Int J Legal Med (2008) 122:421–427
DOI 10.1007/s00414-008-0258-y

ORIGINAL ARTICLE

Null allele sequence structure at the DYS448 locus and implications for profile interpretation

Bruce Budowle · Xavier G. Aranda · Robert E. Lagace · Lori K. Hennessy · John V. Planz · Manuel Rodriguez · Arthur J. Eisenberg

Received: 9 January 2008 / Accepted: 27 May 2008 / Published online: 26 June 2008

Abstract Null alleles can occur with any PCR-based STR typing system. They generally are due to deletions within the target region or primer binding sites or by primer binding site mutations that destabilize hybridization of at least one of the primers flanking the target region. Although not common, null types were detected at the DYS448 locus in seven out of 1,005 unrelated males in the Hispanic population. Of these DYS448 null types, four individuals displayed an apparent duplication at the DYS437 locus. The additional allele observed at the DYS437 locus is in actuality a smaller-sized DYS448 amplicon, which is the result of a deletion of the invariant N42 base pair domain and downstream repeats within the DYS448 locus. Thus, some DYS448 null types are not truly null. A true DYS448 null allele carried numerous primer binding site variants and a large deletion including the N42 base pair domain and surrounding or downstream repeat regions. The presence of null alleles is not a real concern for interpretation of Y STR loci evidence; current methods for interpreting Y STR profiles easily accommodate such phenomena.

Keywords Y STRs · Null alleles · DYS448 · DYS437 · Sequencing · Repeat motif · Interpretation · Statistics

B. Budowle (✉)
Laboratory Division, FBI,
Quantico, VA 22135, USA
e-mail: bruce.budowle@ic.fbi.gov

X. G. Aranda · J. V. Planz · M. Rodriguez · A. J. Eisenberg
UNT Center for Human Identification,
University of North Texas Health Science Center at Ft. Worth,
Ft. Worth, TX 76107, USA

R. E. Lagace · L. K. Hennessy
Applied Biosystems,
Foster City, CA 94404, USA

Introduction

The utilization of male-specific Y-chromosome STR loci are useful molecular tools for the genetic characterization of forensic biological evidence, particularly of male/female DNA mixtures where there are low quantities of male DNA amidst a large background of female DNA [1–11]. The analysis of Y STR loci has been greatly facilitated by the advent of commercial multiplex kits, such as the AmpF*l*STR® Yfiler™ kit (Applied Biosystems, Foster City, CA, USA) [8]. The kit contains reagents for the simultaneous amplification of 17 Y STR loci: DYS19, DYS385 (counted as two loci), DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438, DYS439, DYS456, DYS458, DYS635, DYS448, and Y GATA H4. A haplotype comprised of these 17 loci provides strong exculpatory evidence as only a very small percentage of unrelated males would be expected to share a complete haplotype.

Even with well-designed robust multiplex kits, genetic anomalies may be observed. One anomaly that can occur with any polymerase chain reaction (PCR)-based STR typing system is the generation of null allele phenotypes [12–14]. Null alleles generally can occur due to deletions within the target region or primer binding sites or by primer binding site mutations that destabilize hybridization of at least one of the primers flanking the target region. Initial primer design is based on sequence data housed in public databases [15–18]. Using in silico informatics and empirical testing, individual primers are designed and evaluated, and then multiplex assays are developed. The initial sequence information available for primer design often is limited to one or a few individuals. Many common variations within a population at the target or at the primer binding sites are unknown during initial primer design, resulting in some primers which yield no amplification

product, i.e., a null allele. To assess if null alleles are infrequent and the selected primers are effective for amplifying the alleles at the various loci, validation studies on samples from different population groups are carried out. However, if a null allele is infrequent, it is unlikely to be observed given the typical sample sizes used in population studies. Also, if a null allele is prevalent, but only so in less common population group(s), it may not be detected until that population is tested.

The presence of null alleles is not a real concern for most direct comparison cases and especially not for haploid markers, such as the Y STR loci. In most cases, a null profile at a locus in an evidence sample will be consistent with its true source reference sample, if the same primer sets are used. Of course, as is common with paternity testing, lineage studies must take into consideration the possibility of null alleles as a mismatch between samples may also occur due to mutational events during meiosis.

