J Forensic Sci, Jan. 2004, Vol. 49, No. 1
Paper ID JFS2003209_491
Available online at: www.astm.org

Erin K. Hanson,[1] *M.S. and Jack Ballantyne,*[1,2,3] *Ph.D.*

A Highly Discriminating 21 Locus Y-STR "Megaplex" System Designed to Augment the Minimal Haplotype Loci for Forensic Casework*

ABSTRACT: In order to increase significantly the discriminatory potential of Y-STR systems available to the forensic community, we have developed and validated a 21-locus Y-STR multiplex system. Since the system was designed specifically to augment the European Y chromosome typing community's "minimal haplotype" Y-STR set (MHL) for forensic casework, it contains a novel constellation of markers not contained therein. The system, which we refer to as Multiplex IV (MPIV), permits the co-amplification of DYS 443, DYS 444, DYS 445, DYS 447, DYS 448, DYS 449, DYS 452, DYS 453, DYS 454, DYS 455, DYS 456, DYS 458, DYS 463, DYS 464, DYS 468, DYS 484, DYS 522, DYS 527, DYS 531, DYS 557, and DYS 588. Although the multiplex contains 21 Y-STR loci, of which one is bi-local and one is tetra-local, there are actually 25 sites exhibiting allelic variation, and this has prompted us to use the descriptor "megaplex" to describe the system.

This report describes a number of performance checks that were employed to characterize the system including sensitivity, specificity, discriminatory capacity, and nonprobative casework studies. Although 1 ng of male DNA was found to be the optimal amount of input template, a complete 21-locus profile was obtained with as little as 50 pg of male DNA (i.e., ~8 to 9 diploid cells). The specificity of the system was demonstrated by the lack of significant female DNA derived artifacts when tested using either 300 ng of female DNA alone or an admixture of male/female DNA in which the female component was present in a 100-fold excess. The ability of the system to determine the number of male donors was demonstrated by testing different admixtures of DNA at different ratios from two male donors. Cervicovaginal samples taken up to 48 h post coitus yielded a complete 21-locus Y-STR profile of the semen donor, thus confirming the potential utility of the system for forensic casework. Preliminary estimates of the gene diversity (h) of the individual loci for the Caucasian and African-American population indicated that 15 of the 21 loci possessed an h of $\geq$0.5 in at least one population. Multi-locus haplotype analysis revealed that the 21-plex system could augment the use of the minimal haplotype loci and increase significantly the discriminatory capacity of Y-STR analysis.

KEYWORDS: forensic science, Y-chromosome markers, Y-STR multiplex, Y-STR megaplex, DYS 443, DYS 444, DYS 445, DYS 447, DYS 448, DYS 449, DYS 452, DYS 453, DYS 454, DYS 455, DYS 456, DYS 458, DYS 463, DYS 464, DYS 468, DYS 484, DYS 522, DYS 527, DYS 531, DYS 557, DYS 588, post-coital cervicovaginal swabs, MPIV, minimal haplotype loci (MHL)

The unique biology of the genetic markers present on the non-recombining region of the Y chromosome has resulted in their widespread use in determining patrilineal relationships within and between populations to aid in the understanding of human migration and evolution (1,2). More recently, subsets of such genetic markers, namely Y-chromosome microsatellites or short tandem repeats (STRs), have been incorporated into a variety of multiplex PCR assays for potential forensic casework applications (3–7). Their intended use is not to supplant the current battery of autosomal STR loci, but to apply them to certain defined casework situations whereby the traditional autosomal loci would not be expected to yield sufficient probative information. For example, Y-STRs can be particularly useful when trying to determine the genetic profile of the male donor in a male/female DNA admixture when the female DNA component is present in vast excess (e.g., $\geq$100 x) and when traditional autosomal STR analysis fails or is not possible (8–12).

Autosomal STR analysis may not be possible if the sample contains a male/female mixture of body fluids other than semen, such as in saliva/saliva mixtures, saliva/vaginal secretion mixtures, or fingernail scrapings with cells from the perpetrator. In these types of samples a differential extraction to separate the male and female cells is not possible (unlike sperm-containing samples), and the male component will not be detectable with the PCR-based autosomal STR systems routinely used. This is due to the kinetics of the PCR process itself, which does not permit minor components to be detected at low levels (i.e., $\leq$1/20) because of titration of critical reagents by the major component (11). Autosomal STR analysis may also fail with some semen containing samples in which the sperm are present in very low copy number or are present in an extremely fragile state, such as in extended interval (i.e., >48 h) post-coital samples. Differential extraction of these particular samples may yield no profile from the male donor due to a combination of premature lysis of the sperm's cellular constituents into the non-sperm fraction and to sperm loss during the physical manipulations required of the DNA isolation process.

A set of nine Y-STR loci, commonly referred to as the "minimal haplotype loci" set (MHL), were recommended for forensic use by the European Y chromosome typing community (2). The MHL loci have proved to be a particularly robust set of genetic markers and have been successfully employed in casework analysis (8–12). Despite their utility, however, additional Y-STR loci are required in order to improve the discriminatory capacity of the MHL. For example, MHL analysis of 1,705 individuals (599 African Americans,

[1] Graduate Program in Forensic Biochemistry, University of Central Florida, PO Box 162366, Orlando, FL.
[2] Department of Chemistry, University of Central Florida, PO Box 162366, Orlando, FL.
[3] National Center for Forensic Science, PO Box 162367, Orlando, FL.

* Oral presentation at the 55th Annual American Academy of Forensic Science Conference, Chicago, IL, February 17–22, 2003, as "Highly Discriminating Y-STR Multiplexes to Permit the Determination of 42-Locus Male Haplotypes." Project funded by NIJ.

Received 17 June 2002; and in revised form 19 Sept. 2003; accepted 19 Sept. 2003; published 23 Dec. 2003.

628 European Americans, and 478 Hispanics) by the custodians of an online MHL U.S. database demonstrated that these individuals collectively possessed 1,116 different haplotypes, a discriminatory capacity of approximately 65%, where discriminatory capacity equals the number of individuals/number of different haplotypes (3). Approximately 20% ($n = 220$) of the 1,116 haplotypes were observed more than once, with the most common haplotype being found in 3% ($n = 53$) of individuals.

Although more than 150 STR loci have been described on the Y chromosome, a limited number of additional non-MHL have been incorporated into multiplex systems (13–16). These multiplex systems, which normally incorporate all or most of the MHL, have undergone varying degrees of performance checks to ascertain their suitability for routine use. Yet even fewer have undergone the extensive developmental validation exercises required by national standards prior to their use in forensic casework.

In the present work, we have designed, developed, and tested a novel 21-locus Y-STR multiplex system that, although not incorporating any of the MHL, was designed to augment their use by providing significant additional discriminatory power. This report describes the results of a number of performance checks used to characterize the system including sensitivity, specificity, nonprobative casework, and population studies.

Methods

Preparation of Body Fluid Stains

Body fluids were collected from volunteers using procedures approved by the University of Central Florida's Institutional Review Board. Buccal samples were collected from donors using sterile swabs by swabbing the inside of the donor's mouth. Neat semen samples were provided in sealed plastic tubes and were stored frozen until they were dried onto sterile cotton swabs. Post-coital cervicovaginal swabs were taken from female participants at various specified time periods subsequent to sexual intercourse. Population samples for gene diversity and multi-locus discriminatory capacity studies were obtained from the Virginia Division of Forensic Science (bloodstains), Richmond, VA, Connecticut State Police Forensic Science Laboratory (DNA extracts in ethanol), Waterbury, CT, the South Dakota Crime Laboratory (bloodstains), and the Minnesota Bureau of Criminal Apprehension (DNA extracts). All samples were stored at −20°C until needed.

DNA Isolation and Purification

DNA was extracted from the buccal swabs, semen swabs, and the post-coital cervicovaginal samples using a standard phenol:chloroform method (17). Stains or swabs were cut into small pieces and placed into a Spin-Ease tube (Gibco-BRL, Grand Island, NY). The tubes were incubated overnight in a 56°C water bath using 400 μL of DNA Extraction Buffer (100 m*M* NaCl, 10 m*M* Tris-HCl, pH 8.0, 25 m*M* EDTA, 0.5% SDS), 0.1 mg/mL Proteinase K, and 10% 0.39 *M* DTT (added to semen containing samples). After the overnight incubation, swab or stain fabric was placed into a Spin-Ease basket, the basket inserted back into the original tube, and the samples centrifuged at 14,000 *g* for 5 min to remove the absorbed fluid from the swab material. A volume of phenol/chloroform/isoamyl alcohol equal to the volume of the crude extract was added and vigorously intermixed by shaking. The aqueous layer containing the DNA was removed. Precipitation of the DNA was accomplished by the addition of cold absolute ethanol (two and a half times the volume of the aqueous layer extract) and allowed to progress overnight at −20°C. The DNA was pelleted by centrifugation, washed twice using 70% ethanol, and resolubilized with 100 μL of TE^{-4} (10 m*M* Tris-HCl, 0.1 m*M* EDTA, pH 7.5) overnight at 56°C.

