Int J Legal Med (2000) 113 : 208–214

ORIGINAL ARTICLE

A. Pérez-Lezaun · F. Calafell · J. Clarimón · E. Bosch
E. Mateu · L.Gusmão · A. Amorim · N. Benchemsi
J. Bertranpetit

Allele frequencies of 13 short tandem repeats in population samples from the Iberian Peninsula and Northern Africa

Received: 15 March 1999 / Received in revised form: 25 June 1999

Abstract The 13 short tandem repeat (STR) loci D3S1358, vWA, FGA, D16S539, TH01, TPOX, CSF1PO, D8S1179, D21S11, D18S51, D5S818, D13S317 and D7S820 as well as the amelogenin locus, contained in AmpF*l*STR Profiler Plus and/or AmpF*l*STR Cofiler and/or AmpF*l*STR Green I PCR amplification kits, were studied in four populations from the Iberian Peninsula, Basques, Catalans, Andalusians and Portuguese and two North African populations (Moroccan Arabs and Berbers). The aim of the study was to obtain accurate allele frequency data and other genetic parameters of forensic interest on the main representative human groups living in Iberia and Morocco using an automated method and commercial amplification kits.

Key words STR · Microsatellite · Polymorphism · Iberia · North Africa

Introduction

The AmpF*l*STR Profiler Plus and AmpF*l*STR Cofiler PCR amplification kits provide an easily reproducible and fast laboratory tool for typing the most widely used short tandem repeat loci (STRs) in forensic applications. This specific set of markers is, in fact, the set of STRs that have been approved in the combined DNA index system (CODIS) database in the USA.

A. Pérez-Lezaun · F. Calafell · J. Clarimón · E. Bosch · E. Mateu
J. Bertranpetit (✉)
Unitat de Biologia Evolutiva,
Facultat de Ciències de la Salut i de la Vida,
Universitat Pompeu Fabra, Doctor Aiguader 80,
E-08003 Barcelona, Catalonia (Spain)
e-mail: jaume.bertranpetit@cexs.upf.es,
Tel.: +34-93-5422840, Fax: +34-93-5422802

L. Gusmão · A. Amorim
Instituto de Patologia e Imunologia Molecular,
University of Porto, Porto, Portugal

N. Benchemsi
Centre National de Transfusion Sanguine, Rabat, Morocco

Despite the fact that a considerable amount of information on the allele frequencies of some of these STRs is available in Iberian populations (Cabrero et al. 1995; Pestoni et al. 1995; Alonso et al. 1995; Pancorbo et al. 1996; Iriondo et al. 1999), allele frequencies for the complete set of STRs presented here seems to be so far available only for one Italian population (Garofano et al. 1998) and general US ethnic groups.

The analysis presents data on six population groups from the Iberian Peninsula and Northern Africa. Four Iberian populations were included, namely Basques which seem to be an outlier population in the European genetic continuum (Calafell and Bertranpetit 1994; Comas et al. 1998; Torroni et al. 1998), Catalans, Andalusians and Northern Portuguese. The survey also includes two populations from North Africa, Northern Berbers and Moroccan Arabs, which nowadays represent an important source of foreign immigration to Spain.

Material and methods

Between 64 and 100 chromosomes were analysed for each marker and population under study. Iberian samples were from Catalonia (Girona province), the Basque Country (several towns and villages within the Gipuzkoa province), Andalusia (including several provinces) and Northern Portugal (Porto region). Both North African samples were of Moroccan origin. The Arab sample included 20 immigrant individuals living in the Barcelona area and 30 individuals collected in central Morocco. The Berber sample comprised 50 individual samples collected in North East Morocco (Oujda and Nador). Special care was taken in the assessment of the origin of the individuals included by choosing those whose four grandparents were born in the same region. In all cases DNA was extracted from fresh blood from autochthonous blood donors using a standard phenol-chloroform DNA extraction method.