Null alleles typically are not common within the loci evaluated in the Y-Filer™ kit in the populations tested to date; when they arise, there are few, if any, problematic interpretation issues (see below). However, phenomena that require further evaluation and description sometimes arise. In the course of our Y STR population study on a Hispanic sample population (1,005 father/son pairs), a few samples demonstrated null types (N=7) at the DYS448 locus. Of these, a subset also showed an apparent duplication at the DYS437 locus (Fig. 1). There are two possible explanations for this observation. Either the DYS448 null type and the DYS437 duplication are independent events or the DYS448 null type is the result of a deletion, and the allele is amplified but now is located where DYS437 alleles migrate—with the latter one being the simpler explanation. This paper describes the molecular basis for some of the DYS448 null types (true or apparent) observed and how these observations would impact the calculation of the rarity of an observed Y STR haplotype with this phenomenon.

Materials and methods

Samples

Samples were obtained from fathers and sons from paternity testing cases submitted to DNA Identification Lab at the University of North Texas Health Science

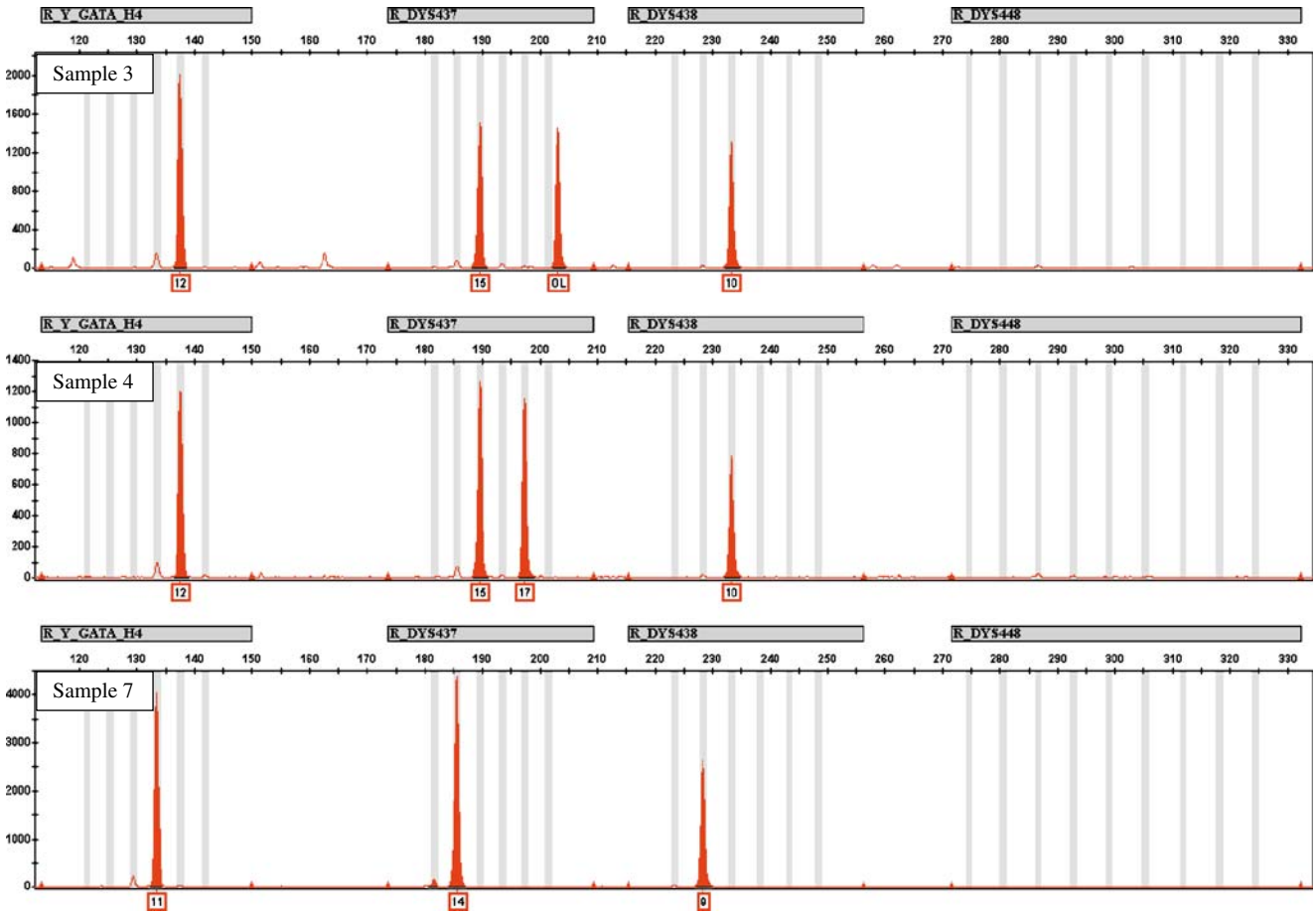

Fig. 1 Electropherograms of profiles for the null DYS448 locus and concomitant DYS437 locus

Center, Ft. Worth, TX, USA. Population affinity was ascribed by self-declaration. Paternity was confirmed by autosomal STR typing, with a minimum PI of 1,000.

Sterile foam-tipped swabs with a 6-in. polypropylene shaft (catalogue no.: 25-1616 2PF, Puritan Hardwood Products Company, LLC, Guilford, ME, USA) were used for buccal sample collection. Each swab was removed from the packaging by grasping the plastic handle and by directly placing the tip into the individual's mouth. The swab was vigorously rubbed against the inner cheek for a minimum of 20 up and down strokes per swab. The swab was allowed to air-dry.

Sample preparation