Differential Cell Lysis for the Recovery of Sperm DNA

Sperm and nonsperm cells were separated using a standard differential lysis protocol, with minor modifications (18). Post-coital cervicovaginal swabs were incubated overnight at 37°C in 400 μL of DNA extraction buffer (100 m*M* NaCl, 10 m*M* Tris-HCl, 25 m*M* EDTA, 0.5% SDS) and 0.1 mg/mL Proteinase K. Swab remnants were removed to a Spin-Ease basket, the basket inserted back into the original tube, and centrifuged at 14,000 *g* for 5 min. The resulting supernatant, containing the nonsperm DNA fraction, was removed into a separate tube for further analysis. The sperm pellet was resuspended in 400 μL of DNA extraction buffer, 0.1 mg/mL Proteinase K, and 40 μL of 0.39 *M* DTT and incubated for 1 h at 56°C. DNA from both the sperm and nonsperm fractions was isolated and purified using the phenol : chloroform method described above.

DNA Isolation and Purification of Dried Blood Samples

The dried bloodstains were incubated overnight at 56°C in 400 μL of DNA extraction buffer (100 m*M* NaCl, 10 m*M* Tris-HCl, 25 m*M* EDTA, 0.5% SDS) and 0.1 mg/mL Proteinase K. The swab pieces were placed into a Spin Ease basket and centrifuged at 14,000 *g* for 5 min. An equal volume of phenol/chloroform/isoamyl alcohol was added to the crude extract. The aqueous phase extracts containing the DNA were purified using Centricon 100™ concentrators (Millipore, Bedford, MA) according to the manufacturer's instructions.

DNA Quantitation

DNA was quantitated using ethidium bromide induced fluorescence on a 1% agarose yield gel, using a reference set of DNA standards of known concentration (19).

Standard PCR Conditions

Standard Reaction—Optimization of the multiplex system resulted in a set of standard conditions. The 25-μL reaction mix contained: 3 ng of template DNA, 0.07 to 0.50 μ*M* primers (see below), 250 μ*M* dNTPs, 1X PCR Buffer II (10 m*M* Tris-HCl, pH 8.3, 50 m*M* KCl), 1.5 m*M* $MgCl_2$, and 2.0 units of AmpliTaq Gold DNA Polymerase (Applied Biosystems, Foster City, CA).

Primers—Primer sequences were obtained from published sources, the Human Genome Database (20), or re-designed using Oligo 6 Primer Analysis Software (Lifescience Software Resource, Long Lake, MN). The forward or reverse primer at each locus was labeled with a fluorescent phosphoramidite dye (Applied Biosystems or Invitrogen, Grand Island, NY). Table 1 lists the primer sequences and associated dye labels. The primer concentrations were as follows: DYS 588–0.072 μ*M*; DYS 455–0.16 μ*M*; DYS 447–0.12 μ*M*; DYS 448–0.36 μ*M*; DYS 527–0.20 μ*M*; DYS 458–0.08 μ*M*; DYS 484–0.16 μ*M*; DYS 557–0.18 μ*M*; DYS 445–0.10 μ*M*; DYS 449–0.12 μ*M*; DYS 453–0.18 μ*M*; DYS 454–0.36 μ*M*; DYS 464–0.14 μ*M*; DYS 444–0.18 μ*M*; DYS 522–0.50 μ*M*; DYS 531–0.12 μ*M*; DYS 456–0.12 μ*M*; DYS 452–0.45 μ*M*; DYS 463–0.18 μ*M*; DYS 443–0.12 μ*M*; DYS 468–0.38 μ*M*.

Cycling Conditions—(1) 95°C for 11 min, (2) 32 cycles: 96°C for 30 s, 50°C for 45 s, 72°C for 45 s, and (3) final extension at 72°C for 120 min.

TABLE 1—*MPIV characteristics.*

Locus	Repeat Motif	PCR Product Size Range (bp)	GDB Accession Number	Primer Sequences
DYS 588	Penta	140–167	11503994	F:FAM-GAATGCAGAACCCTCAAGGA[20] R: AGCCTGGGTGACAGAAACAC[20]
DYS 455	Tetra	167–184	11498125	F:FAM-GGGGTGGAAACGAGTGTT[20] R: ATCTGAGCCGAGAGAATGATA [20]
DYS 447	Penta	204–234	10843977	F:GGTCACAGCATGGCTTGGTT [20] R:FAM-GGGCTTGCTTTGCGTTATCT [20]
DYS 448	Hexa	284–309	10877524	F:FAM-TGTCAAAGAGCTTCAATGGAGA[20] R:TCTTCCTTAACGTGAATTTCCTC [20]
DYS 527	Tetra	330–367	11503872	F:FAM-TCGCAAACATAGCACTTCAG [20] R: TTCTAGGAAGATTAGCCACAACA[20]
DYS 458	Tetra	110–133	11498131	F:VIC-AGCAACAGGAATGAAACTCCAAT[20] R: CCACCACGCCCACCCTCC[20]
DYS 484	Tri	143–155	11503786	F:VIC-CCTATCATCCGCATGGACTT[20] R: CCTGGTTGACAAAGCCAGAT [20]
DYS 557	Tetra	179–206	11503932	F:VIC-TTTTCTGTGCCAAGCCTACA [20] R: TCTAATGCACCTTGAGGGATG [20]
DYS 445	Tetra	250–271	10807129	F:VIC-GAGCTGAGATTATGCCACCAAAA[20] R: AGTTAAGAGCCCCACCTTCCTG[20]
DYS 449	Tetra	345–382	10879367	F:VIC-TGGAGTCTCTCAAGCCTGTTCTA [20] R: CCTGGAAGTGGAGTTTGCTGT[20]
DYS 453	Tetra	122–131	11498119	F:NED-GGGTAACAGAACAAGACAGT[20] R: CTAAAAGTATGGATATTCTTCG[20]
DYS 454	Tetra	195–204	11498123	F:NED-GACATGTAGCTCTTCACTTCAC[20] R: GACTGACCTCACATTGTTGTTA[20]
DYS 464	Tetra	249–283	11499416	F:NED-TTACGAGCTTTGGGCTATG [20] R: CCTGGGTAACAGAGAGACTCTT[20]
DYS 444	Tetra	289–312	10807128	F: TCTAAGGGATCCAAAGGCAGAA [20] R:NED-GTGTGAACCATTTGGCATGTTTA[20]
DYS 522	Tetra	347–364	11503862	F:NED-CCTTTGAAATCATTCATAATGC [20] R: TCATAAACAGAGGGTTCTGG[20]
DYS 531	Tetra	108–117	11503880	F:PET-GACCCACTGGCATTCAAATC [20] R: TGCTCCCTTTCTTTGTAGACG[20]
DYS 456	Tetra	140–164	11498127	F:PET-CCCATCAACTCAGCCCAAAAC [20] R: GGACCTTGTGATAATGTAAGATA[20]
DYS 452	Penta	182–218	11498121	F:PET:TTTATTATACTCAGCTAATTAATTGGTT[20] R:GTGGTGTTCTGATGAGGATAAT[20]
DYS 463	Hexa	222–258	11499418	F:PET- AATTCTAGGTTTGACCAAAGACA [20] R: ATGAGGTTGTGTGACTTGACTG[20]
DYS 443	Tetra	289–311	10807127	F:PET-TTTCATTGGCCACCTGACATTAC[20] R: CGCTTCCATTTACACTTCCTGTG[20]
DYS 468	Tri	354–385	11503754	F:PET-GGGAGTTCCAAACTTTTTCACA R: GGGGGAAGATGACAATGATG

PCR Product Detection—Amplified fragments were detected with the ABI Prism 310 Genetic Analyzer capillary electrophoresis system (Applied Biosystems). A 0.5-μL aliquot of the amplified product was added to 12 μL of deionized formamide (Amresco, Solon, OH) and 0.2 μL of GeneScan 500 LIZ internal lane standard (Applied Biosystems). Tubes containing the above were heated at 95°C for 3 min and snap-cooled on ice for 3 min. Samples were injected using Module G (5-s injection, 15 kV, 60°C) and analyzed with GeneScan Analysis Software v3.7 using Filter Set G (Applied Biosystems).

