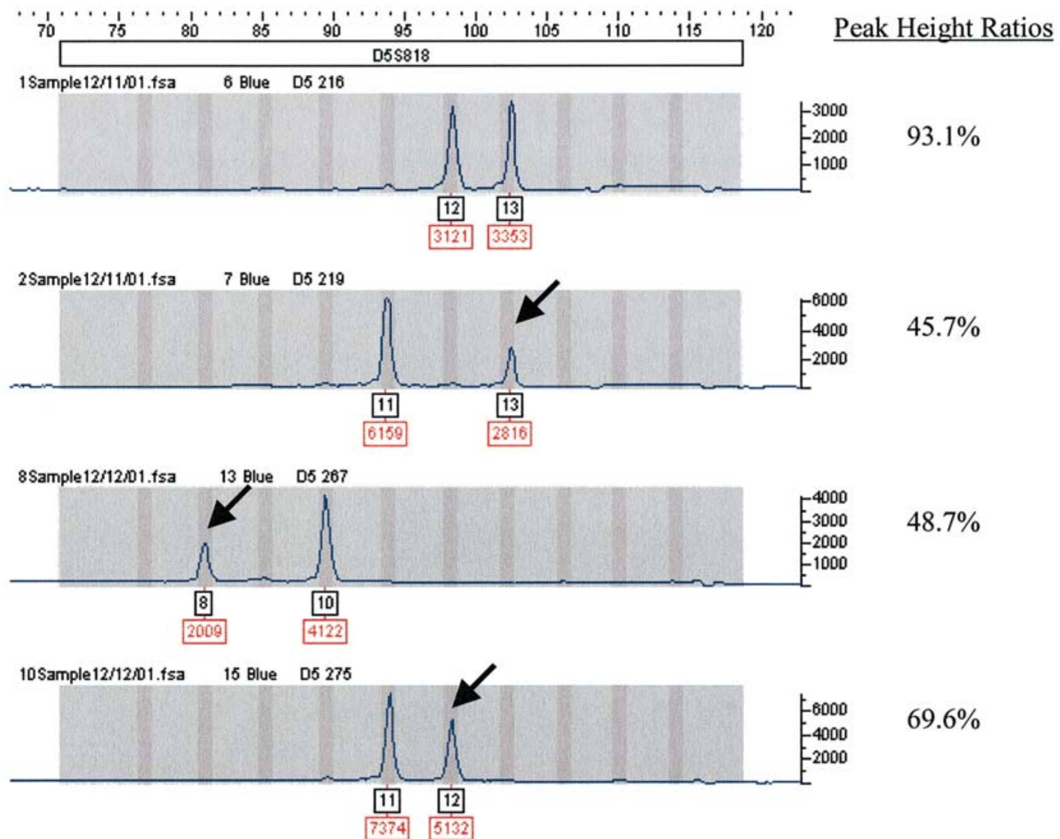


FIG. 3—*D5S818 miniSTR results illustrating a heterozygote peak height imbalance likely due to a polymorphic nucleotide occurring within a primer-binding site. The arrows indicate the low peak height allele. Conditions as in Materials and Methods with ~4 ng of DNA template. Peak labels are allele calls and peak heights in relative fluorescence units.*

Analytical Issues: Dye Blobs and Sizing with the Local Southern Method

When the PCR product size range is reduced to less than 150 bp, several analytical issues arise that are not significant when PCR products range from 100 to 350 bp. One of the biggest problems comes from residual dye molecules or "dye blobs." Residual dye molecules may be left over from the oligonucleotide synthesis process (i.e., were not fully attached) or may result from dyes falling off the primer during the heating and cooling steps of PCR. These artifacts can be virtually filtered out through software or physically removed through a spin column approach as described below.

With the electrokinetic injection process of capillary electrophoresis, any small, charged dye molecules present in the sample will inject onto the capillary along with the labeled PCR products. This residual dye gives rise to colored peaks in the electropherogram that are usually wider and less intense than the true alleles. An analyst can usually distinguish the dye blobs from the true STR alleles based on peak widths. However, as lower amounts of DNA are tested, as is often the case with degraded DNA samples, the ratio of residue dye to amplified product may increase because there is less amplification occurring and the impact of these impurities on the electropherogram may increase. Figure 4 shows the results from a Big Mini amplification where both the STR alleles and the residual dye molecule sizes are noted. Forensic DNA laboratories typically use commercial kits where great effort has been expended to remove the residue dye impurities that exist in research grade experiments such as described here.

