Int J Legal Med (2011) 125:685–694
DOI 10.1007/s00414-010-0539-0

ORIGINAL ARTICLE

Development and validation of I-DNA1: a 15-Loci multiplex system for identity testing

A. Odriozola · J. M. Aznar · D. Celorrio · M. L. Bravo · J. J. Builes · J. F. Val-Bernal · Marian M. de Pancorbo

Received: 12 October 2010 / Accepted: 1 December 2010 / Published online: 22 January 2011

Abstract This paper presents a system for the multiplex amplification of 15 loci, known as I-DNA1, which combines mini and midiSTR technology, with amplicon sizes ranging from 49 to 297 bp. I-DNA1 analyses all the STR loci included in the CODIS and the Interpol Standard Set of loci, nine of the ten European core loci and seven of the eight German core loci, making it suitable for use in identifying humans. Moreover, its high sensitivity and the small size of its amplicons mean that I-DNA1 is potentially highly useful for analysing highly degraded and/or very small DNA samples.

Keywords Forensic science · Mini and midi short tandem repeat · STR · DNA typing · Degraded · Multiplex · DHPLC · CODIS

Adrian Odriozola and Jose Maria Aznar have contributed equally to this paper.

A. Odriozola · J. M. Aznar · D. Celorrio · M. M. de Pancorbo (✉)
Grupo BIOMICs/BIOMICs Research Group. Banco de ADN UPV/EHU. Universidad del País Vasco UPV/EHU. CIEA Centro de Investigación y Estudios Avanzados "Lucio Lascaray", Avda. Miguel de Unamuno, 3, 01006 Vitoria-Gasteiz, Álava, Spain
e-mail: marian.mdepancorbo@ehu.es

M. L. Bravo · J. J. Builes
Instituto de Biología, Universidad de Antioquia, Medellín, Colombia

J. J. Builes
Laboratorios Genes Ltda, Medellín, Colombia

J. F. Val-Bernal
Department of Anatomical Pathology, Marqués de Valdecilla University Hospital, University of Cantabria, Santander, Spain

Introduction

Short tandem repeat (STR) loci are outstandingly useful in forensic genetics because of their highly polymorphic nature [1]. Several core STR loci have been selected to establish national and international databases, such as the German Core Loci (GCL), the Combined DNA Index Systems (CODIS), the Interpol Standard Set of Loci (ISSL) and the European Core Loci (ECL), intended for the comparison of stored genetic profiles with profiles from suspects and/or evidence from crimes [2–4].

Multiplex reactions have been designed to provide highly discriminating tools for human identification. They include a growing number of STR loci [5–12]. Moreover, multiplex polymerase chain reactions (PCRs) have been developed, which produce small amplicons to increase the likelihood of success in the analysis of highly degraded DNA samples [13–20]. In this context, we have developed a new multiplex which we call I-DNA1, capable of amplifying 15 loci in a single reaction, including all the STR loci in CODIS and ISSL, 9 of those in the ECL and 7 of those in the 8 GCL (the exception is SE33).

I-DNA1 simultaneously amplifies seven mini-STR loci plus amelogenin in the range of 50–173 bp (TH01, TPOX, CSF1PO, D13S317, D16S539, VWA and D5S818) and seven midi-STR loci in the range of 151–297 bp (D3S1358, D7S820, D8S1179, HUMFIBRA (FGA), D18S51, D21S11 and D19S433).

We have validated I-DNA1 in accordance with the revised guidelines issued by the Scientific Working Group on DNA Analysis Methods (SGWDAM) [21]. This validation consisted of primer set optimisation tests, concordance studies, sensibility, analysis of degraded DNA samples, determination of stutter percentage, heterozygous peak height ratio (PHR) and mixtures. The results

show that I-DNA1 is a highly versatile, robust, sensitive tool for obtaining human genetic profiles.

Materials and methods

Human DNA control samples

We used 9947A DNA controls from the AmpFlSTR® YFiler kit (Applied Biosystems, Foster City, CA, USA), DNA control 007 (Applied Biosystems, Foster City, CA, USA) and K562 (Promega® Corporation, USA) to set up the PCR conditions and run sensibility tests. These three samples were quantified using a Quantifiler® Human DNA Quantification Kit (Applied Biosystems, Foster City, CA, USA) in an ABI PRISM® 7,000 Sequence Detection System (Applied Biosystems, Foster City, CA, USA), following the specifications of the manufacturer.

Population samples

Peripheral blood samples were taken from 600 healthy individuals: 318 caucasoid Europeans (Basque Country, Spain) and 282 individuals from Colombia (133 negroid and 149 hispanic).

Degraded DNA samples

Thirty formalin-fixed and paraffin-embedded tissue blocks from heart and liver autopsy tissue, dating back to 1980, were obtained from the Anatomical Pathology Department of the Marqués de Valdecilla University Hospital, Spain. The blocks were prepared with the same size of fresh tissue (1-cm wide, 1-cm long, and 0.3 cm of thickness). They were all fixed in 10% buffered formalin. The average period of fixation was 96 h at room temperature. The time between autopsy and tissue sampling was less than 24 h.

DNA extraction and quantification

The DNA extracts from the biological samples taken from hispanic and negroid individuals were obtained by the salting-out method [22]. The DNA from the caucasoid individuals was obtained by proteolytic lysis with proteinase K and organic extraction. DNA from formalin-fixed and paraffin-embedded tissue blocks was extracted by an initial deparaffinisation with xylene, followed by an ethanol rinse and thorough drying. The resulting deparaffinised samples were processed by proteolytic lysis with proteinase K and organic extraction. All the DNA extracts were quantified with PicoGreen® (Invitrogen), and DNA samples from paraffin-embedded tissue were also quantified with Quantifiler® (Applied Biosystems, Foster City, CA, USA).

Primer set optimisation

The PerlPrimer programme (http://perlprimer.sourceforge.net/) was used to design new primers that amplified 14 STR loci (D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D19S433, D21S11, CSF1PO, FGA, TH01, TPOX and vWA) plus amelogenin based on the reference sequences whose Genbank accession numbers are shown in Table 1. The regions flanking the STRs analysed were studied by means of SNPblast (www.ncbi.nlm.nih.gov/SNP/snpblastByChr.html) so as to avoid variable positions in the region where all the primers designed come together. The amplicons extend from 49 to 297 bp, and cover a large number of the alleles on the STRBase website [23]. To prevent incomplete adenylation, we added both a guanine to the 5′ end of the primer reverse and pigtailing [10, 18, 24] (Table 1). These primers are protected by patents P201031269 and P201031270.