Minimal Haplotype Loci (MHL) Analysis—MHL analysis was carried out as described previously (21).

Multiplex System Performance

Multiplex Sensitivity—Different input quantities of template male DNA were tested using the standard MPIV reaction conditions. The amounts tested were: 50 pg, 100 pg, 125 pg, 150 pg, 200 pg, 250 pg, 500 pg, 1 ng, 3 ng, 5 ng, 10 ng, and 100 ng.

Specificity—To evaluate possible female DNA cross-reactivity, DNA from female volunteers was tested using the following amounts: 3 ng, 30 ng, 300 ng, and 1 μg.

Mixture Studies

(1) Male/Female: 3 ng of male DNA was co-amplified with increasing amounts of female DNA in the following ratios: 1/2 (3 ng of male DNA/3 ng of female DNA); 1/10 (3 ng of male DNA / 27 ng of female DNA); 1/100 (3 ng of male DNA/297 ng of female DNA).
(2) Male/Male: DNA from two individual males was combined in the following ratios, 1/2, 1/3, 1/6, and 1/12. In each case 3 ng of the admixed DNA was tested.

Nonprobative Casework

Post-coital cervicovaginal swabs were taken from a female participant at various time intervals after sexual intercourse (immediately (0 h) and 12, 24, 36, and 48 h). Only one set of swabs was taken after an individual act of sexual intercourse to ensure that the amount of

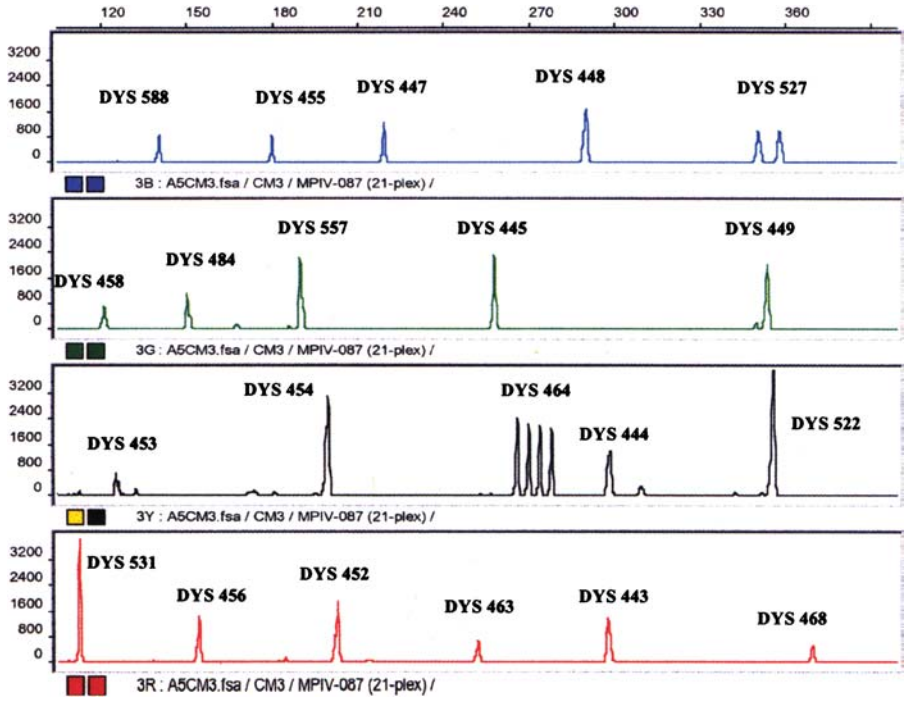

FIG. 1—*Multiplex IV (MPIV). Twenty-one Y- STR loci are co-amplified in a single reaction, separated by capillary electrophoresis, and displayed as an electropherogram. The x-axis represents fragment size (bp) and the y-axis represents signal intensity (relative fluorescence intensity, RFU). Each locus is labeled with a fluorescent dye: top channel—6-FAM (blue); second channel—VIC (green); third channel—NED (yellow); bottom channel—PET (red).*

semen present at the time was not affected by prior removal of sample. Two swabs were taken at each time interval and DNA isolated as described above. One of the swabs was extracted using a differential extraction method to separate the sperm and nonsperm fractions.

Population Studies

Descriptive Statistics—All data were checked for duplicate samples. A small number of duplicate samples ($n = 4$) were removed from the dataset after confirmation by the submitting laboratory of their duplicate status by autosomal STR analysis. The following formulae were used: (1) discriminatory capacity = number of individuals/number of different haplotypes; (2) gene diversity (h), equivalent to the expected frequency of heterozygotes with autosomal diploid loci, was calculated as $(n/(n-1)) \times (1 - \Sigma p_i^2)$, where p_i = allele frequency at the ith locus (22).

Results and Discussion

Description of Y-STR Multiplex System

A candidate set of loci with compatible annealing temperatures and appropriate amplicon size were identified by scanning the >150 non-MHL Y-STRs in the Human Genome Database (GDB) (20). A 21-locus subset was successfully incorporated into a single multiplex system, which is referred to as Multiplex IV (MPIV) since previous multiplex Y-STR systems developed in our laboratory are referred to as MPI, MPII and MPIII. The new loci include DYS 443, DYS 444, DYS 445 (23); DYS 447, DYS 448 (15), (16); DYS 449, DYS 452, DYS 453, DYS 454, DYS 455, DYS 456, DYS 458, DYS 463, DYS 464 (16); DYS 468, DYS 484, DYS 522, DYS 527, DYS 531, DYS 557, and DYS 588 (20). Alleles range in size from 100 to 400 bp (Fig. 1). Although the multiplex system contains 21 Y-STR loci, it comprises a total of 25 sites exhibiting allelic variation (since one locus is bi-local and one is tetra-local), which has prompted the use of the descriptor "megaplex" to denote the system. DYS 527 is the incorporated bi-local locus, whereas DYS 464 is the tetra-local locus in which each individual can possess up to four of nine common alleles (16).

The MPIV primer sequences (see Table 1) were obtained from the GDB (20) with the exception of DYS 468, which had to be re-designed to eliminate artifacts. The resulting DYS 468 amplicon was 12 bp longer than that obtained from using the GDB primer sequences.

Multiplex Optimization

The MPIV PCR reaction conditions were optimized by altering the concentration of critical reagents and the thermocycling conditions and using both male (3 ng) and female (300 ng) DNA as input template. This resulted in the determination of a set of standard reaction conditions described in the Methods Section.

Multiplex Performance

Sensitivity—A novel system developed for forensic use needs to be able to produce a profile with sub-nanogram quantities of template DNA. To test the sensitivity of the MPIV system, various amounts of template male DNA were amplified, ranging from 50 pg to 100 ng. Although 1 to 3 ng of DNA gave optimal signal intensity and inter-locus balance, a full 21-locus male haplotype was obtained with as little as 50 pg of male template DNA (Fig. 2), which is equivalent to approximately 8 to 9 diploid cells (24). The amplification with 50 pg of template DNA was repeated two additional times to ensure that the sensitivity of this multiplex was reproducible. Each of the three amplifications produced full 21-locus male haplotypes (with a >100 RFU threshold). With higher amounts of template DNA, 10 to 100 ng, there was an emergence of nonspecific products, particularly in the green (VIC) and yellow (NED) channels.