Dye blobs can appear with each of the three different dyes used in genotyping: 6-FAM, VIC, and NED. Typically the 6-FAM blue dye blobs can migrate through the capillary at sizes of approximately 62 and 125 bp and thus impact TH01 allele 4 in miniplex 1 (data not shown). A number of VIC residual dye molecules can be present in a sample that impact CSF1PO alleles 7 or 13 and D21S11 allele 32. The major NED residual dye impurities run at around 82 and 89 bp and impact TPOX alleles 9 and 11 among others. These dye blobs may migrate through the capillary slightly differently depending on the electrophoretic conditions in use.

Residual dye molecules may be removed following PCR with a spin column normally used in the removal of dye terminators prior to testing fluorescent DNA sequencing samples. Figure 5 demonstrates the post-PCR removal of all dye blobs with an Edge DTR column following the manufacturer's protocol. While this procedure works well, it is not routinely used as it adds cost and the potential for contamination due to the additional sample manipulation.

Another problem that may arise when using smaller PCR products is the manner in which sizing is performed using an internal lane standard. The default method used by the GeneScan software is Local Southern, which requires two peaks from the internal size standard to be present on either side of a peak being defined. Thus, any STR allele below 75 bp in size must have both the 35 and 50 bp peaks designated from the GS500 ROX size standard in order to be properly defined. These peaks are often not defined by forensic DNA analysts running commercial STR kits, where the smallest allele is greater than 100 bp (thus requiring only the 75 and 100-bp peaks from the size standard), because high levels of PCR primers or primer dimers can be present in a sample and can interfere with the clear designation of the smaller peaks in the GS500 ROX internal size standard. The presence of these inter-

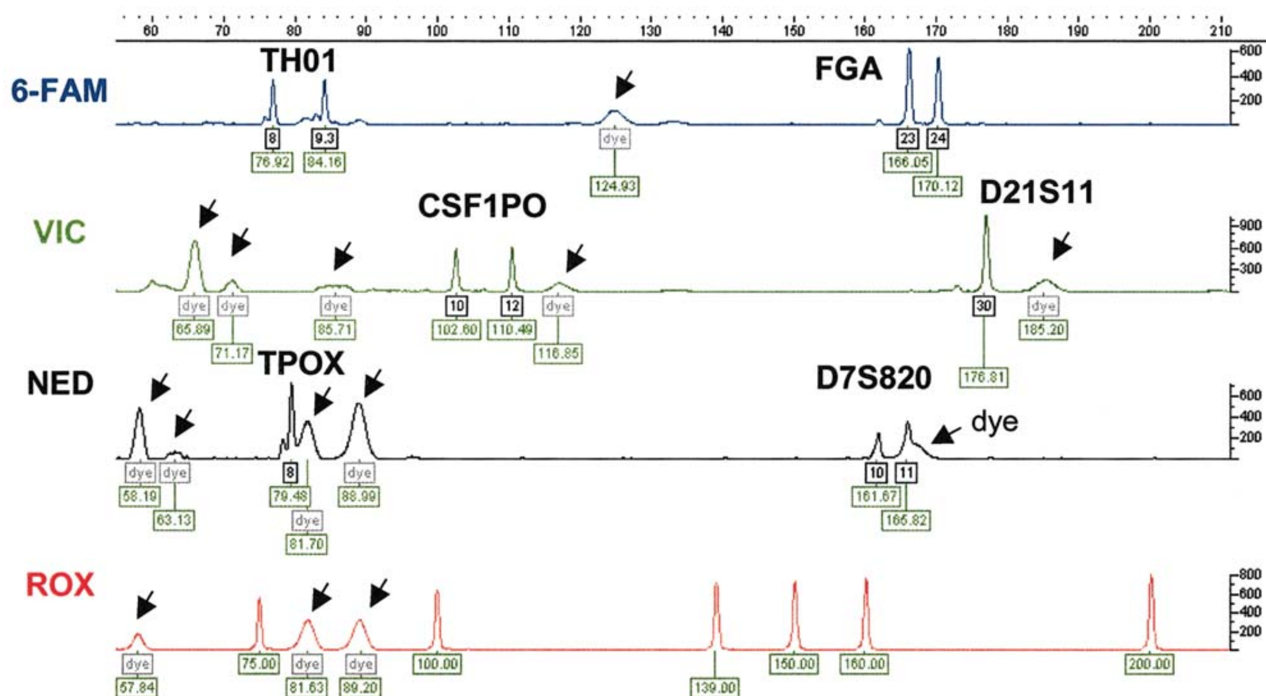


FIG. 4—*Result on ABI 3100 from 0.5 ng 9947A demonstrating residual dyes (arrows) that can cause interferences with particular miniSTR alleles, in this case from the Big Mini assay. Residual dye peak sizes are consistent in size across runs under the same electrophoretic conditions and are always broader in shape than true alleles. The residual dye signals in the red (ROX) channel are bleed-through from the yellow (NED) channel. Peak labels are allele calls and peak sizes in bases relative to GS500 ROX size standard.*