The loci analysed in the study were those included in the AmpF*l*STR Green I, AmpF*l*STR Cofiler and AmpF*l*STR Profiler Plus PCR amplification kits. The tetranucleotide repeat systems D3S1358 (Li et al. 1993), vWA (Kimpton et al. 1992), FGA (Mills et al. 1992), TH01 (Edwards et al. 1992), TPOX (Anker et al. 1992), CSF1PO (Hammond et al. 1994), D8S1179 (Oldroyd et al. 1995), D21S11 (Sharma and Litt 1992), D18S51 (Urquhart et al. 1995), D5S818 (Hudson et al. 1995), D13S317 (Hudson et al. 1995), D7S820 (Green et al. 1991), D16S539 [Cooperative Human

Table 1 Allele frequencies, heterozygosity (Het), polymorphism information content (PIC), power of discrimination (POD), chance of paternity exclusion if mother is known (CE) and chance of paternity exclusion if only one parent and child are typed (CE2) for the STR CSF1PO in six populations from the Iberian Peninsula and North Africa: Basques (Bas), Catalans (Cat), Andalusians (And), Portuguese (Por), Arabs (Ara) and Berbers (Ber)

CSF1PO (2n)	Bas 100	Cat 96	And 72	Por 78	Ara 94	Ber 100
7	0	0	0	0	0.011	0
8	0	0.010	0	0	0.011	0.030
9	0.010	0.021	0	0.026	0	0.020
10	0.280	0.229	0.347	0.308	0.330	0.310
11	0.380	0.333	0.306	0.308	0.309	0.340
12	0.260	0.302	0.236	0.295	0.309	0.230
13	0.050	0.094	0.097	0.064	0.032	0.070
14	0.010	0.010	0.014	0	0	0
15	0.010	0	0	0	0	0
Het	0.707	0.736	0.721	0.719	0.700	0.730
PIC	0.652	0.689	0.669	0.665	0.639	0.681
POD	0.860	0.884	0.870	0.867	0.849	0.879
CE	0.450	0.496	0.468	0.463	0.430	0.486
CE2	0.282	0.322	0.297	0.294	0.268	0.314

Table 2 Allele frequencies, heterozygosity (Het), PIC, POD, CE and CE2 for the STR D13S317 in six populations from the Iberian Peninsula and North Africa

D13S317 (2n)	Bas 100	Cat 98	And 70	Por 78	Ara 94	Ber 98
8	0.320	0.102	0.114	0.090	0.106	0.082
9	0.020	0.092	0.029	0.051	0.032	0.051
10	0.030	0.051	0.057	0.051	0	0.020
11	0.270	0.327	0.371	0.346	0.383	0.316
12	0.200	0.316	0.300	0.295	0.255	0.408
13	0.120	0.051	0.114	0.090	0.149	0.082
14	0.040	0.061	0.014	0.077	0.074	0.041
Het	0.767	0.765	0.742	0.766	0.748	0.715
PIC	0.731	0.732	0.702	0.733	0.711	0.672
POD	0.909	0.912	0.894	0.912	0.899	0.876
CE	0.552	0.561	0.521	0.563	0.530	0.488
CE2	0.374	0.382	0.343	0.384	0.351	0.314

Table 3 Allele frequencies, heterozygosity (Het), PIC, POD, CE and CE2 for the STR D16S539 in two populations from the Iberian Peninsula and North Africa

D16S539 (2n)	Por 78	Ara 94
8	0.026	0.032
9	0.167	0.138
10	0.090	0.043
11	0.256	0.309
12	0.282	0.191
13	0.167	0.234
14	0.013	0.053
Het	0.790	0.789
PIC	0.759	0.758
POD	0.924	0.924
CE	0.588	0.589
CE2	0.410	0.411

Linkage Center (CHLC), accession number 715; Genebank accession number G07925] and the sex-specific amelogenin locus (Sullivan et al. 1993). AmpF*l*STR Profiler Plus and AmpF*l*STR Green I amplification kits were used to test Basque, Catalan, Andalusian and Arab populations. The AmpF*l*STR Cofiler kit was used to analyse Portuguese and Berber populations. Therefore, all populations included in this study were tested for a total of 12 STRs and in addition the marker D16S539 was tested in Portuguese and Berber groups.

Amplifications were performed following the instructions provided in the kit user manual with the recommended DNA amount (1.0–2.5 ng) using a final PCR volume of 25 µl. Electrophoresis of amplified fragments was performed in a 377 ABI PRISM sequencer using 36/48-cm well-to-read plates. GeneScan 672 analysis software was used to track lanes and measure fragment sizes. Genotyper 2.1 × 3 software was used to automatically designate alleles by comparison to locus specific allelic ladders.

Allele frequencies were estimated by direct gene counting. Expected heterozygosity was estimated as $1\text{-}\Sigma p_i^2$ where p_i is the frequency of the *i*th allele in the locus.