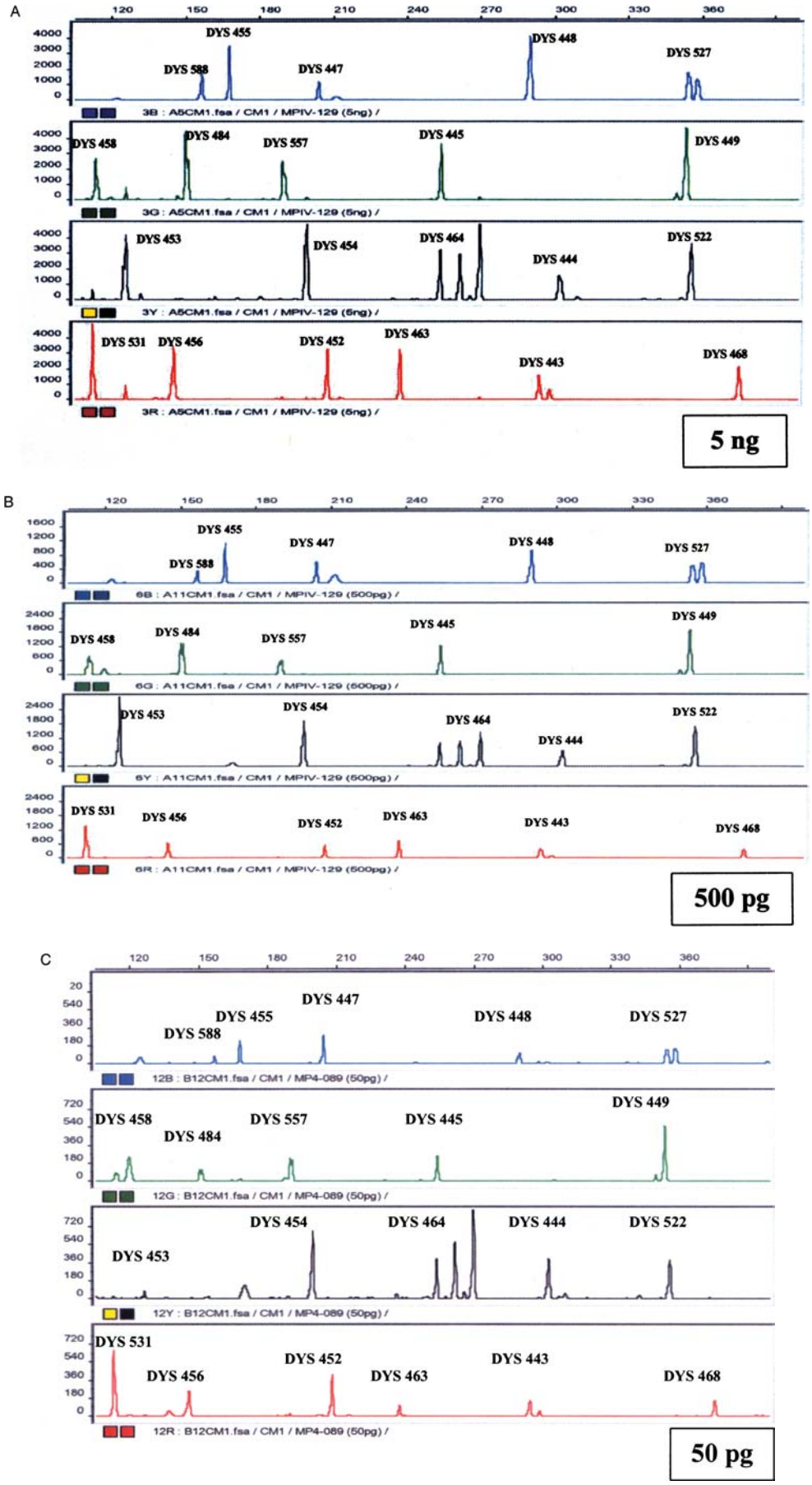

FIG. 2—*MPIV sensitivity. A complete 21-locus male haplotype is obtained with 5 ng of template DNA (A), with 500 pg of template DNA (B), and with 50 pg of template DNA (C).*

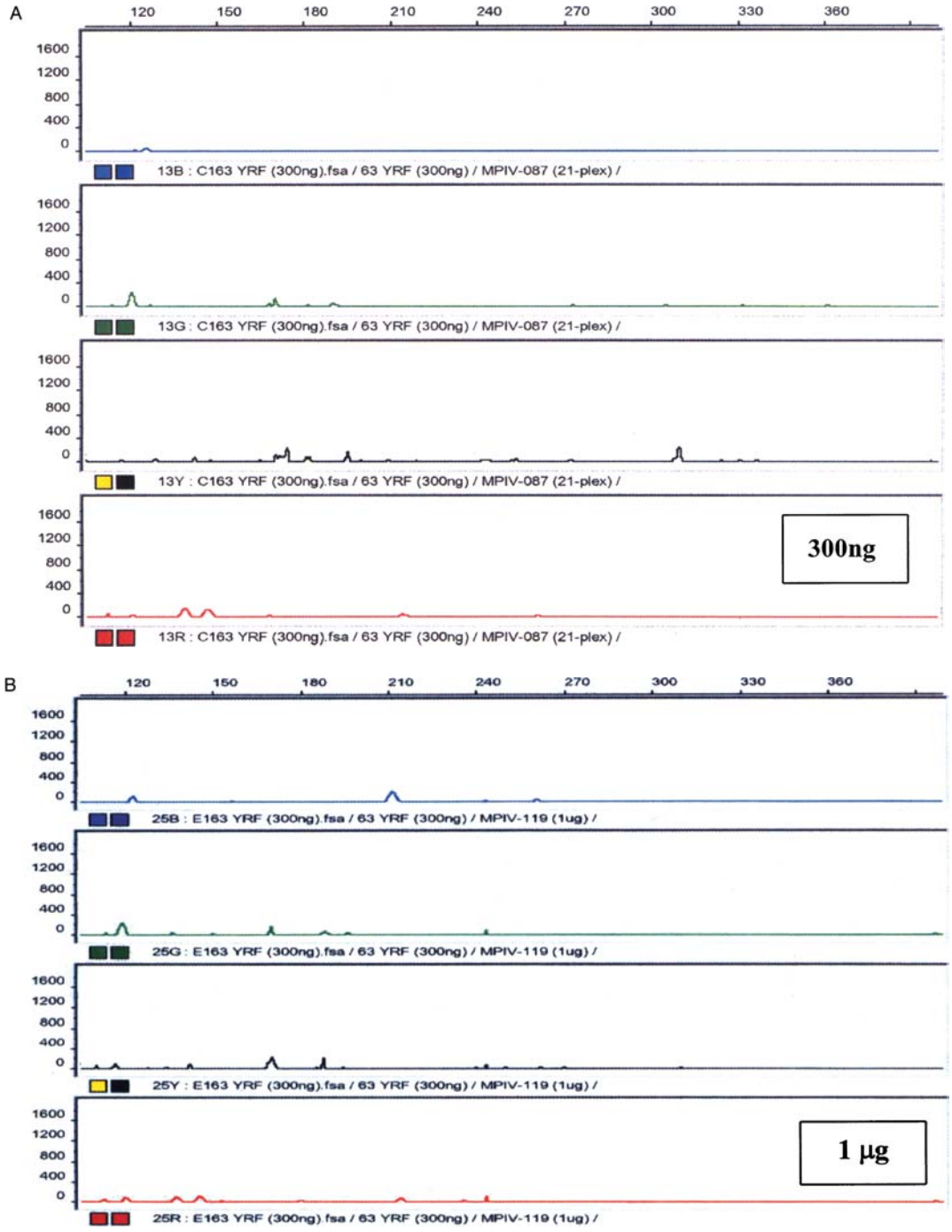

FIG. 3—*Female DNA products. Female DNA produces an insignificant number of artifacts with Multiplex IV using 300 ng of template DNA (A) and using 1 µg of template DNA (B).*

Specificity—A significant benefit of using Y-STR multiplex systems is the specific amplification of the male DNA component of an admixed male/female DNA sample, such as occurs in biological stains recovered during sexual assault investigations. Although there is little genetic recombination between the X and Y chromosomes, a degree of homology exists between both chromosomes due to their common evolutionary history (25). One of the main design objectives of Y-STR multiplex systems is to incorporate primers with enough specificity to ensure preferential amplification of Y- chromosome sequences, thus minimizing or eliminating any interference from the X chromosome. Without such specificity, X chromosome derived artifacts could confound analysis by their co-migration with *bona fide* Y-STR alleles and could reduce analytical sensitivity by titration of critical PCR reagents. Loci were chosen as candidates for incorporation into MPIV only after BLAT searches were performed on their primer sequences to determine the extent of homology with the X chromosome or with any other locations in the genome (26).