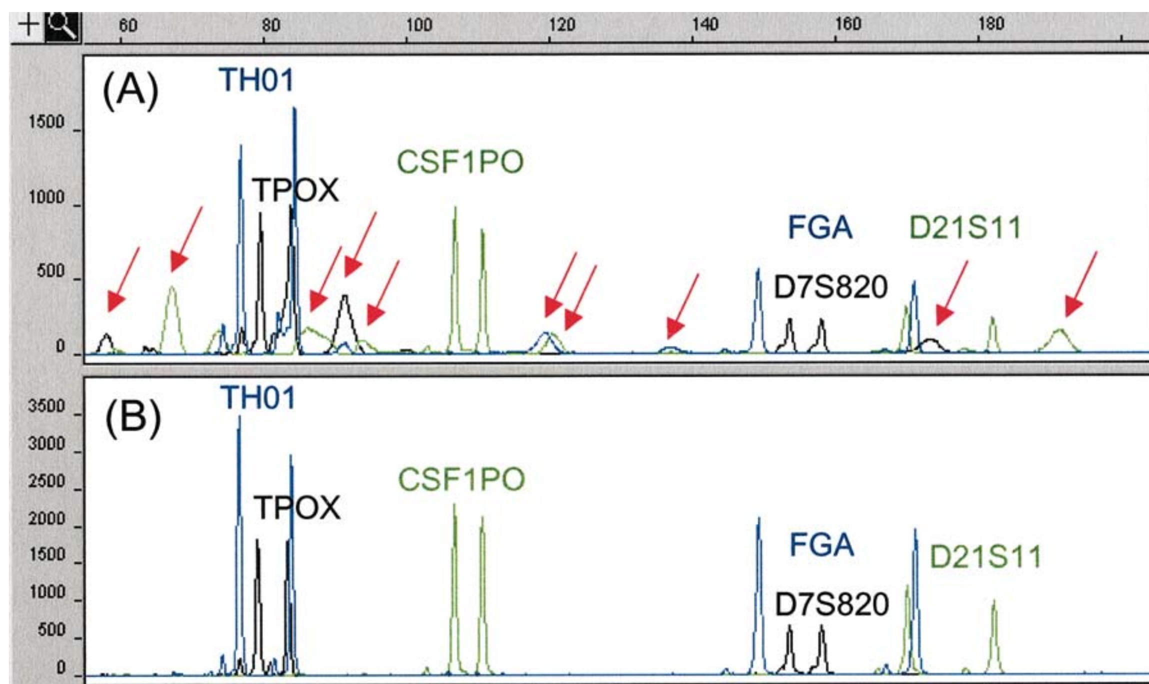


FIG. 5—*Comparison of unpurified* (A) *versus Edge-column-filtered* (B) *PCR products from the Big Mini assay. Samples were amplified from approximately 1 ng of DNA template extracted from a 15-year-old blood stain and were run on an ABI 310 as specified in the Materials and Methods section. Arrows indicate the broad-shaped residual dye peaks that are removed with sample purification.*

fering peaks makes it difficult to properly size the miniplex STRs with Local Southern methods. While it might be possible to produce an internal size standard with peaks just larger than the interfering primer dimers, such a standard is not presently commercially available.

A better approach to sizing small alleles is Global Southern, where the STR alleles are fit to a regression line generated from all peaks in the size standard. There are certain advantages to this approach since Global Southern sizing is less prone to effects resulting from migration shifts of individual peaks in the sizing ladder (31). However, it is still necessary to make certain the minimum allele size is larger than the smallest peak in the size standard and to be aware that size variation will be somewhat greater at the ends of the regression line than in the center. We will address this issue more fully in future validation studies. At present we anticipate that the smallest miniSTR loci, TH01 and TPOX, will be most impacted by this sizing issue.