Hardy-Weinberg (HE) equilibrium was tested for all markers and populations using the Guo and Thompson (1992) exact test with the Arlequin package (Schneider et al. 1997). In those cases where the exact test yielded a significant value, a χ^2-test was applied to assess the homozygosity excess.

Some parameters of forensic interest were calculated for each marker and population. The polymorphism information content (PIC) was calculated as described by Botstein et al. (1980). The power of discrimination (POD) was calculated following Fisher's method (Fisher 1951). The chance of paternity exclusion if the mother is known and typed (CE) was calculated as suggested by Smouse and Chakraborty (1986). The *a priori* probability of paternity exclusion if only one parent and child are typed (CE2), (equations 12 and 14 in Chakraborty and Jin 1993) was also calculated.

Allele association was tested with a likelihood ratio test (Slatkin and Excoffier 1996) as implemented in the Arlequin package, which was also used to test for population differentiation (Raymond and Rousset 1995).

Table 4 Allele frequencies, heterozygosity (Het), PIC, POD, CE and CE2 for the STR D18S51 in six populations from the Iberian Peninsula and North Africa

D18S51 (2n)	Bas 100	Cat 98	And 70	Por 78	Ara 94	Ber 98
10	0	0.010	0	0.013	0	0
11	0.050	0.061	0.014	0.026	0.011	0.010
12	0.240	0.133	0.071	0.141	0.096	0.163
13	0.080	0.133	0.143	0.103	0.170	0.102
13.2	0	0	0	0	0	0.010
14	0.120	0.143	0.100	0.192	0.085	0.163
15	0.090	0.173	0.143	0.128	0.106	0.112
16	0.110	0.133	0.214	0.128	0.255	0.143
17	0.200	0.122	0.114	0.115	0.117	0.112
18	0.030	0.031	0.029	0.077	0.043	0.102
19	0.020	0.041	0.086	0.038	0.043	0.041
20	0.040	0.010	0.057	0.026	0.032	0.020
21	0.020	0.010	0.029	0.013	0.032	0.020
Het	0.856	0.875	0.873	0.877	0.858	0.878
PIC	0.840	0.862	0.860	0.865	0.844	0.865
POD	0.964	0.971	0.971	0.973	0.965	0.972
CE	0.714	0.747	0.744	0.752	0.721	0.752
CE2	0.553	0.593	0.591	0.601	0.562	0.600

Table 5 Allele frequencies, heterozygosity (Het), PIC, POD, CE and CE2 for the STR D21S11 in six populations from the Iberian Peninsula and North Africa

D21S11 (2n)	Bas 100	Cat 98	And 68	Por 78	Ara 94	Ber 98
24.2	0	0.010	0	0	0	0
26	0	0.010	0	0	0.011	0
27	0.010	0.010	0.029	0.038	0.064	0.010
28	0.050	0.204	0.103	0.154	0.074	0.173
28.2	0	0	0	0	0	0.010
29	0.300	0.235	0.206	0.167	0.170	0.235
29.2	0	0	0	0	0	0.010
30	0.290	0.296	0.265	0.282	0.255	0.204
30.2	0.030	0.010	0.015	0.026	0.032	0.010
31	0	0.020	0.147	0.064	0.064	0.051
31.2	0.060	0.051	0.059	0.141	0.096	0.092
32	0	0.031	0.029	0.026	0	0.010
32.2	0.120	0.071	0.118	0.077	0.117	0.061
33.2	0.120	0.041	0.029	0.026	0.064	0.092
34	0	0	0	0	0.011	0
34.2	0.020	0.010	0	0	0.021	0.020
35	0	0	0	0	0.021	0.010
36	0	0	0	0	0	0.010
Het	0.790	0.804	0.835	0.836	0.863	0.849
PIC	0.761	0.779	0.815	0.817	0.850	0.832
POD	0.927	0.936	0.953	0.954	0.968	0.960
CE	0.598	0.623	0.676	0.679	0.731	0.701
CE2	0.421	0.449	0.506	0.510	0.575	0.538

Results and discussion

The 13 short tandem repeat (STR) loci D3S1358, vWA, FGA, D16S539, TH01, TPOX, CSF1PO, D8S1179, D21S11, D18S51, D5S818, D13S317 and D7S820 as well as the amelogenin locus, all contained in AmpF*l*STR Profiler Plus and AmpF*l*STR COfiler PCR amplification kits, were studied in four populations from the Iberian Peninsula (Basques, Catalans, Andalusians and Portuguese) and two North African populations from Morocco. Tables 1–13 show the allele frequencies and other parameters of forensic interest for all the loci and populations studied.