The intensity of the amplicons and the absence of non-specific amplicons for each locus were checked by analysing the products of PCR by denaturing high-performance liquid chromatography (DHPLC), using a DNAsep Cartridge (Transgenomic® WAVE® System 4500, Glasgow, UK). Amplified product (10 μl) was migrated at 40°C on a linear gradient from 38.6% to 61.1% of buffer B (25% acetonitrile and TEAA 0.1 M) for 10 min.

PCR amplification, electrophoresis and data analysis

We used 5 μl of Qiagen® Master Mix from a Qiagen® Multiplex PCR Kit (Qiagen, Valencia, CA, USA), 4 μl of the primer mix and 1 μl of DNA template (1 ng/μl). The concentrations of each pair of primers are shown in Table 1. Amplification performance was tested in various thermocyclers: Biorad™ ICycler, C1000™ and Mycycler™ (BioRad, Hercules, CA, USA) and a GeneAmp® 9700 PCR System (Applied Biosystems, Foster City, CA, USA), under different annealing temperature conditions: 57, 58.9, 60.5, 61.8 and 63°C, with different numbers of cycles: 28, 29, 30, 31 and 32 and with final extension times of 30, 45, 60 and 90 min. Two microliter of each amplified product was analysed, mixed with 9 μl of Hi-Di™ Formamide (Applied Biosystems®, Foster City, CA, USA) and 0.5 μl of GeneScan™ 500 LIZ® Size Standard (Applied Biosystems®, Foster City, CA, USA). Following denaturation of the amplified products (6 min at 95°C), they were cooled (4 min at 4°C) and separated in an ABI PRISM 3130 Genetic Analyzer (Applied Biosystems®, Foster City, CA, USA). The results of the electrophoresis were analysed using version 4.0 Genmapper software (Applied Biosystems®, Foster City, CA, USA).

Table 1 Genbank accession, allele range, fragment sizes, primer sequence and concentration for each locus of I-DNA1

Locus	Genbank accession	Allele range	PCR fragment (bp)	Primer sequence[a] (5′–3′)	[Primer] (μm)
CSF1PO	X14720	6–15	75–111	6FAM-ACTGCCTTCATAGATAGAAGAT **G**CCCTGTTCTAAGTACTTCCT	0.08
D5S818	AC008512	7–17	127–167	6FAM-TTAATAGCAAGTATGTGACAAGG **G**ACATTTGTATCTTTATCTGTATCCTT	0.19
D7S820	AC004848	6–15	183–219	6FAM-GAATTATAACGATTCCACATTTATCC ATAAAGGGTATGATAGAACACTTG	0.13
D21S11	M84567	24–38	240–296	6FAM-CCCTGATTCTTCAGCTTGTA **G**CAGAGACAGACTAATAGGAGG	0.1
TPOX	M68651	6–16	61–101	NED-CTTAGGGAACCCTCACTG **G**CAGCGTTTATTTGCCCAA	0.08
VWA	M25858	11–24	121–173	NED-TCAGTATGTGACTTGGATTGA **G**TAGGTTAGATAGAGATAGGACAGA	0.07
D8S1179	AF216671	8–20	181–229	NED-TTGTATTTCATGTGTACATTCGT TTGTGTTCATGAGTATAGTTTCAC	0.09
D19S433	G08036	5.2–18.2	245–297	NED-CATGTTGGCACATTCCTG AGTTCTTTAGCAGTGATTTCTG	0.19
TH01	D00269	5–14	57–93	VIC-AACACAGACTCCATGGTG **G**TTCCTCCCTTATTTCCCT	0.04
D16S539	AC024591	5–16	105–149	VIC-CCTCTTCCCTAGATCAATACA ATCTGTAAGCATGTATCTATCATC	0.11
D3S1358	NT_005997	9–20	161–205	VIC-TGTAGTGAGCTATGATTCCC **G**TATTCCCTGTGCCTTTG	0.09
D18S51	AP001534	7–27	213–293	VIC-GTCTCAGCTACTTGCAGG **G**GAGATGTCTTACAATAACAGTTG	0.38
D13S317	AL353628	5–17	72–120	PET-CGCCTATCTGTATTTACAAATACAT **G**GACAGAAAGATAGATAGATGATTGAT	0.06
Amelogenin	Sullivan et al. [34].	X, Y	121,127	PET-CCCTGGGCTCTGTAAAGA **G**GGCTTGAGGCCAACCAT	0.11
FGA	M64982	16–51.2	151–293	VIC-CTCACAGATTAAACTGTAACCA TTGTCTGTAATTGCCAGC	0.31

[a] These primers are protected by patents P201031269 and P201031270

Allelic ladder

The allelic ladder was developed by selecting and mixing individuals with different genotypes and by amplifying them all together. Each allele included in the allelic ladder was confirmed by Identifiler® (Applied Biosystems, Foster City, CA, USA) in an AB3130 Genetic Analyzer (Applied Biosystems), using the GenMapper 4.0 programme by Applied Biosystems.

Precision and accuracy

The precision of the allelic ladder from I-DNA1 was calculated from multiple injections of ladder run in an ABI PRISM 3130 Genetic Analyzer (Applied Biosystems®, Foster City, CA, USA). To assess sizing accuracy, the standard deviation in size in pairs of bases was calculated for each sample allele with regard to the corresponding allelic ladder allele size (±0.40 bp) produced in Genmapper (Applied Biosystems, Foster City, CA, USA) for 162 caucasoid individuals, amplified under standard conditions.

Concordance studies

A comparison was run between genetic profiles analysed with I-DNA1 and Identifiler®. For analysis via I-DNA1 1 ng of DNA template from each individual was used. The Identifiler® analysis was conducted under the optimum conditions suggested by the manufacturer.