The buccal cells were eluted from the swab into 1 ml Isotone® III buffer (Beckman) contained in a 2.0-ml Dolphine (nipple) microcentrifuge tube (Costar, cat no. 3213, Corning, Corning, NY, USA). Cells were eluted by swirling and pelleted by centrifugation (1 min at 1,500×*g*). A 15-μl aliquot of Coomassie Brilliant Blue G250 (1 mg/ml in H_2O) was added to the eluted cells to aid visualization of the cell pellet for subsequent spotting onto the FTA® matrix. The majority of the supernatant was decanted and the pellet resuspended in the remaining residual buffer. A 15-μl aliquot was spotted directly onto the target circles on FTA® Micro Sheet™ (Whatman, Florham Park, NJ, USA). The FTA® Micro Sheet™ was air-dried for a minimum of 1 h prior to processing.

The FTA® Micro Sheet™ was placed into a Hybriboat™, and 35 ml of FTA® Purification Reagent containing 10 μg/ml of proteinase K was added and incubated at 65°C for 15 min. The reagent was discarded, and the FTA® Micro Sheet™ was washed according to the manufacturer's recommendations.

Y STR typing

The PCR amplification was performed using the AmpF*l*STR® Yfiler™ kit (Applied Biosystems) according to the manufacturer's instructions, except that a 12.5-μl reaction volume was employed with a 1.2-mm FTA® punch serving as the template DNA. Amplification was performed in an ABI PRISM® GeneAmp® 9700 Gold-plated or Silver block Thermal Cycler (Applied Biosystems) using the 9600 emulation mode for 28 cycles. PCR products were separated and detected on an ABI PRISM® 3100 and 3130 xl Genetic Analyzer (Applied Biosystems) following the manufacturer's recommendations. Prior to electrophoresis, 1.5 μl of the amplified product or allelic ladder and 0.3 μl of GeneScan™-500 LIZ® size standard (Applied Biosystems) were added to 8.7 μl of deionized Hi-Di™ formamide (Applied Biosystems), denatured at 95°C for 5 min, and then chilled on ice for 5 min. Samples were injected for 15 s at 3 kV in performance-optimized polymer (POP-4™; Applied Biosystems) using the GeneScan 36vb_POP4 Dye Set G5 Module for both instruments, and run time was 1,500 s. The data were collected using the ABI PRISM® 3100 Data Collection Software v1.1 (Applied Biosystems) and ABI PRISM® 3130 xl Genetic Analyzer Data Collection Software v3.0, respectively. Electrophoresis results were analyzed with GeneMapper® ID software v3.2 (Applied Biosystems). Allele peaks were called when the peak heights were greater than or equal to 50 relative fluorescence units.