To assess female DNA cross reactivity, MPIV was tested with increasing quantities of female DNA (3 to 1000 ng). MPIV proved to be highly specific for the Y chromosome in that no significant female DNA products were observed up to 300 ng of female DNA (Fig. 3*A*). However, a number of artifacts of low signal intensity began to appear with increased quantities of female DNA (>300 ng),

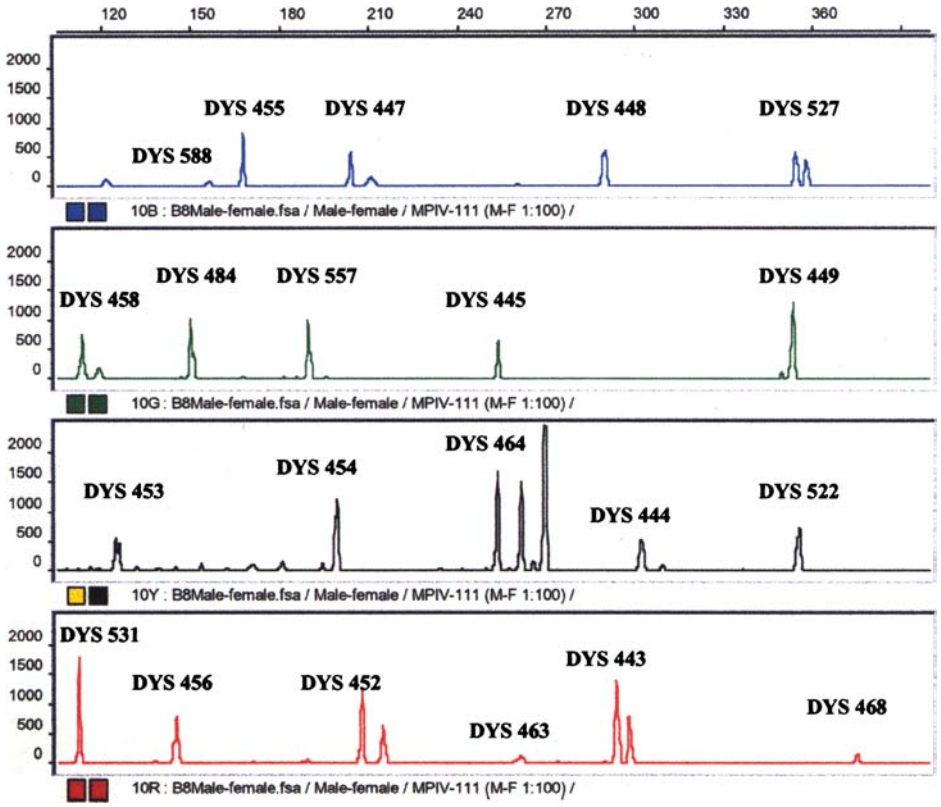

FIG. 4—*Male/Female admixed DNA samples. MPIV profile obtained at a male/female DNA ratio of 1/100.*

although most were not located within the allelic ranges of any of the MPIV loci. One yellow channel artifact, however, was produced by increased quantities of female DNA (300 ng) and was located at the extreme end of the DYS 444 allelic range (Fig. 3*A*).

Mixtures

Male/Female—To test the ability of MPIV to determine the genetic profile of the male donor in the presence of large quantities of female DNA, a situation which is more akin to *bona fide* forensic casework, a series of samples were prepared in which 3 ng of male DNA was admixed with varying quantities of female DNA and the total admixture amplified in a single reaction. A full 21-locus Y-STR profile was obtained when the male DNA component comprised 1/2, 1/10th, and 1/100th (Fig. 4) of the total. A monomorphic artifact peak, of unknown origin, at 210 bp began to appear in the red channel at these increased female DNA concentrations but did not preclude the ability to discern the male donor's profile. Importantly, these experiments demonstrate that even in the presence of an excess of female DNA it is possible, with MPIV, to obtain a Y-STR profile of the male donor.

Male/Male—Often there is more than one male contributor to a forensic casework sample. With traditional autosomal STR analysis, discerning the number of individual contributors in an admixture becomes increasingly difficult as the number of contributors exceeds two. However, since Y-STRs are inherited in a hemizygous manner resulting in one allele being found at most loci (i.e., excluding bi- and tetra-local loci), determining the number of male contributors in a sample is made relatively facile.

In order to demonstrate the detection of multiple male contributors, and also to determine the limits of detection of the minor component, male/male DNA admixtures from two unrelated individuals were tested at various ratios including 1/2, 1/3, 1/6, and 1/10, with 3 ng total DNA being amplified. For the 1/10 male/male sample, the minor component could not be seen. For the 1/1 (Fig. 5), 1/3, and 1/6 samples, two profiles could easily be discerned, with two allelic peaks being seen at a majority of the loci. Due to multiple, overlapping alleles, it is more difficult to determine the number of donors with the bi- and tetra-local loci, and their use for this purpose is not recommended.

Post-Coital Cervicovaginal Samples—A novel DNA typing system needs to be tested using simulated or nonprobative casework material prior to use in actual casework. Since Y-STR typing can have a significant impact on the analysis of rapes and other sexual assaults, we decided to concentrate our initial efforts in this area. A series of post-coital vaginal swabs were collected from a female subject at 0 to 48 h (in 12-h intervals) after intercourse. The participant was instructed to carry out normal daily functions to allow for regular semen drainage and degradation to most accurately simulate the type of samples encountered in forensic casework. Using a differential extraction, the sperm and nonsperm DNA fractions were separated and amplified using MPIV. Approximately 3 to 5 ng of template DNA was amplified for each of the sperm fractions, and approximately 100 ng of template DNA was amplified for each of the nonsperm fractions.

As predicted, the DNA recovered immediately after intercourse yielded complete 21-locus MPIV male profiles from both the sperm and nonsperm fractions (data not shown). The 12-h sperm fraction produced a full 21-locus Y-STR haplotype with relatively strong, artifact-free allelic signals (data not shown). Importantly for casework applications, the sensitivity and specificity of the system was such that full 21-locus Y-STR profiles were obtained 24 (Fig. 6*A*) and 48 h (Fig. 6*B*) after intercourse. Although some of the loci in

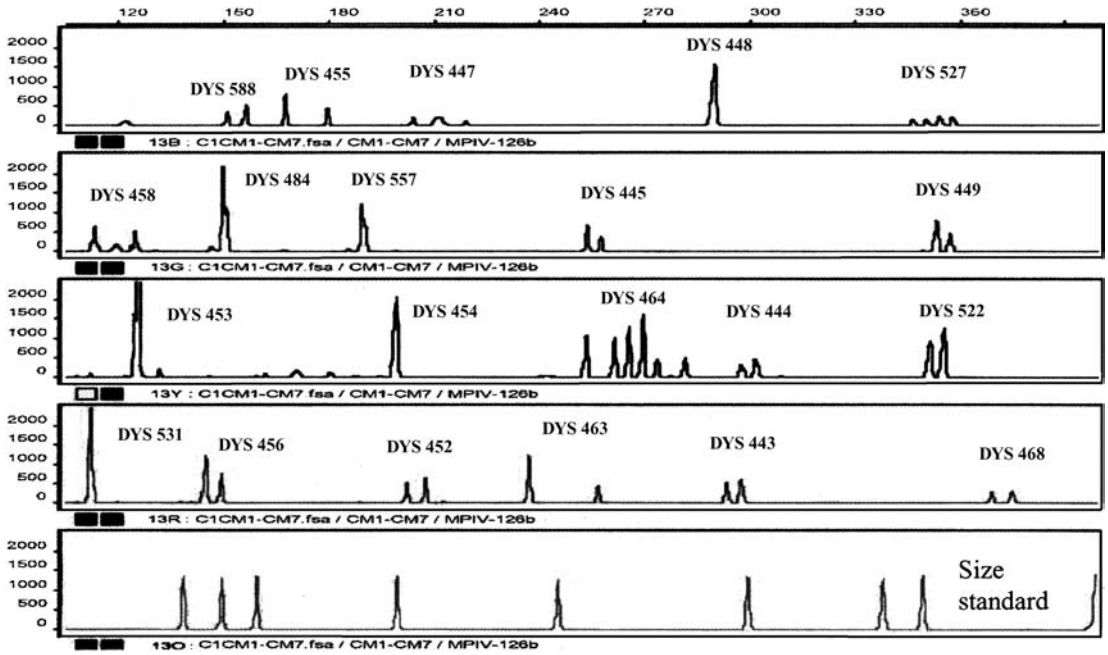

FIG. 5—*Male/Male admixed DNA samples. MPIV profile of DNA from two males (1/2 ratio).*