Sensibility

To determine the sensibility of I-DNA1, the 9947A DNA control samples from the AmpFlSTR® YFiler kit (Applied Biosystems, Foster City, CA, USA), Control DNA 007 (Applied Biosystems, Foster City, CA, USA) and K562 (Promega® Corporation, USA) were analysed in triplicate. DNA concentrations of 30 ng, 20 ng, 10 ng,

1.6 ng, 1 ng, 400 pg, 200 pg, 100 pg, 50 pg and 25 pg were used.

Analysis of degraded DNA samples

Previous studies have suggested that DNA extracted from formalin-fixed and paraffin-embedded tissue may be highly degraded [25]. To test the ability of I-DNA1 to analyse highly degraded samples, we selected 30 formalin-fixed and paraffin-embedded tissue blocks, which were 30 years old. DNA samples extracted from these samples were analysed by Identifiler® and I-DNA1. Each PCR reaction was performed in the same number of cycles (30 cycles) and in the same concentrations of template DNA (1 ng).

Determination of stutter percentage, heterozygous peak height ratio and mixtures

The percentage of stutters in homozygotic and heterozygotic individuals differing in size by more than 4 bp was determined by calculating the stutter peak height as a percentage of the true allele height. The PHR was calculated by dividing the lower of the peaks by the height of the higher for heterozygotic individuals at each locus.

Mixtures were studied by establishing a threshold based on the stutter percentage to distinguish between stutter and the minor contributor. To check the efficacy of that threshold, two DNA samples were mixed with the following ratios: 1:1, 1:3, 1:5, 1:7, 1:10, 1:15, 1:20, 3:1, 5:1, 7:1, 10:1, 15:1 and 20:1, with the quantity of final template DNA being held at 1 ng.

Results and discussion

I-DNA1 is a multiplex system for the PCR amplification of 14 STR loci plus amelogenin. Its design combines the amplification of seven miniSTRs, seven midiSTRs and amelogenin, using five-dye chemistry. This enables 15 loci to be analysed in a single reaction, the PCR products of which are no bigger than 297 bp. I-DNA1 covers all the STR loci in CODIS plus D19S433, which makes I-DNA1 a highly discriminating system for identifying humans.

The results for the validation of IDNA in primer set optimisation tests, PCR amplification parameter optimisation, precision and accuracy, concordance studies, sensibility, analysis of degraded DNA samples, determination of stutter percentage, heterozygous PHR and mixtures are discussed below.

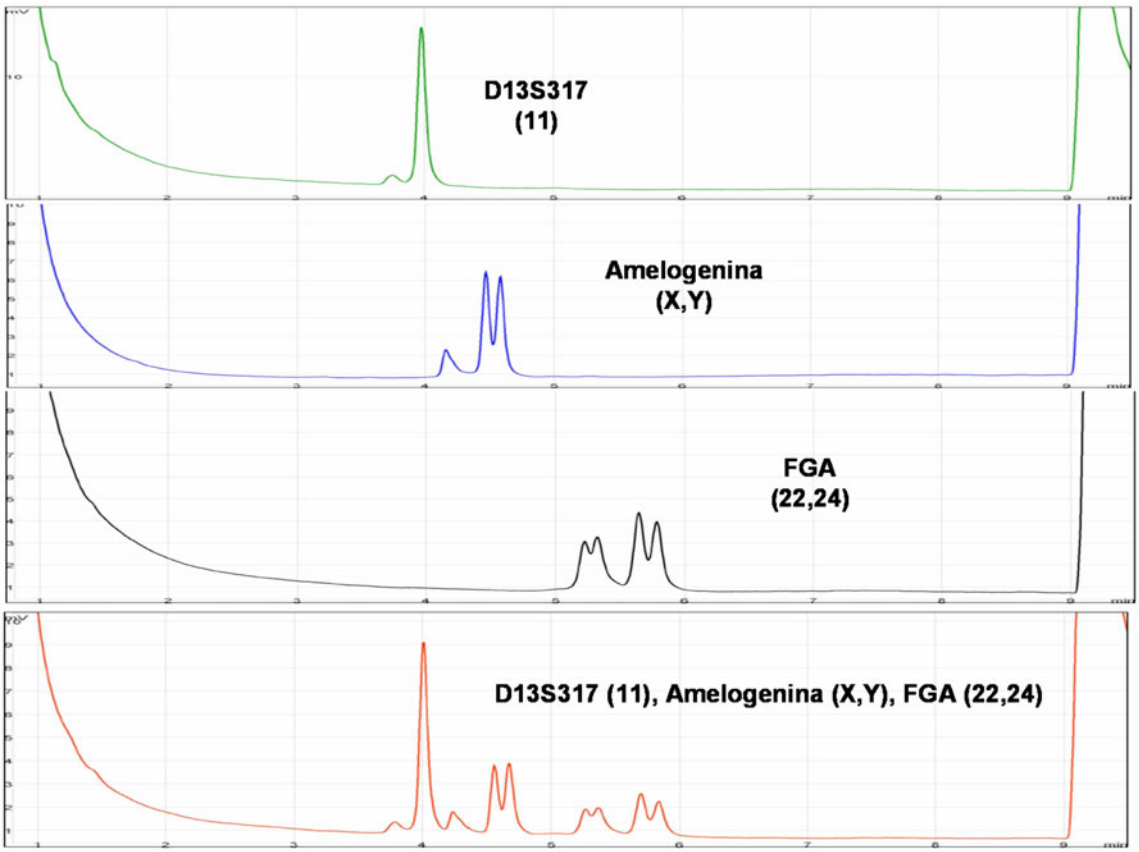

Fig. 1 Chromatograms showing the spectra obtained after DHPLC analysis of the amplificates for D13S317 (**a**), Amelogenin (**b**) and FGA (**c**) in singleplex reactions and jointly in the triplex reaction (**d**)