Sequencing

Sequencing reactions were performed to analyze the alleles of the DYS448 null types. The region was amplified using the forward primer GAAGGTGGGTTTTAGTTGGCTATG and the reverse primer TGGAGGTTATTTCTTGAT TCCCTG. PCR amplifications were carried out using the same protocol as above, except that 800 nM of each primer was used. The PCR products were cloned using the TOPO TA Cloning® kit (Invitrogen, Carlsbad, CA, USA) and sequenced with BigDye® Terminator v 1.1 Cycle Sequencing Kit (Applied Biosystems) using M13 forward and reverse primers following the recommendations of the manufacturer. Capillary electrophoresis was performed on an ABI PRISM® 3130 xl Genetic Analyzer with Data Collection Software v3.0. The resulting sequencing data were analyzed with the software DNA Sequencing Analysis 5.2 (Applied Biosystems).

Results and discussion

Null types at the locus DYS448 have been reported previously [19–24]. Roewer et al. [23] identified seven out of 99 Kalmyk individuals with null types. This population study represents the highest DYS448 null allele frequency in a sample population reported to date. Several authors suggested that the null types may be due to deletions, although sequencing was not performed to verify that a deletion was the cause of the null type. Null types have been reported primarily in Asian populations; however, Sanchez et al. [24] identified two individuals out of 247 Catalonians who were null at the DYS448 locus. They recommended further studies to determine if the null types were due to a large deletion or a primer binding site mutation(s). In our study, seven out of 1,005 unrelated Hispanics from Texas demonstrated a DYS448 null type [the population genetic data on these 1005 Y STR haplotypes will be provided in a separate paper (manuscript in preparation)]. Unlike the other null alleles reported to date, four of the DYS448 null males also showed an additional allele at the DYS437 locus (Fig. 1; Table 1).

Table 1 Y STR haplotypes for the DYS448 null males

Sample[a]	DYS389I Allele 1	DYS389II Allele 1	DYS390 Allele 1	DYS456 Allele 1	DYS19 Allele1	DYS385 Allele 1	DYS385 Allele 2	DYS458 Allele 1	DYS437 Allele 1	DYS437[b] Allele 2
3	14	29	23	15	16	12	12	19	15	18.2
4	14	29	23	15	17	12	16	19	15	17
5	14	29	23	15	16	12	12	19	15	18.2
6	14	29	23	15	16	12	12	19	15	17
7	13	29	23	15	14	15	18	15	14	–
8	13	29	23	15	14	15	18	15	14	–
9	13	19	23	15	14	15	18	15	14	–

[a]Samples 1 and 2 were reference samples that do not have the DYS448 null type and therefore are not displayed in the table.
[b]The extra DYS437 alleles for samples 3–6 are actually DYS448 alleles with a deletion.

Note that the Y STR profiles are similar for the four individuals with the apparent DYS448 null type and DYS437 duplicate alleles as are the profiles for the three individuals displaying the DYS448 null type only. This observation supports that each phenomenon likely resides within a specific haplogroup [23], although haplogroup determination was not performed in this study.

The basic repeat motif of the DYS448 locus is the hexanucleotide repeat AGAGAT [25]. There are two polymorphic domains separated by an invariant 42-bp region. The basic structure at the locus is:

$$(\mathrm{AGAGAT})_m \mathrm{N42} (\mathrm{AGAGAT})_n$$

where m and n are integers representing the number of repeats, and N42 is the sequence AT(AGAGAT)AG(AGAGAT)$_3$(AGAT)$_2$(AGAGAA). Two individuals, a father/son pair (designated samples 1 and 2, respectively, herein), that exhibited allele 18 at the DYS448 locus were selected and sequenced to serve as a reference for alignment (Table 2). The sequences obtained for samples 1 and 2 were consistent with the basic structure of the locus and the same as that reported by Tang et al. [26]. The only difference was that the son had a SNP at the fifth repeat resulting in the hexanucleotide AGAGGT.