Performance with Degraded DNA Samples

To evaluate the newly developed miniSTR systems, tests were conducted on partially degraded DNA samples. A set of 92 Chelex-extracted aged bloodstains that had been stored at room temperature for 14 or 15 years prior to DNA extraction and typing was examined with both the PowerPlex® 16 kit and miniplexes 1, 2, and 3. The top panel in Fig. 6 demonstrates that the larger loci in the PowerPlex® 16 kit, such as CSF1PO and Penta D (inset), are greatly reduced in signal when these aged bloodstains were tested. The bottom panel of Fig. 6 shows improved signals for CSF1PO alleles with amplification using the miniSTR primers because the allele sizes are reduced by almost 200 bp.

The current approach of 28-cycle PCR works well with the primers in Table 2 down to approximately 0.5 ng of template DNA. Increasing the number of PCR cycles or adding more DNA polymerase may reliably type lower levels of DNA, but it is important to keep in mind the forensic issues involved when amplifying low-copy number DNA (32).

Conclusions

This paper outlines the initial efforts made to define new primer sets for the commonly used STR markers in human identity testing. We have developed these miniSTR primer sets into several miniplexes containing between three and six loci. Comparison studies in over 100 samples have verified that these miniSTR primers can provide fully concordant results to commercial STR kits and can provide improved signals from degraded DNA specimens. We anticipate that these new primer sets will prove useful in a variety of scenarios in which the current set of STR multiplex kits fail. In particular, the miniplex sets should prove valuable in the analysis of samples where allele dropout and reduced sensitivity of larger alleles occurs. The miniplexes also provide a potential alternative to mitochondrial sequencing in the forensic analysis of degraded DNA. Because of the difference in primer sequence, they also provide a check for the presence of allele dropout due to problems with primer binding of standard STR sets. Lastly, the miniplexes have great potential in application to high-speed microfluidic and mass spectrometric approaches. In general miniSTRs offer a new potential tool for recovering useful information from samples that generated partial profiles with commercial STR kits that can be completely compatible with reference data.