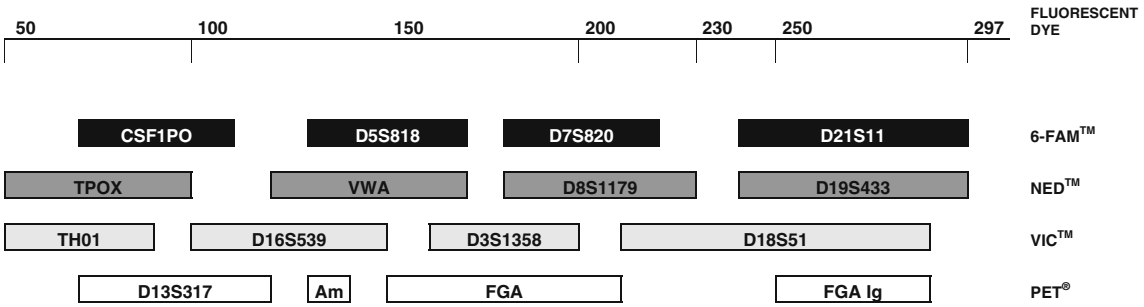

Fig. 2 Design strategy of I-DNA1. The *boxes* represent the expected fragment sizes of each locus. Loci labelled with different fluorescent dyes are shown in different *grey scales*. The respective dyes are assigned on the right side of the figure

Primer set optimisation

Ninety-six primers were designed, from which those capable of producing the smallest amplicons (no bigger than 300 bp in all cases) were selected, with melting temperatures between 57.5°C and 62.5°C, hybridising in regions without SNPs (single nucleotide polymorphisms) and Indels (insertions/deletions) described to date in National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/snp/). In the first stage of optimisation, singleplex PCR amplifications of each locus were produced.

Primer set optimisation consisted of analysing the amplified products both separately and in multiplex form. The usual procedure is to analyse amplicons marked with a DNA genetic analyser [26]. However, here, DHPLC technology was used to assess the intensity of amplicons and check that there were no non-specific amplifications. The advantage of DHPLC is that there is no need for prior marking of primers.

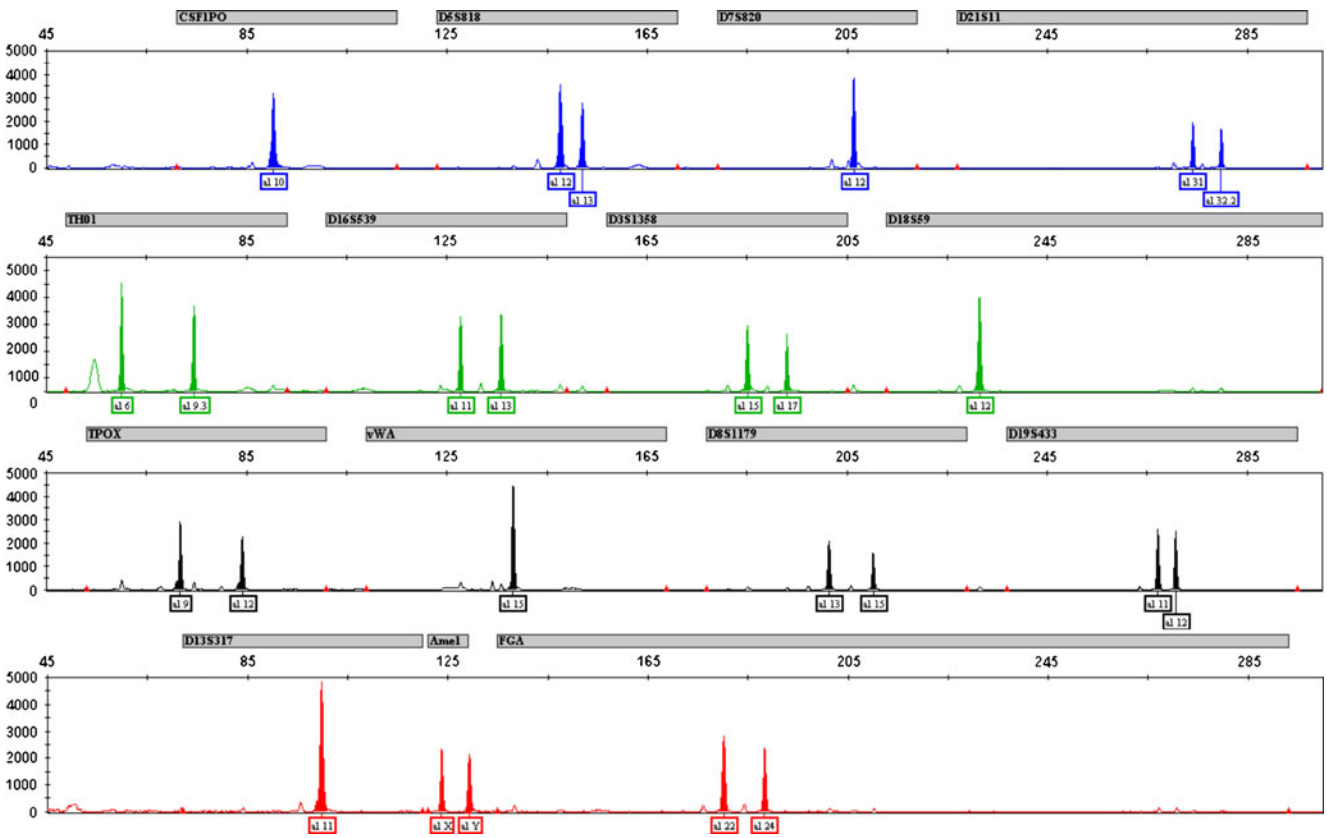

Fig. 3 Typical electropherogram of 1-ng amplified DNA. *Signals* represent the alleles indicated below the respective peaks. *Boxes* above peaks indicate each locus name

This method was used in the first (separate) phase of PCR product optimisation as shown in Fig. 1a, b and c. None of the pairs of primers selected in this paper is used in any of the kits currently validated. Table 1 shows primer sequences and a broad range of alleles for each locus, in which the main population groups are considered as per the information on STRbase [23].

In the second stage of optimisation, various sets of primers were tested via DHPLC in a multiplex format (triplex and quadruplex reaction) with the additional aim of checking that there is no overlap of PCR products. The ranges of the alleles of the adjacent loci were spaced in size, so as to avoid overlaps between alleles from different loci, and thus ensure correct genotyping, as shown in Fig. 1d.