All seven null profiles were sequenced at the DYS448 locus. In all seven samples, there were deletions encompassing at a minimum the N42 region and the AGAGAT repeats downstream, and one sample (sample 9) had a deletion encompassing upstream repeats as well (all alignments were based on allele 18, i.e., samples 1 and 2). For samples 3–6, a DYS448 amplicon was generated, but due to the deletion smaller amplicons were created that migrated within the range of the DYS437 locus and were typed as an OL (18.2) and 17 allele, respectively. Thus, these four samples did not have DYS448 null alleles. They were designated “apparent null” alleles. Samples 7–9 also had a deletion in the same region. But there were also a number of nucleotide differences in the flanking regions, compared with the sequences of reference samples 1 and 2. In fact, the flanking region, and thus primer binding site, of samples 7–9 was very different in sequence composition compared with the sequences observed in samples 1–6. Samples 7–9 had a true null allele with the Y-Filer™ assay; no amplification occurred due to primer mismatch. These three samples were designated “true null” alleles. From an evolutionary perspective, extant data were insufficient to determine whether the deletion occurred before or after the numerous substitutions in the flanking region of samples 7–9 and their lineage. Therefore, while a deletion was found in all null alleles sequenced herein, albeit a very limited sampling, one cannot rule out the possibility that some DYS448 null alleles may not exhibit the deletion and are solely the result of a primer mismatch.

A null type observed at one locus, such as that shown for samples 7–9, does not present an interpretation problem for single source samples or mixtures. Indeed, null types have been observed in forensic DNA typing systems since their inception [27, 28] and can easily be addressed both for interpretation of the evidence and for assessing the rarity of a Y STR haplotype. For single source samples (and depending on signal strength across the profile), one could reasonably call the null type at the DYS448 locus. Alternatively, if no allele is observed, the locus could be deemed inconclusive and not used for inclusion or exclusion purposes when compared with a reference sample. This practice applies to any haplotype displaying a null type at a locus.

When an inclusion is obtained, the rarity of the Y STR haplotype can be calculated using existing practices. The forensically important Y STR loci reside in the non-recombining portion of the Y chromosome. Thus, the assumption of independence between and among loci does not apply for the Y STR loci. For statistical calculations, the Y STR haplotypes essentially are treated as alleles of a single locus when using the current method of choice for statistical assessment, the counting method [29, 30]. The counting method entails: (1) searching a population database(s) for the number of profiles matching the Y

DYS438 Allele 1	DYS448 Allele 1	Y_GATA_H4 Allele 1	DYS391 Allele 1	DYS392 Allele 1	DYS393 Allele 1	DYS439 Allele 1	DYS635 Allele 1
10	0	12	10	11	13	12	23
10	0	12	10	11	13	12	23
10	0	12	10	11	13	12	23
10	0	12	10	11	13	12	23
9	0	11	9	13	13	12	21
9	0	11	9	13	13	12	21
9	0	11	9	13	13	12	21

STR haplotype derived from the evidence, (2) counting the number of times the Y STR haplotype is observed in a database(s), (3) dividing that count by the total number of profiles in the reference database(s), and (4) placing a confidence interval on the proportion of count/total profiles in a database(s). A Y STR haplotype, such as that observed for samples 7–9, can be searched in the database(s), and only those database entries matching at all of the other loci of the haplotype and containing a null type at the DYS448 locus would be counted. Alternatively, the locus can be deemed inconclusive, and the DYS448 locus is not considered when performing the database search and count for the rest of the haplotype loci.