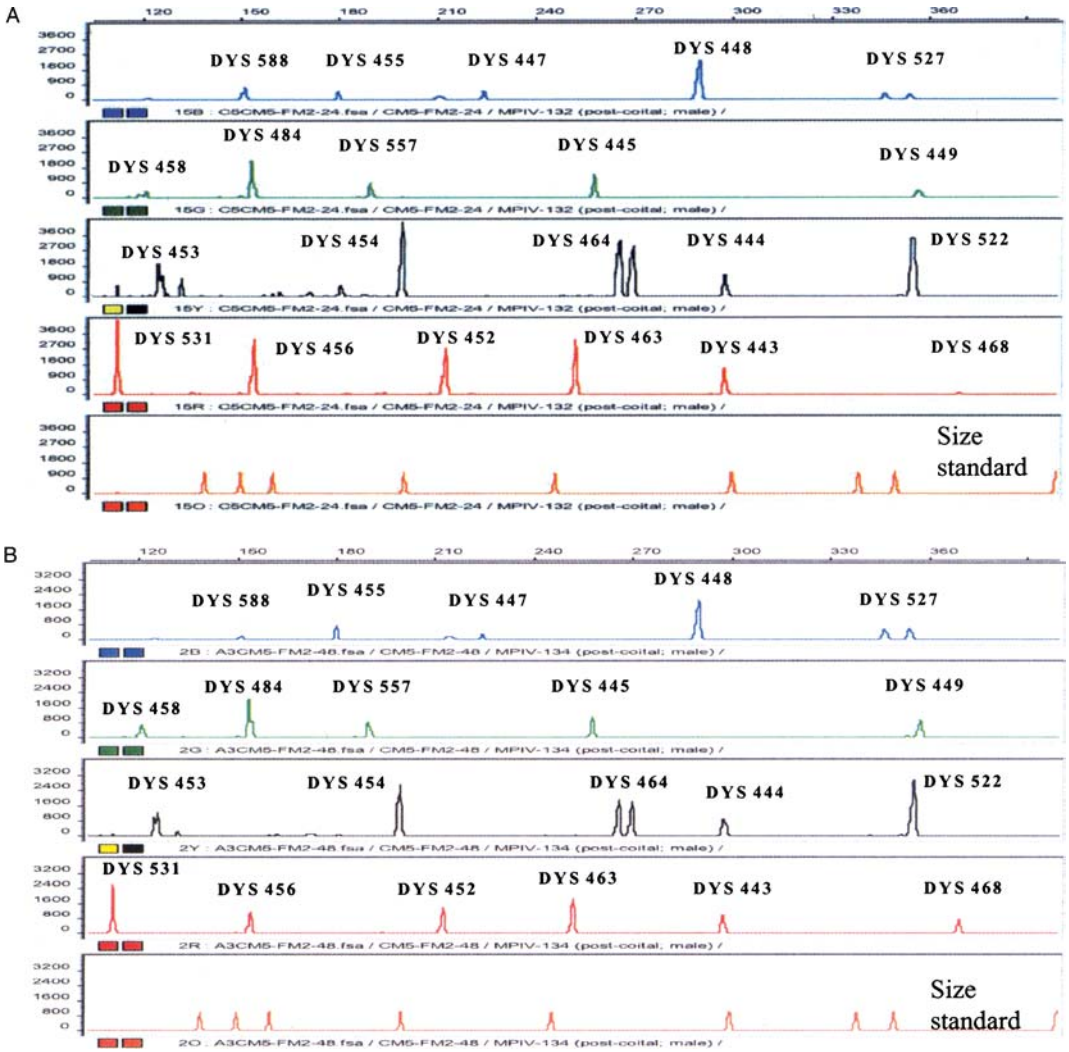

FIG. 6—*Post-coital cervicovaginal samples. MPIV profiles obtained from 1 ng of sperm fraction DNA recovered 24 h (A) or 48 h (B) after intercourse.*

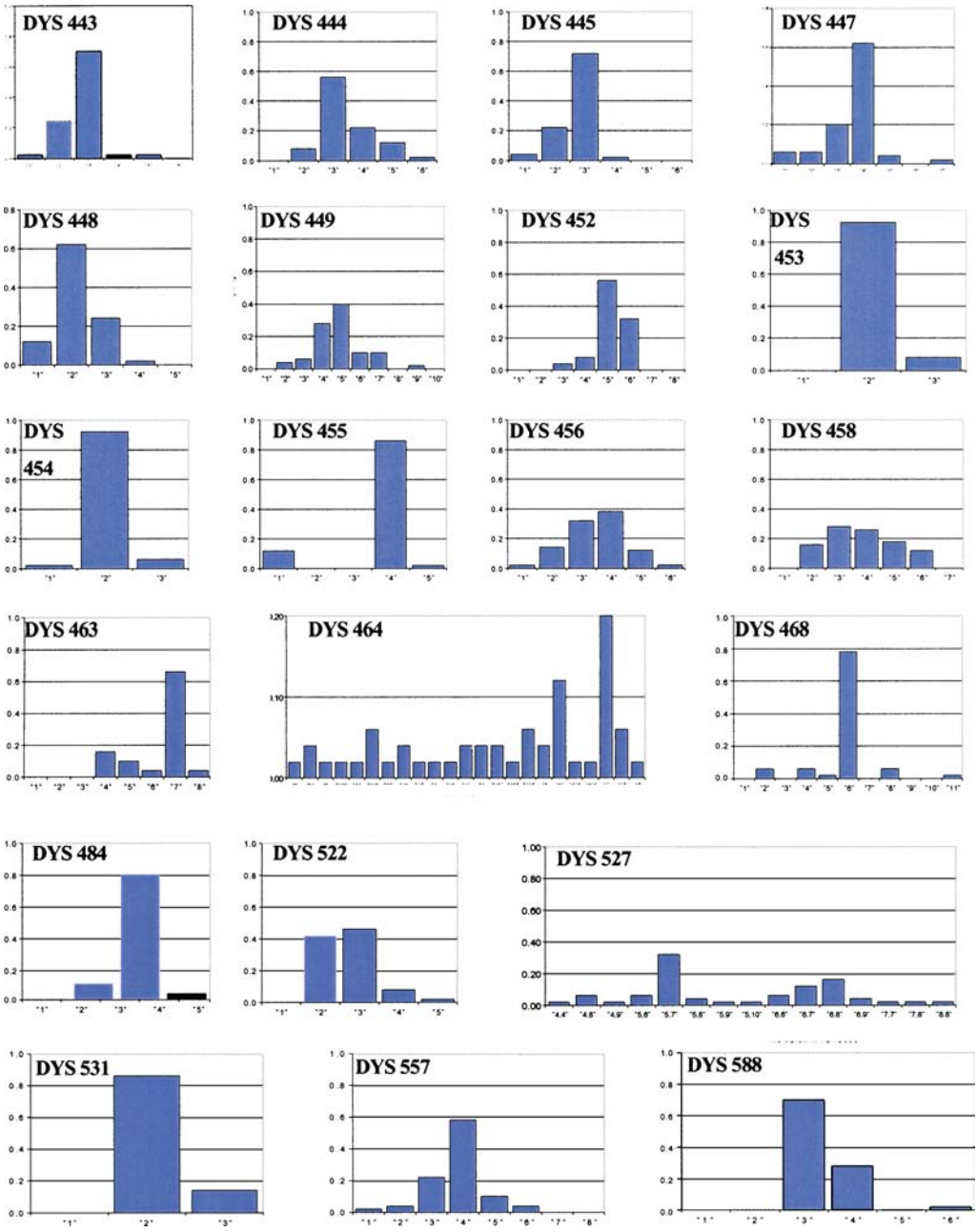

FIG. 7—*Allele frequency distributions of the 21 MPIV Y-STR loci in the Caucasian population (n = 50).*

the latter samples possessed lower signal intensities, they still met the minimum threshold requirements for allelic designation (i.e., RFU >100 together with a good signal to noise ratio).

Population Studies

An additional requirement of a Y-STR genetic marker system for forensic analysis is that it needs to be able to discriminate effectively between unrelated males, and the frequency of occurrence of any multilocus haplotype (i.e., the coincidence probability) should be as small as possible. In order to evaluate the efficacy of MPIV for individual discrimination, a preliminary population study was carried out.

Allele Frequency Distributions—The allele frequency distributions for each of the 21 MPIV loci were determined for both the Caucasian ($n = 50$) and the African American ($n = 51$) populations and are graphically illustrated in Figs. 7 and 8, respectively. The multi-local markers, DYS 527 and DYS 464, are plotted using multi-allelic haplotypes instead of individual alleles. The allelic distributions at several of the loci appear to differ significantly between the two populations (e.g., DYS 448, DYS 468, DYS 484, DYS 531, and DYS 557).