Finally, three quadruplex PCR reactions (quadruplex A: CSF1PO, D5S818, D7S820 and D21S11; quadruplex B: TH01, D16S539, D3S1358 and D18S51; quaduplex C: TPOX, vWA, D8S1179 and D19S433) and one triplex PCR reaction (triplex D: D13S317, Amelogenin and FGA) were selected. Once the best set of primers for a single multiplex amplification of the 14 STR loci and amelogenin was determined, the primers were marked with fluorochromes. The forward primers of the three quadruplex reactions A, B and C were marked with 6-FAM™, VIC™ and NED™, respectively, and the forward primers of the triplex reaction D were marked with PET® (Applied Biosystems, Foster City, CA, USA) (Table 1). Thus, only those primers actually included in the final reaction referred to as I-DNA1 were marked. Figure 2 shows the final result of the strategy in the design of primers for I-DNA1.

Using DHPLC technology in optimising the amplification of multiplex reactions is, therefore, a fast, inexpensive alternative [27] to analyse amplicons via a genetic analyser, as it does away with the need to use primers marked in the initial phases of optimisation [28].

PCR amplification parameter optimisation

Amplification parameters were optimised by assessing the peak height data, the balance between them in heterozygosis and incomplete adenylation obtained from the electropherograms of the PCR products.

The concentration of each pair of primers was tested between 0.0375 μM and 0.75 μM. Peak height was observed to increase, as primer concentration increased. Moreover, at high concentration imbalance in heterzygotes, artefacts FAM87, FAM96, FAM158, VIC53, VIC105, VIC121 and VIC180 and incomplete adenylation were observed.

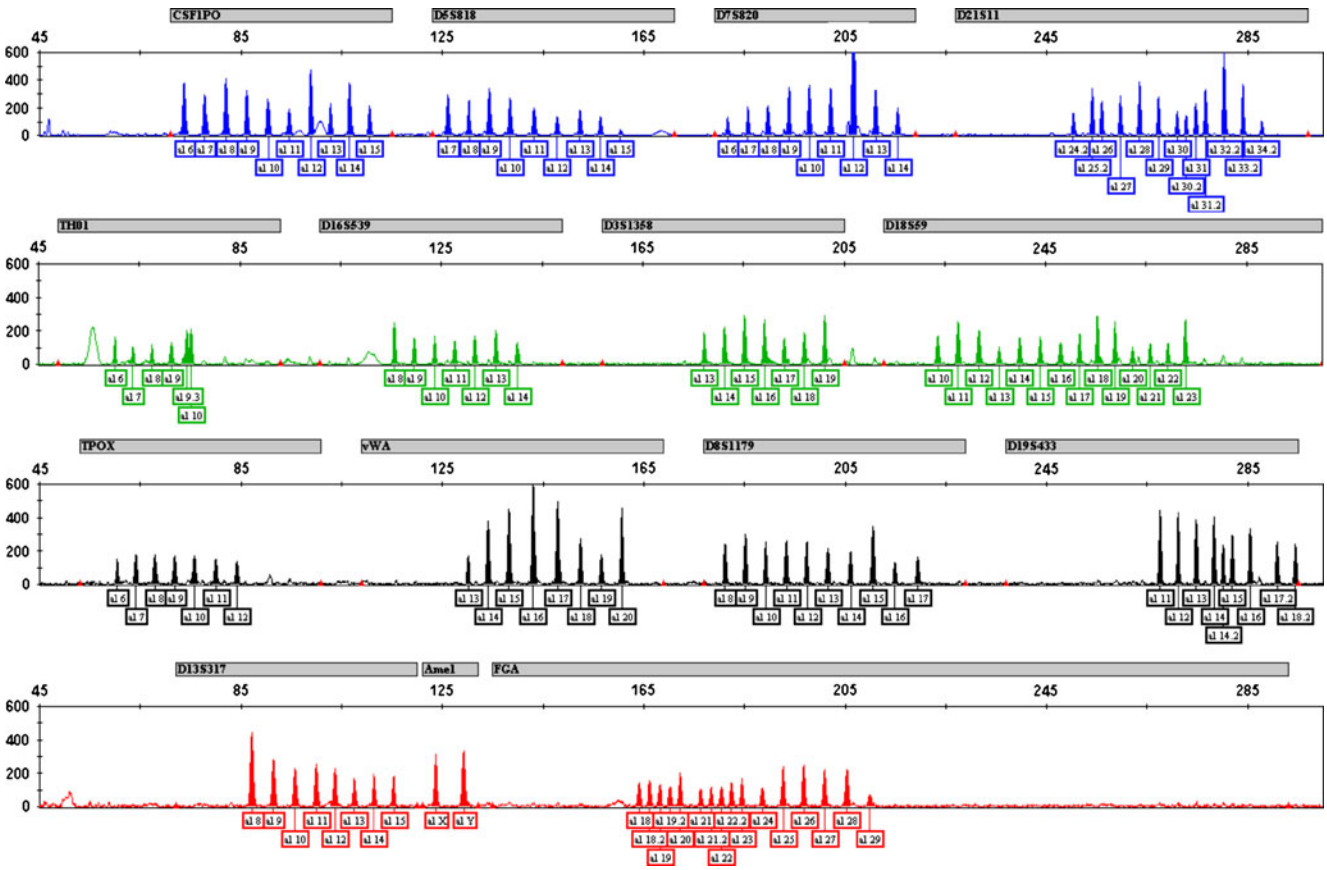

Fig. 4 Allelic ladder for the I-DNA1 multiplex system. The allele designation of each locus is assigned under the respective signals

The annealing temperature of the multiplex PCR reaction was then optimised. The resulting increase in annealing temperature gave rise to a gradual decrease in incomplete adenylation, peak height and imbalance in heterozygotes. The best results were obtained with an annealing temperature of 60.5°C.

Increasing the number of cycles resulted in a general increase in peak height and incomplete adenylation, but had no significant effect on the heterozygote balance. The most balanced results were obtained at 30 cycles. In an attempt to eliminate incomplete adenylation, the final extension time was increased, guanine was added to the 5′ end of the reverse primer (marked as G in boldface, Table 1), and pig-tailing was also added. Increasing final extension time resulted in no significant changes, but adding a guanine to the 5′ end of the reverse primer eliminated incomplete adenylation completely, and worked, in fact, even better than adding pig-tailing. No significant variations in the quality of the results appeared when different thermocyclers were used.