For a mixed profile, a null type is not problematic either. Consider a two-person, mixed profile displaying two alleles at the other 16 loci, and the DYS448 locus shows only one allele, for example, an allele 21 (and assume that allele dropout due to stochastic effects during PCR is not a consideration). Those haplotypes that share alleles with the mixed profile and cannot be excluded as part contributors would be counted. The mixed profile at the DYS448 locus is either the result of the contributors both having allele 21 or of one of the contributors having an allele 21 and the other having a null allele. Using the same counting method principles as above and the mixture statistical approaches described by Budowle et al. [29, 30], all profiles that cannot

Table 2 Sequence in the relevant flanking and repeat region of the DYS448 locus for normal and null alleles

Sequence**
Sample 1 (Genbank accession number EU682948—18 allele reference for alignment) ATCGCGAGACAGAAAGGGAGATAGAGACATGGA TAA(AGAGAT)$_{11}$N42(AGAGAT)$_7$AGAGAGGTAAAGATAGAGATAAATTTCCAGACCGGCCAGAA
Sample 2 (Genbank accession number EU682949—18 allele reference with SNP variation from biological father)ATCGCGAGACAGAAAGG GAGATAGAGACATGGATAA(AGAGAT)$_4$AGAG**G**T (AGAGAT)$_6$N42 (AGAGAT)$_7$AGAGAGGTAAAGATAGAGATAAATTTCCAGACCGG CCAGAA
Sample 3 and sample 5 (Genbank accession number EU682950; Genbank accession number EU682951—apparent null)[a] ATCGCGAGACAGA AAGGGAGATAGAGACATGGATAA(AGAGAT)$_{11}$—[84 bp del relative to 18 allele (N42 and AGAGAT$_7$)]—AGAGAGGTAAAGATAGAGA TAAATTTCCAGACCGGCCAGAA
Sample 4 (Genbank accession number EU682953—apparent null) ATCG**T**GAGACAGAAAGGGAGATAGAGACATGGATAA(AGAGAT)$_{10}$—[90 bp del relative to 18 allele (N42 and AGAGAT$_8$)]—AGAGAGGTAAAGATAGAGATAAATTTCCAGACCGGCCAGAA
Sample 6 (Genbank accession number EU682952—apparent null) ATCGCGAGACAGAAAGGGAGATAGAGACATGGATAA(AGAGAT)$_{10}$—[90 bp del relative to 18 allele (N42 and AGAGAT$_8$)]—AGAGAGGTAAAGATAGAGATAAATTTCCAGACCGGCCAGAA
Sample 7 Genbank accession number EU682954—true null) AT**AAT**G**G**GA**T**AGAAAG(4 bp del)AT(14 bp Del) (AGAGAT)$_8$A**C**AGATAGAG ATA**C**AGATA**G**AGA—[82 bp del relative to 18 allele (N42 and AGAGAT$_6$)]—**T**AGATAAAGATAGAGAT**T**AA**C**TTCC**T**GAC**A**GGC**T**AGAA
Sample 8 (Genbank accession number EU682955—true null) AT**AAT**G**G**GA**T**AGAAAG (4 bp del)AT(14 bp Del)(AGAGAT)$_5$AGAGA**A** (AGAGAT)$_2$ A**C**AGATAGAGATA**C**AGATA**G**—[82 bp del relative to 18 allele (N42 and AGAGAT$_6$)]—AGA**T**AG**A**TAAAGATAGAGAT**T**A A**C**TT CC**T**GAC**A**GGC**T**AGAA
Sample 9 (Genbank accession number EU682956—true null) A**CA**G**T**GAGACAGAAAG (4 bp del) AT (34 bp del) AGATAGAGATAGAGA **A**AGA(2 bp del)TAG—[106 bp del relative to 18 allele (AGAT + AGAGAT$_3$, N42 and AGAGAT$_6$)]—AGA**TG**G**A**TAAA**C**ATAGAGATAAA **C**TTC**AT**GAC**A**GGCCAGAA

While samples 3 and 5 have the same Y STR profile and DYS448 allele sequence, respectively, their 13 autosomal loci profiles are quite different.
[a] Sequences were evaluated and aligned with reference to allele 18 using MEGA software version 4.0 [31].
**SNPs are identified by boldface and underlining.

be excluded at the 16 loci and carrying a DYS448 allele 21 and a null allele would be haplotypes counted for a statistical assessment. The same approach would apply to any profile having any locus that carries a null type.