Out of 74 Hardy-Weinberg equilibrium tests, 4 were significant: CSF1PO in Arabs and Andalusians and TPOX and D5S818 in Arabs. When the Bonferroni correction for multiple tests was applied, none of the previous *p*-values could be considered significant. Moreover, a

Table 6 Allele frequencies, heterozygosity (Het), PIC, POD, CE and CE2 for the STR D3S1358 in six populations from the Iberian Peninsula and North Africa

D3S1358 (2n)	Bas 100	Cat 98	And 70	Por 78	Ara 94	Ber 100
12	0.010	0	0	0	0	0
13	0.020	0	0.014	0.013	0	0.020
14	0.160	0.092	0.086	0.154	0.021	0.060
15	0.340	0.286	0.171	0.321	0.372	0.270
16	0.150	0.265	0.329	0.179	0.245	0.280
17	0.150	0.194	0.257	0.154	0.160	0.220
18	0.170	0.143	0.143	0.128	0.181	0.140
19	0	0.020	0	0.051	0.011	0.010
20	0	0	0	0	0.011	0
Het	0.784	0.781	0.769	0.798	0.743	0.777
PIC	0.754	0.747	0.732	0.771	0.701	0.741
POD	0.923	0.918	0.910	0.932	0.892	0.914
CE	0.583	0.571	0.553	0.608	0.512	0.563
CE2	0.403	0.392	0.374	0.429	0.334	0.385

Table 7 Allele frequencies, heterozygosity (Het), PIC, POD, CE and CE2 for the STR D5S818 in six populations from the Iberian Peninsula and North Africa

D5S818 (2n)	Bas 100	Cat 98	And 68	Por 78	Ara 94	Ber 98
8	0	0	0	0.013	0.074	0.020
9	0.040	0.010	0.029	0.026	0.053	0.010
10	0.130	0.102	0.088	0.064	0.096	0.102
11	0.370	0.398	0.265	0.321	0.330	0.194
12	0.260	0.306	0.382	0.397	0.319	0.418
13	0.180	0.163	0.221	0.154	0.117	0.224
14	0.020	0.020	0.015	0.026	0.011	0.031
Het	0.744	0.710	0.726	0.710	0.758	0.725
PIC	0.703	0.661	0.680	0.662	0.722	0.684
POD	0.894	0.867	0.879	0.868	0.905	0.884
CE	0.517	0.465	0.486	0.470	0.545	0.497
CE2	0.339	0.294	0.312	0.299	0.367	0.319

Table 8 Allele frequencies, heterozygosity (Het), PIC, POD, CE and CE2 for the STR D7S820 in six populations from the Iberian Peninsula and North Africa

D7S820 (2n)	Bas 100	Cat 98	And 64	Por 78	Ara 94	Ber 96
7	0.050	0.061	0	0	0.021	0.021
8	0.170	0.194	0.109	0.179	0.138	0.135
9	0.060	0.092	0.188	0.179	0.149	0.125
10	0.410	0.235	0.344	0.346	0.415	0.229
11	0.170	0.245	0.156	0.231	0.117	0.229
12	0.110	0.153	0.188	0.038	0.117	0.229
13	0.030	0.020	0	0.013	0.032	0.021
14	0	0	0.016	0.013	0.011	0.010
Het	0.755	0.811	0.775	0.761	0.758	0.807
PIC	0.725	0.784	0.741	0.723	0.730	0.779
POD	0.910	0.937	0.916	0.905	0.913	0.935
CE	0.554	0.625	0.565	0.540	0.561	0.617
CE2	0.372	0.448	0.385	0.361	0.378	0.440

χ^2-test for homozygosity excess did not yield significant values. Thus, equilibrium may be assumed for all loci in all populations.