Individual Loci Gene Diversity—The gene diversity (h) for each locus, equivalent to the expected frequency of heterozygotes with autosomal loci and the discrimination potential (DP) for hemizygous loci such Y-STRs, is shown in Fig. 9. Fifteen, 13, and 9 of the 21 MPIV loci possess an h of >0.50, >0.60, and >0.70, respectively, for at least one population, demonstrating the high degree of discriminating ability afforded by most of the individual MPIV loci. One locus, the tetra-local DYS 464, has the highest h known for a Y-STR (>0.9).

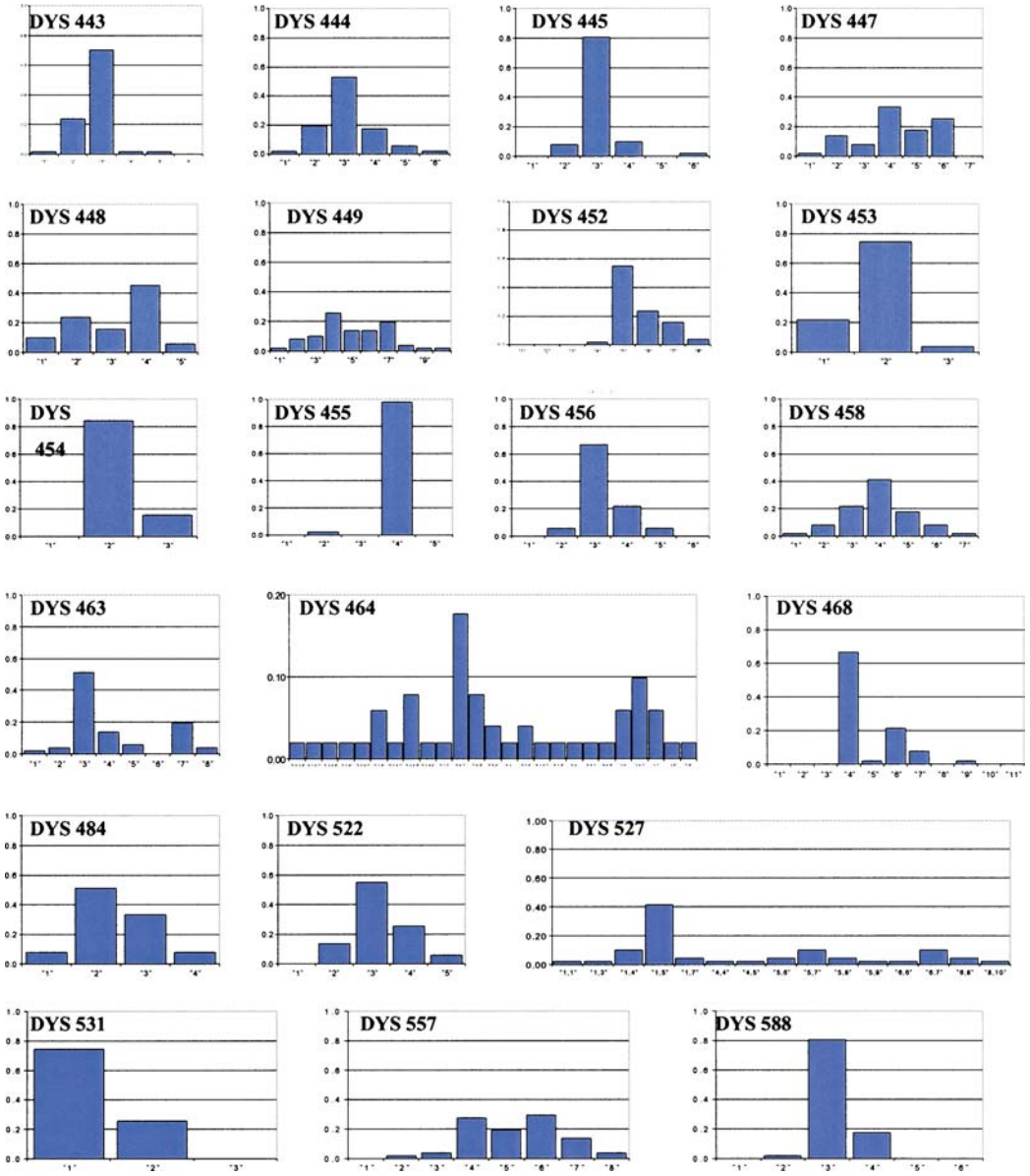

FIG. 8—*Allele frequency distributions of the 21 MPIV Y-STR loci in the African-American population (n = 51).*

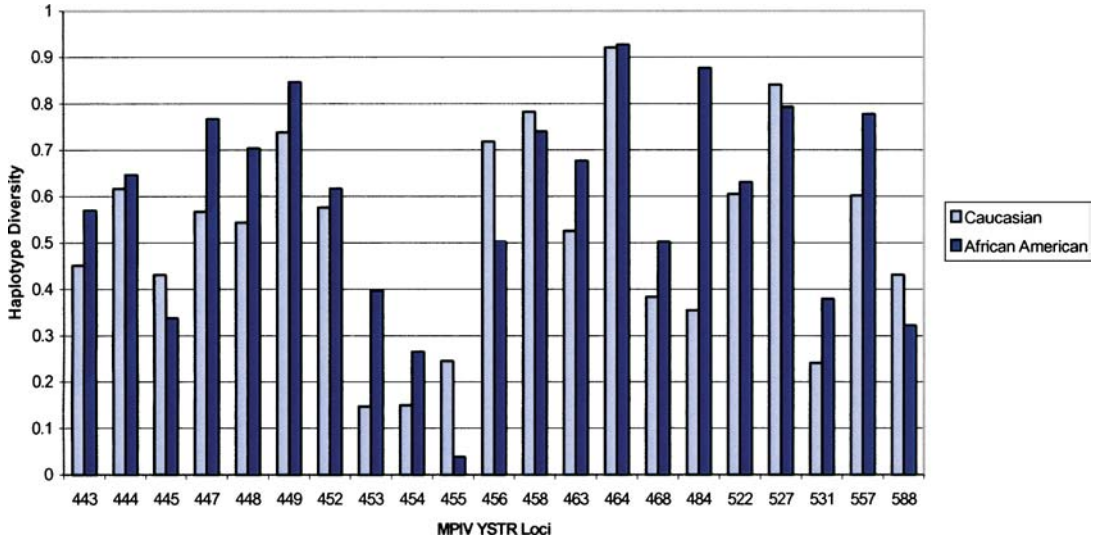

FIG. 9—*Gene diversity values for the multiplex IV loci in the Caucasian and African-American populations.*

TABLE 2—*Distribution of 21 locus Y-STR haplotypes for two major U.S. populations.*

	Caucasian	African-American	Total
Number of Individuals	50	51	101
Number of Haplotypes	50	51	101
Discriminatory Capacity	100%	100%	100%
No. of Haplotypes Observed Only Once (%)	50 (100%)	51 (100%)	101 (100%)
Occurrence of Most Frequent Haplotype (%)	1 (2.0%)	1 (2.0%)	1 (0.99%)

TABLE 3—*MPIV augments the discrimination potential afforded by the MHL loci.*

Group Number	Samples with Same MHL Profile	Samples with Same MHL + MPIV Profile
1	2	0
2	10	0
3	2	0
4	2	0
5	2	0
6	5	2
7	7	0
8	3	0
9	3	0
10	3	0
11	2	0
12	3	0
13	4	0
14	3	0
15	2	0
16	3	0
17	2	0
18	2	0
19	2	0
Total	62	2

MPIV Multi-Locus Discrimination Potential—Among 50 Caucasian samples, 50 unique MPIV Y-STR haplotypes were observed, whereas 51 unique MPIV haplotypes were found among the 51 African American samples (Table 2). Also, among the 101 total Caucasian and African-American samples, 101 unique MPIV haplotypes were found, indicating no duplicated haplotypes between the two population groups. Thus, the discriminatory capacity for MPIV was found to be 100% within the limited population samples tested, but, of course, this figure will decrease with an increased sample population size.

MPIV Augmentation of the MHL Discriminatory Power—In order to ascertain the degree to which MPIV might aid the MHL in individual discrimination, 330 individuals from three major population groups (Caucasian ($n = 139$), African American ($n = 120$), and Hispanic ($n = 71$)) were first genotyped using the MHL. Of the 330 samples tested, 62 (~19%) did not possess a unique MHL haplotype. These 62 individuals could be classified into 19 groups (each possessing a different haplotype) that consisted of 2 to 10 individuals (Table 3). Subsequently MPIV analysis was carried out on the same 62 individuals. All individuals, except for one pair, were distinguishable by the additional testing (see Table 3). Thus, of the 330 individuals, only two (~0.6%) did not possess unique MHL + MPIV haplotypes.