The intensities of the markers analysed by I-DNA1 were balanced at around 2,500 RFUs in the analysis of 1 ng of DNA to strike a balance between signal intensity and absence of overlap between signals corresponding to different fluorochromes: 6-FAM™ (2,243±764 RFUs), VIC™ (2,742±748 RFUs), NED™ (2,313±426 RFUs) and PET® (2,417±247 RFUs).

Once the above parameters had been assessed, it was possible to determine the optimum primer concentration (Table 1) for a single, balanced, specific multiplex reaction capable of amplifying 14 STR loci plus amelogenin from 1 ng of template DNA (Fig. 3). The amplification conditions were as follows: one initial denaturing cycle lasting 15 min. at 95°C, 30 × 30 s/cycles at 95°C, 90 s. at 60.5°C and 1 min. at 72°C, followed by a final extension cycle of 30 min. at 60°C.

Allelic ladder

The allelic ladder was developed by selecting and mixing individuals with different genotypes to include as many alleles as possible. In step one, DNA samples from selected individuals were mixed and amplified individually in simplex reactions for each locus. In step two, amplified fragments from each 15 loci were mixed and adjusted to give balanced signal intensity for all loci. Rare alleles for the high molecular weight fragments of the FGA were not included in the allelic ladder, because they are not present in the samples analysed in this study (Fig. 4).

Precision and accuracy

Allelic ladder precision was assessed from injections of ladder in independent electrophoresis runs. The standard deviation (SD) of the mean was calculated and shown to be within ≤0.07 bp.

Accuracy was assessed by analysing 162 individuals to measure the deviation of each sample allele size from the corresponding allelic ladder. No alleles outside the ±0.40 bp range in the allelic ladder were detected in any of the loci studied.

Preliminary concordance studies

Concordance studies were carried out by comparing 2,430 allele calls obtained by I-DNA1 and Identifiler®. Identifiler® was selected because it is widely used [29] and widely validated [7]. Only one discordance was found, which gives a concordance level of 99.9% between the results from the two kits. That single discordance was for an individual genotyped as homozygotic (15/15) for

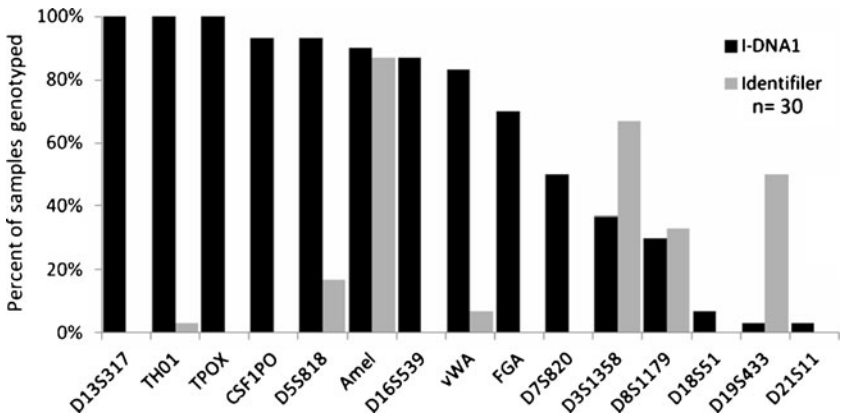

Fig. 5 Percent of DNA samples from 30-year-old paraffin-embedded samples (n=30) genotyped successfully for each locus with I-DNA1 (*black columns*) and/or Identifiler ® (Applied Biosystems, Foster City, CA, USA) (*grey columns*). Loci are distributed from left to the right according to their amplicon size which are analysed with I-DNA1. Direct correlation between amplification efficiency and amplicon size can be established. These results demonstrated that I-DNA1 provide higher throughput than Identifiler ® (Applied Biosystems, Foster City, CA, USA), analysing reduced size DNA fragments

Table 2 Relative stutter intensities (%) of STR loci analysed with I-DNA1

I-DNA1	Average	Median	SD	Minimun	Maximun
CSF1PO	6.3	5.4	2.6	2.7	13.0
D13S317	6.2	6.1	2.3	2.0	13.2
D16S539	9.4	9.3	1.3	6.6	12.6
D18S59	8.6	8.5	2.0	5.2	12.7
D19S433	8.5	8.4	1.4	5.9	12.5
D21S11	11.4	11.2	2.2	7.7	16.9
D3S1358	10.1	10.2	1.9	6.1	14.1
D5S818	5.7	6.7	3.4	5.0	10.9
D7S820	7.2	7.0	2.1	2.8	12.4
D8S1179	9.2	8.9	2.2	5.2	15.0
FGA	8.4	8.4	2.0	3.9	12.9
TH01	4.0	3.8	2.1	0.8	11.2
TPOX	4.2	3.9	1.6	1.0	9.3
VWA	8.7	9.2	3.1	1.0	14.6

SD standard deviation

D3S1358 by Identifiler®, and as heterozygotic (12/15) by I-DNA1. The sample was analysed in triplicate by both systems, and the results were the same. Similar discordances reported in other studies have proved to be due to mutations which probably coincided with the 3′ end binding region of one of the primers [30, 31].

Sensibility

The sensibility of an STR multiplex is determinant in obtaining STR profiles with a sufficient probability of identity [32]. This study, therefore, sought experimental conditions in which I-DNA1 would provide high levels of sensibility. I-DNA1 enables 14 STR loci plus amelogenin to be amplified from as little as 100 pg of DNA, and even enabled 12 loci to be analysed with just 25–100 pg of DNA, while the remaining three STR markers (D5S818, D3S1358 and D18S51) showed allele losses due to stochastic effects. The sensibility of I-DNA1 was compared with that of other multiplex systems that include at least 13 STR loci from CODIS: PowerPlex16 system (Promega® Corporation, USA) and Identifiler®. Both these systems require larger quantities of template DNA than I-DNA1 (250 pg [7, 8]).

