International Congress Series 1239 (2003) 131–135

Multiplex STR genotyping: comparison study, population data and new sequence information

C. Alves [a,*], L. Gusmão [a], L. Pereira [a,b], A. Amorim [a,b]

[a] *IPATIMUP, Instituto de Patologia e Imunologia Molecular da Universidade do Porto, R. Dr. Roberto Frias, s/n, 4200 Porto, Portugal*
[b] *Faculdade de Ciências, Universidade do Porto, 4050 Porto, Portugal*

Abstract

A simultaneous study was performed using Perkin-Elmer Profiler Plus (PP)/GenePrint CTTv and Promega Powerplex 16 (P16) kits, which share 12 STR loci, in two population samples from North Portugal (*N*=110) and Mozambique (*N*=110). Population data for the three STRs included only in the Powerplex 16 kit (D16S539, Penta D and Penta E) are reported for both populations. No deviations from Hardy–Weinberg equilibrium were observed for all markers in both population samples. A comparative analysis between our data and Promega's (Caucasian-Americans and African-Americans) is also presented. Inconsistencies between genotypes obtained by both kits were detected for D5S818 (3/110 individuals from North Portugal) and D8S1179 (1/110 individuals from Mozambique) and are due to polymorphisms in the primer annealing regions. For D5S818 locus, the inconsistencies are due to a Promega primer annealing failure. In both populations, other point mutations outside the repeat region and structural differences from the consensus repeat motif were also detected. For D8S1179, the observed inconsistency is due to a Perkin-Elmer primer annealing failure.

Keywords: STR; PCR; Kits; Null alleles; North Portugal; Mozambique

1. Introduction

Human STR genotyping for population and forensic studies is nowadays mainly performed using commercial PCR amplification kits available from different manufacturers. The most recent ones amplify more than 10 loci, forensically very informative and well characterised by the scientific community. Although from different manufacturers, these multiplex kits share most of the STR loci now commonly used, which are amplified,

* Corresponding author.

0531-5131/03
PII: S0531-5131(02)00623-4

however, with distinct primer sets. If a primer is unable to anneal to the DNA template due to a polymorphism in its annealing region, then false homozygotes most likely appear.

Discrepancies between genotypes due to different primer sets has become a common finding in many laboratories (e.g. Refs. [1–7]) and only draws our attention to the need for caution when evaluating allele frequencies and STR profiles in population and forensic studies.

In this work, the same set of samples from North Portugal and Mozambique were genotyped using Perkin-Elmer's Profiler Plus (PP)/GenePrint CTTv and Promega's Powerplex 16 (P16) kits. The molecular bases for the discrepancies found were evaluated by sequencing analysis using new sets of primers.

Population data (North Portugal and Mozambique) for the STRs present only in the P16 kit (D16S539, Penta D and Penta E) are reported and compared to Promega's data.

2. Material and methods

A total of 110 unrelated donors from North Portugal as well as 110 from Mozambique were sampled by venipuncture or buccal swabs and DNA was extracted according to Ref. [8].

PCR amplification and genotyping were performed according to manufacturers instructions using the ABI 310 automatic sequencer and the GeneScan software.

Allele frequencies were estimated for the D16S539, Penta D and Penta E loci in samples from North Portugal (N=102–104) and Mozambique (N=105–109). Hardy–Weinberg equilibrium was assayed by an exact test [9], using GENEPOP software [10]. Expected heterozygosity (He) was calculated according to Nei [11]. Population differentiation was tested using the Arlequin software [12].

Sequencing analysis was performed using the following primer pairs:

D5S818 P1: 5′-TCTAATTAAAGTGGTGTCCCAGATA (forward);
P2: 5′-AGCCAAGTGATTCCAATCATAG (reverse).
D8S1179 P1: 5′-ATTGCAACTTATATGTATTTTTGTATTTCATG (forward)
P2: 5′-ACCAAATTGTGTTCATGAGTATAGTTTC (reverse).

The samples presenting null alleles in this study together with control samples with amplifiable alleles were amplified with the new sets of primers [6]. The PCR conditions were carried out according to the P16 protocol. PCR products were prepared and sequenced according to Ref. [6].

3. Results and discussion

North Portugal and Mozambique population data for the three STRs present only in the P16 kit (D16S539, Penta D and Penta E) are shown in Table 1. No deviations from Hardy–Weinberg equilibrium were observed for all markers in both populations. Data for the remaining 12 STRs for these populations have been previously published [13–15].

Table 1
Population data for D16S539, Penta D and Penta E STR loci

Allele	North Portugal			Mozambique		
	D16S539 (*N*=104)	Penta D (*N*=102)	Penta E (*N*=103)	D16S539 (*N*=109)	Penta D (*N*=107)	Penta E (*N*=105)
2.2	–	0.015	–	–	0.164	–
5	–	–	0.053	–	0.037	0.110
6	–	–	–	–	0.005	–
7	–	–	0.150	–	0.033	0.152
8	0.034	0.025	0.010	0.023	0.145	0.167
9	0.139	0.191	–	0.261	0.145	0.043
10	0.019	0.083	0.097	0.206	0.187	0.057
11	0.303	0.221	0.136	0.261	0.117	0.043
11.3	–	–	–	–	0.005	–
12	0.303	0.181	0.233	0.161	0.117	0.119
13	0.173	0.201	0.102	0.087	0.037	0.124
14	0.029	0.054	0.053	–	0.005	0.043
15	–	0.025	0.039	–	0.005	0.038
16	–	0.005	0.024	–	–	0.057
17	–	–	0.039	–	–	0.024
18	–	–	0.024	–	–	0.024
19	–	–	0.024	–	–	–
20	–	–	0.010	–	–	–
21	–	–	0.005	–	–	–
He	0.773	0.834	0.878	0.790	0.869	0.897

A comparative study between our northern Portuguese sample and that reported by Promega for Caucasian-Americans showed no significant differences for the three loci. However, our Mozambique sample showed significant differences for both D16S539 (P=0.000) and Penta D (P=0.000) loci when compared with African-American data.