The discriminatory potential (DP) of a genetic marker system is defined as the probability that two individuals chosen at random would be distinguishable by the system (27–29). In the mixed population sample above, 54,285 pair-wise comparisons (i.e., $(330 \times 329)/2$) were carried out. In the case of the MHL loci, 109 pair-wise matches were observed giving a probability of match of approximately 1 in 500 and a DP of 0.998. When the MHL loci were augmented by MPIV, only one pair of individuals could not be distinguished, giving a probability of match of 1 in 54,000 and a DP of 0.999982. Thus, this preliminary study indicates that the use of MPIV in conjunction with MHL results in a hundredfold improvement in the ability to discriminate between two unrelated individuals compared to the use of MHL only. It is recognized that these DP estimates are based upon the particular admixed population sample described above and are expected to differ somewhat in other racial/ethnic groups and other admixed population samples. Nevertheless, the data support the notion that the use of MPIV would increase significantly the DP afforded by MHL testing only.

Conclusions

A highly discriminating Y chromosome STR "megaplex" system, Multiplex IV (MPIV), has been developed that permits the robust co-amplification of 21 novel Y-STR loci. Based upon a number of studies, including specificity, sensitivity, discriminating capacity and performance with nonprobative casework specimens, the system has demonstrated its potential forensic casework utility as an augmentation of the MHL.

Acknowledgments

The authors would like to gratefully acknowledge the National Institute of Justice for their support and funding for this project. The authors would like to thank Ms. Paulina Berdos for contributing unpublished population data for the minimal haplotype loci.

References

1. de Knijff P, Kayser M, Caglia A, Corach D, Fretwell N, Gehrig C, et al. Chromosome Y microsatellites: population genetic and evolutionary aspects. Int J Legal Med 1997;110(3):134–49. [PubMed]
2. Kayser M, Caglia A, Corach D, Fretwell N, Gehrig C, Graziosi G, et al. Evaluation of Y-chromosomal STRs: a multicenter study. Int J Legal Med 1997;110(3):125–33. [PubMed]
3. Kayser M, Brauer S, Willuweit S, Schadlich H, Batzer M, Zawacki J, et al. Online Y-chromosomal short tandem repeat haplotype reference database (YHRD) for U.S. populations. J Forensic Sci 2002;47(3):513–9.
4. Roewer L, Krawczak M, Willuweit S, Nagy M, Alves C, Amorim A, et al. Online reference database of European Y-chromosomal short tandem repeat (STR) haplotypes. Forensic Sci Int 2001;118(2–3):106–13. [PubMed]
5. Corach D, Filgueira Risso L, Marino M, Penacino G, Sala A. Routine Y-STR typing in forensic casework. Forensic Sci Int 2001;118(2–3):131–5.
6. Prinz M, Sansone M. Y chromosome-specific short tandem repeats in forensic casework. Croat Med J 2001;42(3):288–91. [PubMed]
7. Sinha S, Budowle B, Arcot S, Richey S, Chakrabor R, Jones M, et al. Development and validation of a multiplexed Y-chromosome STR genotyping system, Y-PLEX 6, for forensic casework. J Forensic Sci 2003;48(1):93–103. [PubMed]
8. Betz A, Bassler G, Dietl G, Steil X, Weyerman G, Pflug W. DYS STR analysis with epithelial cells in a rape case. Forensic Sci Int 2001; 118(2–3):126–30. [PubMed]
9. Gusmao L, Gonzalez-Neira A, Pestoni C, Brion M, Lareu M, Carracedo A. Robustness of the Y STRs DYS 19, DYS 389 I and II, DYS 390 and DYS 393: optimization of a PCR pentaplex. Forensic Sci Int 1999; 106(3):163–72. [PubMed]
10. Dekairelle A, Hoste B. Application of a Y-STR-pentaplex PCR (DYS 19, DYS 389I and II, DYS 390 and DYS 393) to sexual assault cases. Forensic Sci Int 2001;118(2–3):122–5. [PubMed]

11. Prinz M, Boll K, Baum H, Shaler B. Multiplexing of Y chromosome specific STRs and performance for mixed samples. Forensic Sci Int [PubMed] 1997;85(3):209–18.
12. Martin P, Albarran C, Garcia O, Garcia P, Sancho M, Alonso A. Application of Y-STR analysis to rape cases that cannot be solved by autosomal analysis. Prog in Forensic Gen 2000;8:526–8.
13. Ayub Q, Mohyuddin A, Qamar R, Mazhar K, Zerjal T, Mehdi S, et al. Identification and characterization of novel human Y-chromosomal microsatellites from sequence database information. Nucleic Acids Res [PubMed] 2000;28:e8.
14. White P, Tatum O, Deaven L, Longmire J. New, male-specific microsatellite markers from the human Y chromosome. Genomics 1999;57: [PubMed] 433–7.
15. Butler J, Schoske R, Vallone P, Kline M, Redd A, Hammer M. A novel multiplex for simultaneous amplification of 20 Y chromosome STR [PubMed] markers. Forensic Sci Int 2002;129(1):10–24.
16. Redd A, Agellon A, Kearney V, Contreras V, Karafet T, Park H, et al. Forensic value of 14 novel STRs on the human Y chromosome. Forensic [PubMed] Sci Int 2002;130(2–3):97–111.
17. Comey C, Presely K, Smerick J, Sobieralski C, Stanley D, Baechtel F. Extraction strategies for amplified fragment length polymorphism analysis. J Forensic Sci 1994;39:1254–69.
18. Gill P, Jeffreys A, Werrett D. Forensic application of DNA "fingerprints." [PubMed] Nature 1985;318:577–9.
19. Kline M, Duewer D, Redman J, Butler J. NIST mixed stain study 3: DNA quantitation accuracy and its influence on short tandem repeat multiplex [PubMed] signal intensity. Anal Chem 2003;75(10):2463–9.
20. http://www.gdb.org.
21. Hall A, Ballantyne J. Development of an 18-Loci Y-STR system for forensic casework. Anal Bio Chem 2003;376:1234–46.
22. Nei M. Molecular evolutionary genetics. Columbia (NY): Univ. Press, 1987.
23. Iida R, Tsubota T, Sawazaki K, Masuyama M, Matsuki T, Yasuda T, et al. Characterization and haplotype analysis of the polymorphic Y-STRs DYS 443, DYS 444 and DYS 445 in a Japanese population. Int J Leg Med 2002;116(3):191–4.
24. Kirby L. Chapter 4: Specimens. In: DNA fingerprinting: an introduction. New York: Stockton Press, 1990;53.
25. Lahn B, Pearson N, Jegalian K. The human Y chromosome, in the light of evolution. Nat Rev Genet 2001;2:207–16. [PubMed]
26. http://www.genome.ucsc.edu.
27. Sensabaugh G. Biochemical matches of individuals. In: Suferstein, editor. Forensic science handbook. Englewood Cliffs (CA): Prentice-Hall, 1982;338–403.
28. Fischer R. Standard calculations for evaluating a blood-group system. Heredity 1951;5:95–102. [PubMed]
29. Selvin S, Grunbaum B, Myhre B. The probability of non-discrimination or likelihood of guilt of an accused: criminal identification. J Forensic Sci Soc 1983;23(1):27–33. [PubMed]

Additional information and reprint requests:
Jack Ballantyne, Ph.D.
Associate Professor
Department of Chemistry
University of Central Florida
Bldg 5, 4000 Central Boulevard
Orlando, FL 32816-2366

J Forensic Sci, Nov. 2004, Vol. 49, No. 6
Paper ID JFS12480
Published 6 Oct. 2004
Available online at: www.astm.org

Erratum

Erratum/Correction of Hanson EK, Ballantyne J. A Highly Discriminating 21 Locus Y-STR "Megaplex" System Designed to Augment the Minimal Haplotype Loci for Forensic Casework. J Forensic Sci 2004 Jan;49(1):40–51.

It has come to the attention of the Journal that there is a misprint for the date for receipt of the manuscript. It should be 17 June 2003, not 2002 as printed.

The Journal regrets this error. Note: Any and all future citations of the above-referenced paper should read: Hanson EK, Ballantyne J. A Highly Discriminating 21 Locus Y-STR "Megaplex" System Designed to Augment the Minimal Haplotype Loci for Forensic Casework. [Published erratum appears in J Forensic Sci 2004 Nov;49(6)] J Forensic Sci 2004 Jan;49(1):40–51.