Degraded DNA samples

The size of amplicons in a STR multiplex system is critical for obtaining genetic profiles from highly degraded biological samples [32]. I-DNA1 was specifically designed to obtain genome profiles from small fragments of DNA. In this study, we analysed DNA from 30-year-old paraffin-embedded samples (n=30) with I-DNA1 and Identifiler® widely used for analysing DNA from highly degraded samples [16]. The average number of STR loci typed successfully, using I-DNA1 was 9.5±2.8 loci. In contrast, analysis of the samples with Identifiler® yielded only typeable results with an average of 2.63±1.82 loci (Fig. 5).

In all samples, a direct relationship between amplification efficiency and the size of PCR product fragments was evident. Due to their reduced size (50–173 bp), the seven miniSTRs included in I-DNA1 were successfully genotyped in more than 80% of the samples. However, allele typing tended to encounter more difficulties in STR loci longer than 200 bp and was unsuccessful in loci longer than approximately 300 bp.

These results show that I-DNA1 is more successful in typing degraded samples than Identifiler®.

Table 3 Peak height ratio of STR loci analysed with I-DNA1

I-DNA1	Number of observations	Mean	Median	SD	Minimun	Maximun
CSF1PO	390	84.2	86.3	11.4	49.0	100.0
D13S317	437	75.8	77.2	14.8	40.7	99.8
D16S539	429	92.1	94.1	7.2	63.7	100.0
D18S59	473	84.0	85.6	8.5	55.9	100.0
D19S433	430	81.7	83.7	8.9	51.7	99.6
D21S11	445	85.7	87.4	9.9	49.1	99.6
D3S1358	424	89.1	90.3	7.5	58.8	100.0
D5S818	399	84.2	85.7	10.4	52.7	100.0
D7S820	430	89.2	91.0	9.2	58.0	100.0
D8S1179	421	81.9	84.1	11.2	50.5	99.7
FGA	475	88.0	89.7	8.0	59.8	99.8
TH01	436	90.9	94.6	9.7	57.2	100.0
TPOX	359	88.2	90.6	10.1	54.3	99.9
vWA	457	87.1	88.1	8.3	60.2	99.9

SD standard deviation

Determination of stutter percentage, heterozygous peak height ratio and mixtures

The stutter peaks that result from strand slippage usually show up as peaks 4 bp lower than the true allele. A high percentage of stutter may lead to mixtures being misinterpreted due to confusion with the minor contributor [33]. The results for each locus after 162 caucasoid individuals who had been analysed are shown in Table 2. In all loci it was observed that the percentage of stutter increased as the number of repetitions increased. In general, the percentage of stutter obtained with I-DNA1 is slightly higher than that reported for the Minifiler™ PCR Amplification Kit (Applied Biosystems, Foster City, CA, USA) [16] and slightly lower than for *genRES*MPX-3 [5]. The lowest average percentage of stutter was observed for locus TH01 (4.0±2.1), which coincides with the results reported by Schlenk et al. [5], and the highest was in D21S11 (11.4±2.2). Altogether, the low percentages of stutter obtained for I-DNA1 help to ensure the correct genotyping of samples and the proper interpretation of mixtures.

PHR levels define the balance that exists in the amplification of heterozygotes. For I-DNA1, the calculation of the PHR gave low-standard deviation levels and average values very close to the mean for all the loci analysed in 600 individuals (Table 3). In general, I-DNA1 gives a good balance of alleles in heterozygosis. The only exception was locus D13S317, with a mean of 75.8% and a standard deviation of 14.8. The average PHR was similar to that reported for the Minifiler™ Kit PCR and Identifiler® kits (85.86%, 87.87% and 86.87%, respectively).

The DNA samples for the mixture test were selected for their high heterozygosity, in an attempt to prevent alleles of the minor contributor from being located in stutter positions of the alleles of the major contributor. The threshold for distinguishing between the minor contributor and stutter for the major contributor in each locus analysed was set at the mean percentage of stutter plus three times the standard deviation. With this threshold, it proved possible to genotype the minor contributor at 1:10 ratios in all the STR loci included in I-DNA1 except D19S433 and D7S820, which we managed to genotype at a ratio of 1:7. The minor contributor was genotyped up to ratios of 1:15 in markers CSF1PO, D5S818 and D8S1179, and as far as 1:20 in TH01 and TPOX.

The low percentage of stutter, the high PHR and the low standard deviation of the system mean that in mixtures from two contributors I-DNA1 enables a minor contributor to be detected at 1:10. I-DNA1, therefore, seems eminently suited to determining the genotypes that make up a mixture and estimating their proportions.

Conclusion

In short, the validation of I-DNA1 has shown that is a highly sensitive, robust system for obtaining individual genetic profiles. The small size of the amplicons in I-DNA1 and the high sensibility of the system show that the likelihood of identification with this system is greater than with other STR multiplex reactions when the samples analysed are highly degraded. In addition, the low percentage of stutter and the high PHR of I-DNA1 mean that this system has a high capacity for analysing mixtures. In all, I-DNA1 is, potentially, a useful tool for human identification.

Acknowledgements This work was funded by the Basque Government (IT-424/07). Technical support and staff time provided by DNA Bank of University of Basque Country from SGIker (UPV/EHU, MICINN, GV/EJ, ESF), as well as Basque BioBank of BIO foundation (BIOEF) are gratefully acknowledged. JMA and AO are supported by grants from the Basque Government (Dpto. de Educación, Universidades e Investigación).