The genotyping inconsistencies here described are due to a PP primer annealing failure for the D8S1179 locus and to a P16 primer annealing failure in the case of the D5S818 locus caused by sequence variations in the binding regions, resulting in false homozygotes (Table 2).

In order to assess if changes in the P16 PCR protocol for the D5S818 locus and the PP PCR protocol for the D8S1179 locus would amplify the null alleles, the annealing temperatures recommended in both cases (60 and 59 °C, respectively) were reduced. For the D5S818 locus, amplification of the null alleles with the P16 system occurs at 59 °C annealing temperature with half the peak height of the normal allele. In the case of the

Table 2
Genotype inconsistencies found in this work

STR system	Profiler Plus genotype	Powerplex 16 genotype	Sample
D5S818	11–12	12	North Portugal (*N*=110)
	10–12	12	
	10–13	13	
D8S1179	14	14–16	Mozambique (*N*=110)

Table 3
Sequence data obtained in this work (OUR) for D5S818 and D8S1179 STR alleles compared with GenBank sequences (GDB)

D5S818 (GenBank sequence G08446)

```
                                    1                                                  23
GDB:P1-atctgtactaataaaagtatattttaatagcaagtatgtgacaagggtgattttcctctttggtatccttatgtaat
OUR:P1-.........................c.............................................ca....
                                    4                          5
       attttga(agat)n---------- agaggtataaataaggatacagataaagatacaaatgttgtaaactgtgg-P2
       .......(agat)nacat(agat)3...t.............................c.............-P2
```

D8S1179 (GenBank sequence G08710)

```
GDB:P1-tgtacattcgta(tcta)1-2(tctg)0-2(tcta)n ttccccacagtgaaaataatctacaggataggtaaataaatt
OUR:P1-.....................................................................................
                    6
       aaggcatattcacgcaatgggatacgatacagtgatgaaaatgaactaattatagctacgt-P2
       ............a.............................................-P2
```

Numbers 1 to 6 denote positions where point mutations were detected in this work.

D8S1179 locus with the PP system, slight amplification of the null allele occurs at 58 °C and is well observed at 57 °C. Control samples were also amplified at lower annealing temperatures and there was no significant loss of specificity for all loci in both multiplex kits.

For the D5S818 locus, a total of 14 north Portuguese alleles (9 to 14 repeats) and 18 Mozambique alleles (8 to 15 repeats) were sequenced. Sequencing analysis of the three null alleles with the new set of primers revealed that the reason for non-amplification is due to a T to C substitution at the 36th position downstream of the last AGAT repeat (position 5, Table 3), inside the P16 reverse primer annealing region (position 197 of GenBank sequence G08446). Other point mutations outside the repeat region and, in some alleles, structural differences from the consensus repeat motif were detected. A few of these were only observed in samples from Mozambique: substitution 1 and 3 (Table 3) and the ACAT interruption in two Mozambique alleles (13 and 15 repeats). One more repeat is generated by substitution 4, present in 4/14 N. Portuguese and 6/18 Mozambique alleles analysed. The presence of C instead of T (position 2) revealed to be much more frequent than the GenBank sequence in both populations (12/14 for North Portugal and 16/18 for Mozambique).

For the D8S1179 locus, a total of nine northern Portuguese alleles (8 and 10 to 17 repeats) and 12 Mozambique alleles (10 to 17 repeats) were sequenced. Sequencing of the null allele revealed a G to A substitution at the 56th position downstream of the last TCTA repeat (position 6, Table 3), which must lie inside PP reverse primer annealing region (position 147 of GenBank sequence G08710). The same mutation was recently described [16]. No other point mutations were detected in the flanking regions.

Acknowledgements

This work was supported by FCT and "Programa de financiamento Plurianual de Unidades de I and D (POCTI) Quadro Comunitário de Apoio III". We thank Cindy Sprecher and Rita Weispfenning from Promega for their collaboration and for the D5S818 primers.

References

[1] L. Gusmão, et al., Int. J. Leg. Med. 108 (1996) 313–315.
[2] M.C. Kline, et al., J. Forensic Sci. 43 (1998) 250.
[3] S. Walsh, J. Forensic Sci. 43 (1998) 1103.
[4] B. Budowle, et al., J. Forensic Sci. 44 (1999) 1277.
[5] B. Budowle, Profiles DNA 3 (2000) 10.
[6] C. Alves, et al., Int. J. Leg. Med. 115 (2001) 97–99.
[7] L. Boutrand, et al., Int. J. Leg. Med. 114 (4–5) (2001) 295–297.
[8] J. Singer-Sam, et al., Amplifications: A Forum for PCR Users 3 (1989) 11.
[9] S.W. Guo, E.A. Thompson, Biometrics 8 (1992) 361–372.
[10] M. Raymond, F. Rousset, J. Hered. 86 (1995) 91–102.
[11] M. Nei, Genetics 89 (1978) 583–590.
[12] S. Schneider, et al., University of Geneva, 2000.
[13] A. Pérez-Lezaun, et al., Int. J. Leg. Med. 113 (2000) 208–214.
[14] A. Amorim, et al., Forensic Sci. Int. 115 (1–2) (2001) 119–121.
[15] C. Alves, et al., Forensic Sci. Int. 119 (1) (2001) 146–148.
[16] G.R. Han, et al., Int. J. Leg. Med. 115 (2001) 45–47.