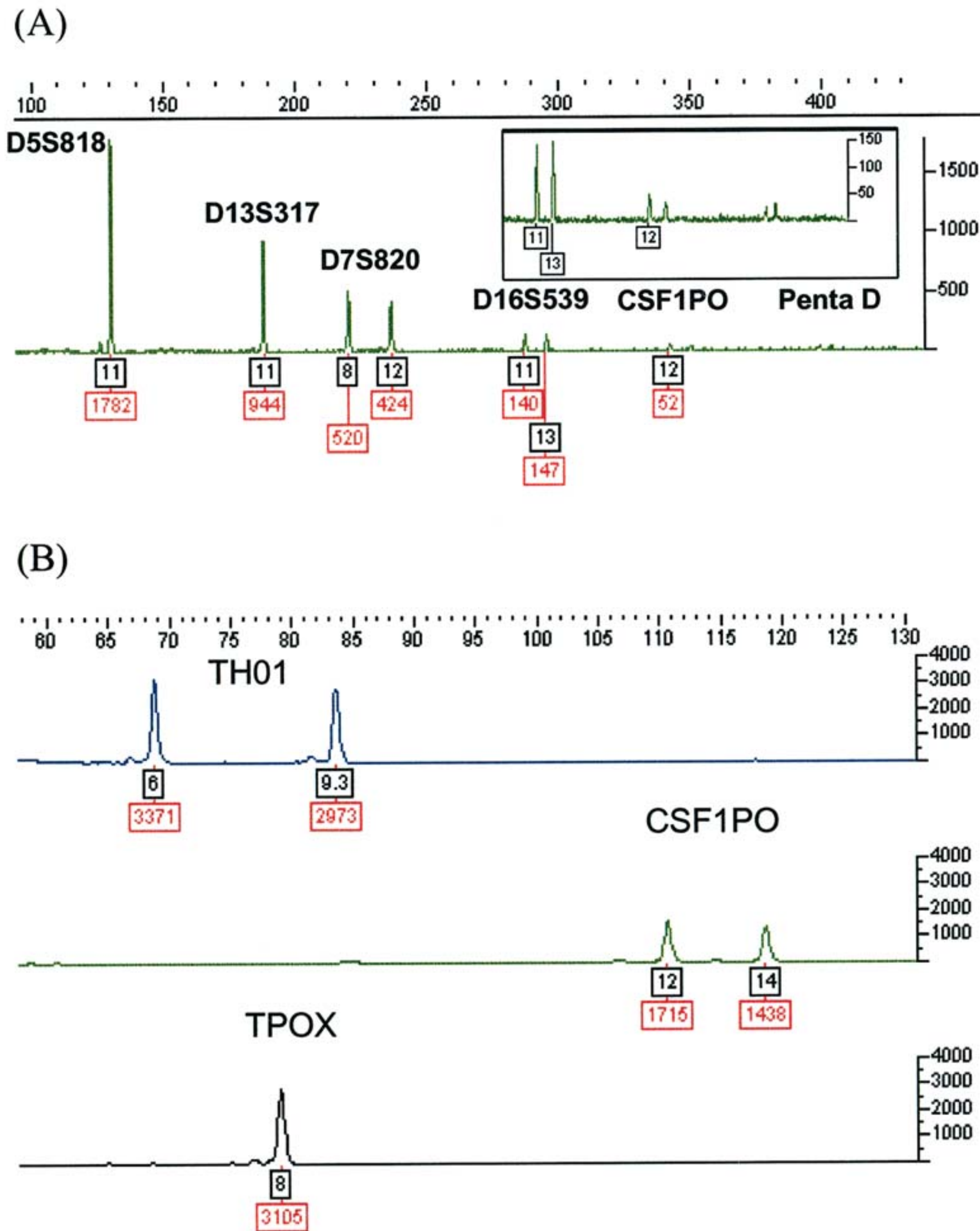


FIG. 6—*Comparison of STR allele peak heights obtained from PowerPlex® 16* (A) *and miniplex 1 loci* (B) *with equivalent amounts of the same aged bloodstain sample. Peak labels are allele calls and peak heights are in relative fluorescence units (RFUs). The second allele for CSF1PO with the PowerPlex® 16 kit is below the 50 RFU peak detection threshold and therefore not labeled by the software.*

Acknowledgments

The authors express appreciation for the encouragement of Bob Shaler and Lisa Forman in this work and the technical assistance of Peter Vallone, Margaret Kline, Kerry Opel, and Denise Chung.

References

1. Lygo JE, Johnson PE, Holdaway DJ, Woodroffe S, Whitaker JP, Clayton TM, et al. The validation of short tandem repeat (STR) loci for use in forensic casework. Int J Legal Med 1994;107:77–89.
2. 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(6):1277–86.
3. Butler JM. Forensic DNA typing: biology and technology behind STR markers. London: Academic Press, 2001.
4. Bar W, Kratzer A, Machler M, Schmid W. Postmortem stability of DNA. Forensic Sci Int 1988;39(1):59–70.
5. Whitaker JP, Clayton TM, Urquhart AJ, Millican ES, Downes TJ, Kimpton CP, et al. Short tandem repeat typing of bodies from a mass disaster: high success rate and characteristic amplification patterns in highly degraded samples. BioTechniques 1995;18(4):670–7.
6. Krenke BE, Tereba A, Anderson SJ, Buel E, Culhane S, Finis CJ, et al. Validation of a 16-locus fluorescent multiplex system. J Forensic Sci 2002;47(4):773–85.
7. Butler JM, Levin BC. Forensic applications of mitochondrial DNA. Trends Biotech 1998;16:158–62.
8. Wiegand P, Kleiber M. Less is more—length reduction of STR amplicons using redesigned primers. Int J Legal Med 2001;114(4–5):285–7.
9. Tsukada K, Takayanagi K, Asamura H, Ota M, Fukushima H. Multiplex short tandem repeat typing in degraded samples using newly designed primers for the TH01, TPOX, CSF1PO, and vWA loci. Legal Med 2002;4:239–45.