Observing a Y STR profile such as seen for samples 3–6 might, at first glance, create a concern for interpretation and statistical analysis. It is not possible, without sequencing the samples, to unequivocally assign alleles when a DYS448 null type and two alleles at the DYS437 locus are observed. But specific allele assignment is unnecessary. For interpreting whether the reference sample is included or excluded as a possible source of the DNA evidence, evaluation of allele sharing between the profiles is still the recommended method of comparison. Because a haplotype is treated as a single entity for counting, the null and two allele patterns observed just represent another haplotype. Haplotypes can be counted if all alleles match or cannot be excluded. Duplications at the DYS437 locus are rare [7]; therefore, it is unlikely that many profiles with two allele patterns at the DYS437 locus will reside in the reference population database(s). If so, the result would be a more conservative count because haplotypes with shifted DYS448 alleles and duplicated DYS437 alleles that appear identical will be combined for counting purposes.

In conclusion, DYS448 null alleles have been sequenced and their molecular composition described. They have deletions that encompass the N42 region and downstream repeat units, with one sample also having upstream repeats deleted. Amplicon size of these DYS448 alleles will result in a change in electrophoretic migration, and the alleles will be observed in the DYS437 locus range in the electropherogram. Therefore, these null types are not the result of true null alleles; they are apparent null types. A true DYS448 null allele may occur when there are variants in the flanking region where either primer binds. In this case, no extra allele will be observed at locus DYS437. Observing a null profile at the DYS448 locus solely or in concert with an extra allele at the DYS437 locus should not present interpretation problems; current methods for interpreting Y STR profiles easily accommodate such phenomena.

Acknowledgment This is publication number 08-02 of the Laboratory Division of the Federal Bureau of Investigation. Names of commercial manufacturers are provided for identification only, and inclusion does not imply endorsement by the Federal Bureau of Investigation.

References

1. Corach D, Filgueira Risso L, Marino M, Penacino G, Sala A (2001) Routine Y-STR typing in forensic casework. Forensic Sci Int 118:131–135
2. Gusmão L, Butler JM, Carracedo A (2006) DNA Commission of the International Society of Forensic Genetics (ISFG): an update of the recommendations on the use of Y-STRs in forensic analysis. Int J Leg Med 120:191–200
3. Hohoff C, Dewa K, Sibbing U, Hoppe K, Forster P, Brinkmann B (2007) Y-chromosomal microsatellite mutation rates in a population sample from northwestern Germany. Int J Leg Med 121:359–363
4. Honda K, Roewer L, de Knijff P (1999) Male DNA typing from 25-year-old vaginal swabs using Y chromosomal STR polymorphisms in retrial request case. J Forensic Sci 44:868–872
5. Jobling MA, Pandya A, Tayler-Smith C (1997) The Y chromosome in forensic and paternity testing. Int J Leg Med 110:118–124
6. Kayser M, de Knijff P, Dieltjes P et al (1997) Applications of microsatellite-based Y-chromosome haplotyping. Electrophoresis 18:1602–1607
7. Krenke BE, Viculis L, Richard ML et al (2005) Validation of a male-specific, 12-locus fluorescent short tandem repeat (STR) multiplex. Forensic Sci Int 148:1–14
8. Mulero JJ, Chang CW, Calandro LM, Green RL, Li Y, Johnson CL, Hennessy LK (2006) Development and validation of the AmpFlSTR Yfiler PCR amplification kit: a male specific, single amplification 17 Y-STR multiplex system. J Forensic Sci 51:64–75
9. Park MJ, Lee HY, Chung U, Kang SC, Shin KJ (2006) Y-STR analysis of degraded DNA using reduced-size amplicons. Int J Leg Med 121:152–157
10. Prinz M, Boll K, Baum H, Shaler B (1997) Multiplexing of Y-chromosome specific STRs and performance of mixed samples. Forensic Sci Int 85:209–218
11. Sinha S, Budowle B, Chakraborty R et al (2004) Utility of the Y-STR Y-PLEX™ 6 and Y-PLEX™ 5 in forensic casework and 11 Y-STR haplotype database for three major population groups in the United States. J Forensic Sci 49:691–700
12. Budowle B, Sprecher CJ (2001) Concordance study on population database samples using the PowerPlex™ 16 Kit and AmpFlSTR® Profiler Plus™ Kit and AmpFlSTR® COfiler™ Kit. J Forensic Sci 46:637–641
13. Budowle B, Masibay A, Anderson SJ et al (2001) STR primer concordance study. Forensic Sci Int 124:47–54
14. Leibelt C, Budowle B, Collins P et al (2003) Identification of a D8S1179 primer binding site mutation and the validation of a primer designed to recover null alleles. Forensic Sci Int 133:220–227
15. Chang CW, Mulero JJ, Budowle B, Calandro LM, Hennessy LK (2006) Identification of a novel polymorphism in the X-chromosome region homologous to the DYS456 locus. J Forensic Sci 51:344–348
16. Collins FS, Brooks LD, Chakravarti A (1998) A DNA polymorphism discovery resource for research on human genetic variation. Genome Res 8:1229–1231
17. Fredman D, Munns G, Rios D et al (2004) HGVbase: a curated resource describing human DNA variation and phenotype relationships. Nucleic Acids Res 32:D516–D519
18. Wheeler DL, Barrett T, Benson DA et al (2005) Database resources of the National Center for Biotechnology Information. Nucleic Acids Res 33:D39–D45
19. Chang YM, Perumal R, Keat PY, Kuehn DL (2007) Haplotype diversity of 16 Y-chromosomal STRs in three main ethnic populations (Malays, Chinese and Indians) in Malaysia. Forens Sci Int 167:70–76
20. Park MJ, Kyoung-Jin S, Kim NY, Yang WI, Cho SH, Lee HY (2008) Characterization of deletions in the DYS385 flanking region and null alleles associated with AZFc microdeletions in Koreans. J Forens Sci 53:331–334
21. Parkin EJ, Kraayenbrink T, Opgenort JR, van Driem GL, Tuladhar NM, de Knijff P, Jobling MA (2007) Diversity of 26-locus Y-STR haplotypes in a Nepalese population sample: isolation and drift in the Himalayas. Forens Sci Int 166:176–181