References

1. Butler JM (2007) Short tandem repeat typing technologies used in human identity testing. Biotechniques 43(4):ii–v
2. Martin PDSH, Schneider PM (2001) A brief history of the formation of DNA databases in forensic science within Europe. Forensic Sci Int 119:225–231
3. LS AC (2004) International perspectives on forensic DNA databases. Forensic Sci Int 146(Suppl):S119–S121
4. C-R F (2004) Forensic DNA databases. Forensic Sci Int 146 (Suppl):S143–S144
5. Schlenk J, Seidl S, Braunschweiger G, Betz P, Lederer T (2004) Development of a 13-locus PCR multiplex system for paternity testing. Int J Legal Med 118(1):55–61
6. Cotton EA, Allsop RF, Guest JL, Frazier RR, Koumi P, Callow IP, Seager A, Sparkes RL (2000) Validation of the AMPFISTR SGM plus system for use in forensic casework. Forensic Sci Int 112(2–3):151–161
7. Collins PJ, Hennessy LK, Leibelt CS, Roby RK, Reeder DJ, Foxall PA (2004) Developmental validation of a single-tube amplification of the 13 CODIS STR loci, D2S1338, D19S433, and amelogenin: the AmpFISTR Identifiler PCR Amplification Kit. J Forensic Sci 49(6):1265–1277
8. Krenke BE, Tereba A, Anderson SJ, Buel E, Culhane S, Finis CJ, Tomsey CS, Zachetti JM, Masibay A, Rabbach DR, Amiott EA, Sprecher CJ (2002) Validation of a 16-locus fluorescent multiplex system. J Forensic Sci 47(4):773–785
9. Coticone SR, Oldroyd N, Philips H, Foxall P (2004) Development of the AmpFISTR SEfiler PCR amplification kit: a new multiplex containing the highly discriminating ACTBP2 (SE33) locus. Int J Legal Med 118(4):224–234
10. BJ HCR, Vallone PM (2009) A 26plex autosomal STR assay to aid human identity testing. J Forensic Sci 54(5):1008–1015
11. Wallin JM, Buoncristiani MR, Lazaruk KD, Fildes N, Holt CL, Walsh PS (1998) TWGDAM validation of the AmpFISTR blue PCR amplification kit for forensic casework analysis. J Forensic Sci 43(4):854–870
12. LaFountain MJ, Schwartz MB, Svete PA, Walkinshaw MA, Buel E (2001) TWGDAM validation of the AmpFISTR Profiler Plus

and AmpFISTR COfiler STR multiplex systems using capillary electrophoresis. J Forensic Sci 46(5):1191–1198
13. Grubwieser P, Mühlmann R, Berger B, Niederstätter H, Pavlic M, Parson W (2006) A new "miniSTR-multiplex" displaying reduced amplicon lengths for the analysis of degraded DNA. Int J Legal Med 120(2):115–120
14. Müller K, Klein R, Miltner E, Wiegand P (2007) Q8–a short amplicon multiplex including the German DNA database systems. Forensic Sci Int Genet 1(2):205–207
15. Klaus Bender PMS (2006) Validation and casework testing of the BioPlex-11 for STR typing of telogen hair roots. Forensic Sci Int 161(1):52–59
16. Mulero JJ, Chang CW, Lagace RE, Wang DY, Bas JL, McMahon TP, Hennessy LK (2008) Development and validation of the AmpFISTR MiniFiler PCR Amplification Kit: a MiniSTR multiplex for the analysis of degraded and/or PCR inhibited DNA. J Forensic Sci 53(4):838–852
17. Müller K, Klein R, Miltner E, Wiegand P (2007) Improved STR typing of telogen hair root and hair shaft DNA. Electrophoresis 28 (16):2835–2842
18. Butler JM, Shen Y, McCord BR (2003) The development of reduced size STR amplicons as tools for analysis of degraded DNA. J Forensic Sci 48(5):1054–1064
19. Coble MD, Butler JM (2005) Characterization of new miniSTR loci to aid analysis of degraded DNA. J Forensic Sci 50(1):43–53
20. Wiegand P, Kleiber M (2001) Less is more–length reduction of STR amplicons using redesigned primers. Int J Legal Med 114(4–5):285–287
21. Ellegren H (2004) Microsatellites: simple sequences with complex evolution. Nat Rev Genet 5(6):435–445
22. Miller SA, Dykes DD, Polesky HF (1988) A simple salting out procedure for extracting DNA from human nucleated cells. Nucl Acids Res 16:1215
23. Ruitberg CM, Reeder DJ, Butler JM (2001) STRBase: a short tandem repeat DNA database for the human identity testing community. Nucleic Acids Res 29(1):320–322
24. Brownstein MJ, Carpten JD, Smith JR (1996) Modulation of non-templated nucleotide addition by Taq DNA polymerase: primer modifications that facilitate genotyping. Biotechniques 20 (6):1004–1006, 1008-1010
25. Farrugia A, Keyser C, Ludes B (2010) Efficiency evaluation of a DNA extraction and purification protocol on archival formalin-fixed and paraffin-embedded tissue. Forensic Sci Int 194(1–3): e25–e28
26. Tagliaro F, Pascali J, Fanigliulo A, Bortolotti F (2010) Recent advances in the application of CE to forensic sciences: a update over years 2007–2009. Electrophoresis 31(1):251–259
27. Shinka T, Naroda T, Tamura T, Sasahara K, Nakahori Y (2001) A rapid and simple method for sex identification by heteroduplex analysis, using denaturing high-performance liquid chromatography (DHPLC). J Hum Genet 46(5):263–266
28. Devaney JM, Girard JE, Marino MA (2000) DNA microsatellite analysis using ion-pair reversed-phase high-performance liquid chromatography. Anal Chem 72(4):858–864
29. Hill CR, Kline MC, Mulero JJ, Lagace RE, Chang CW, Hennessy LK, Butler JM (2007) Concordance study between the AmpFISTR MiniFiler PCR amplification kit and conventional STR typing kits. J Forensic Sci 52(4):870–873
30. Budowle B, Aranda XG, Lagace RE, Hennessy LK, Planz JV, Rodriguez M, Eisenberg AJ (2008) Null allele sequence structure at the DYS448 locus and implications for profile interpretation. Int J Legal Med 122(5):421–427
31. Zhai XD, Xue XQ, Mo YN, Zhao GS, Ai HW, Ye Y, Wang Z, Hou YP (2010) False homozygosities at CSF1PO loci revealed by discrepancies between two kits in Chinese population. Int J Legal Med 124(5):457–458
32. Budowle B, Eisenberg AJ, van Daal A (2009) Validity of low copy number typing and applications to forensic science. Croat Med J 50(3):207–217
33. Walsh PS, Fildes NJ, Reynolds R (1996) Sequence analysis and characterization of stutter products at the tetranucleotide repeat locus vWA. Nucleic Acids Res 24(14):2807–2812
34. Sullivan KM, Mannucci A, Kimpton CP, Gill P (1993) A rapid and quantitative DNA sex test: fluorescence-based PCR analysis of X-Y homologous gene amelogenin. Biotechniques 15(4):636–638, 640–631