10. Sparkes R, Kimpton CP, Watson S, Oldroyd NJ, Clayton TM, Barnett L, et al. The validation of a 7-locus multiplex STR test for use in forensic casework: (I) mixtures, ageing, degradation and species studies. Int J Legal Med 1996;109:186–94.
11. Yoshida K, Sekiguchi K, Kasai K, Sato H, Seta S, Sensabaugh GF. Evaluation of new primers for CSF1PO. Int J Legal Med 1997;110(1):36–8.
12. Szibor R, Kayser M, Roewer L. Identification of the human Y-chromosomal microsatellite locus DYS19 from degraded DNA. Am J Forensic Med Pathol 2000;21(3):252–4.
13. Hellmann A, Rohleder U, Schmitter H, Wittig M. STR typing of human telogen hairs—a new approach. Int J Legal Med 2001;114(4–5): 269–73.
14. Butler JM, Li J, Shaler TA, Monforte JA, Becker CH. Reliable genotyping of short tandem repeat loci without an allelic ladder using time-of-flight mass spectrometry. Int J Legal Med 1998;112(1):45–9.
15. Butler JM, Becker CH. Improved analysis of DNA short tandem repeats with time-of-flight mass spectrometry. Washington, DC: Science and Technology Research Report, National Institute of Justice, 2001 Oct.; see http://www.ojp.usdoj.gov/nij/pubs-sum/188292.htm.
16. Schmalzing D, Koutny L, Adourian A, Belgrader P, Matsudaira P, Ehrlich D. DNA typing in thirty seconds with a microfabricated device. Proc Natl Acad Sci USA 1997;94:10273–8.
17. Budowle B, Moretti TR, Niezgoda SJ, Brown BL. CODIS and PCR-based short tandem repeat loci: law enforcement tools. Proceedings of the Second European Symposium on Human Identification, 1998;73–88; see http://www.promega.com/geneticidproc/eusymp2proc/17.pdf.
18. Ruitberg, CM, Reeder, DJ, Butler, JM. STRBase: a short tandem repeat DNA database for the human identity testing community, Nucleic Acids Res 2001;29:320–2; see http://www.cstl.nist.gov/biotech/strbase.
19. Rozen S, Skaletsky HJ. Primer3 (primer 3_www.cgi v 0.2); 1998. Code available at http://www.genome.wi.mit.edu/cgi-bin/primer/primer3_www.cgi.
20. Butler JM, Ruitberg, CM, Vallone, PM. Capillary electrophoresis as a tool for optimization of multiplex PCR reactions. Fresenius J Anal Chem 2001;369:200–5.
21. 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 375:333–42.
22. Kline MC, Duewer DL, Redman JW, Butler JM, Boyer DA. Polymerase chain reaction amplification of DNA from aged blood stains: quantitative evaluation of the "suitability for purpose" of four filter papers as archival media. Anal Chem 2002;74(8):1863–9.
23. Egyed B, Furedi S, Angyal M, Boutrand L, Vandenberghe A, Woller J, Padar Z. Analysis of eight STR loci in two Hungarian populations. Forensic Sci Int 2000;113(1–3):25–7.
24. Alves C, Amorim A, Gusmao L, Pereira L. VWA STR genotyping: further inconsistencies between Perkin-Elmer and Promega kits. Int J Legal Med 2001;115(2):97–9.
25. Budowle B, Masibay A, Anderson SJ, Barna C, Biega L, Brenneke S, et al. STR primer concordance study. Forensic Sci Int 2001;124(1):47–54.
26. Han GR, Song ES, Hwang JJ. Non-amplification of an allele of the D8S1179 locus due to a point mutation. Int J Legal Med 2001; 115(1):45–7.
27. Boutrand L, Egyed B, Furedi S, Mommers N, Mertens G, Vandenberghe A. Variations in primer sequences are the origin of allele dropout at loci D13S317 and CD4. Int J Legal Med 2001;114(4–5):295–7.
28. Butler JM, Li J, Monforte JA, Becker CH. GeneTrace Systems Inc., assignee. DNA typing by mass spectrometry with polymorphic DNA repeat markers. U.S. Patent 6,090,558, 2000 July 18.
29. Brownstein MJ, Carpten JD, Smith JR. Modulation of non-templated nucleotide addition by Taq DNA polymerase: primer modifications that facilitate genotyping. BioTechniques 1996;20(6):1004–10.
30. Alves C, Gusmao L, Pereira L, Amorim A. Multiplex STR genotyping: comparison study, population data and new sequence information. Poster P24 at 19th International Congress of the International Society for Forensic Genetics, Munster, Germany, August 2001.
31. Hartzell B, Graham K, McCord B. Response of short tandem repeat systems to temperature and sizing methods. Forensic Sci Int 2003;133: 228–34.
32. Gill P, Whitaker J, Flaxman C, Brown N, Buckleton J. An investigation of the rigor of interpretation rules for STRs derived from less than 100 pg of DNA. Forensic Sci Int 2000;112(1):17–40.

Additional information and reprint requests:
John Butler, Ph.D.
Research Chemist
National Institute of Standards and Technology
100 Bureau Dr., Mail Stop 8311
Gaithersburg, MD 20899-8311
E-mail: john.butler@nist.gov