22. Mizuno N, Nakahara H, Sekiguchi K, Yoshida K, Nakano M, Kasai K (2008) 16 Y chromosomal STR haplotypes in Japanese. Forens Sci Int 174:71–76
23. Roewer L, Krüger C, Willuweit S et al (2007) Y-chromosomal STR haplotypes in Kalmyk population samples. Forensic Sci Int 173:204–209
24. Sánchez C, Barrot C et al (2007) Haplotype frequencies of 16 Y-chromosome STR loci in the Barcelona metropolitan area population using Y-Filer kit. Forensic Sci Int 72:211–217
25. Redd AJ, Agellon AB, Kearney VA et al (2002) Forensic value of 14 novel STRs on the human Y chromosome. Forensic Sci Int 130:97–111
26. Tang JP, Hou YP, Li YB, Wu J, Zhang J, Zhang HJ (2003) Characterization of eight Y-STR loci and haplotypes in a Chinese Han population. Int J Legal Med 117:263–270
27. Budowle B, Giusti AM, Waye JS et al (1991) Fixed bin analysis for statistical evaluation of continuous distributions of allelic data from VNTR loci for use in forensic comparisons. Amer J Hum Genet 48:841–855
28. Chakraborty R, de Andrade M, Daiger SP, Budowle B (1992) Apparent heterozygote deficiencies observed in DNA typing data and their implications in forensic applications. Ann Hum Gen 56:45–57
29. Budowle B, Sinha SK, Lee HS, Chakraborty R (2003) Utility of Y-chromosome STR haplotypes in forensic applications. Forensic Sci Rev 15:153–164
30. Budowle B, Ge J, Chakraborty R (2007) Basic principles for estimating the rarity of Y-STR haplotypes derived from forensic evidence. Eighteenth International Symposium on Human Identification 2007. Promega Corporation, Madison, Wisconsin http://www.promega.com/ussymp18proc/default.htm
31. Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol Bio Evol 24:1596–1599