Allele association was tested for all possible pairs of loci in each population, giving a total of 420 tests. Of those, 23 were statistically significant with $p < 0.05$. Only three pairs of loci yielded significant allele association in more than one population. VWA and FGA (Andalusians, Arabs), FGA and D13S317 (Andalusians and Basques) D13S317 and TPOX (Catalans and Portuguese). CSF1PO and D5S818 markers map on the same chromosome (5q33.3–34 and 5q21–31, respectively), nevertheless they

Table 9 Allele frequencies, heterozygosity (Het), PIC, POD, CE and CE2 for the STR D8S1179 in six populations from the Iberian Peninsula and North Africa

D8S1179 (2n)	Bas 100	Cat 100	And 68	Por 78	Ara 94	Ber 98
8	0.020	0	0	0	0.010	0.041
9	0.010	0	0	0.013	0.010	0.010
10	0.050	0.070	0.088	0.064	0.074	0.102
11	0.110	0.110	0.074	0.115	0.096	0.143
12	0.070	0.100	0.147	0.128	0.074	0.102
13	0.230	0.370	0.279	0.244	0.223	0.194
14	0.290	0.220	0.279	0.256	0.266	0.194
15	0.210	0.100	0.088	0.154	0.202	0.173
16	0.010	0.020	0.029	0.026	0.032	0.020
17	0	0.010	0.015	0	0.010	0.020
Het	0.799	0.777	0.800	0.817	0.817	0.851
PIC	0.770	0.749	0.773	0.792	0.793	0.833
POD	0.931	0.922	0.933	0.942	0.942	0.960
CE	0.608	0.584	0.613	0.637	0.640	0.700
CE2	0.431	0.404	0.437	0.462	0.466	0.535

Table 10 Allele frequencies, heterozygosity (Het), PIC, POD, CE and CE2 for the STR FGA in six populations from the Iberian Peninsula and North Africa

FGA (2n)	Bas 100	Cat 98	And 70	Por 78	Ara 94	Ber 100
16	0	0	0	0	0	0.010
17	0	0	0	0	0	0.010
18	0.020	0.041	0.014	0.013	0	0.010
19	0.100	0.041	0.043	0.090	0.074	0.030
20	0.160	0.133	0.200	0.141	0.106	0.130
21	0.190	0.173	0.186	0.167	0.160	0.170
22	0.150	0.204	0.229	0.244	0.128	0.120
22.2	0	0	0.014	0	0	0
23	0.210	0.143	0.100	0.064	0.160	0.240
23.2	0	0	0.014	0.013	0	0
24	0.060	0.143	0.114	0.103	0.181	0.150
25	0.110	0.102	0.071	0.077	0.117	0.100
26	0	0.020	0.014	0.077	0.053	0.020
27	0	0	0	0.013	0.011	0.010
28	0	0	0	0	0.011	0
Het	0.846	0.856	0.842	0.858	0.866	0.848
PIC	0.826	0.839	0.823	0.843	0.852	0.830
POD	0.957	0.962	0.956	0.965	0.967	0.959
CE	0.688	0.709	0.686	0.718	0.729	0.696
CE2	0.521	0.545	0.519	0.557	0.570	0.531

Table 11 Allele frequencies, heterozygosity (Het), PIC, POD, CE and CE2 for the STR TH01 in six populations from the Iberian Peninsula and North Africa

TH01 (2n)	Bas 100	Cat 98	And 72	Por 78	Ara 94	Ber 100
5	0	0	0.028	0	0.011	0
6	0.200	0.224	0.222	0.205	0.170	0.240
7	0.090	0.122	0.111	0.205	0.255	0.190
8	0.110	0.102	0.139	0.141	0.117	0.190
9	0.210	0.173	0.181	0.167	0.319	0.270
9.3	0.370	0.367	0.292	0.269	0.096	0.090
10	0.020	0.010	0.028	0	0.032	0.010
11	0	0	0	0.013	0	0.010
Het	0.758	0.759	0.800	0.795	0.780	0.789
PIC	0.722	0.723	0.771	0.764	0.747	0.756
POD	0.906	0.906	0.931	0.926	0.919	0.922
CE	0.543	0.542	0.606	0.592	0.575	0.582
CE2	0.363	0.362	0.428	0.413	0.396	0.403

Table 12 Allele frequencies, heterozygosity (Het), PIC, POD, CE and CE2 for the STR TPOX in six populations from the Iberian Peninsula and North Africa

TPOX (2n)	Bas 98	Cat 98	And 72	Por 78	Ara 94	Ber 100
6	0	0	0	0.013	0	0.010
7	0	0	0.028	0	0.021	0
8	0.520	0.531	0.444	0.526	0.489	0.410
9	0.061	0.143	0.139	0.141	0.202	0.130
10	0.092	0.061	0.083	0.038	0.085	0.120
11	0.286	0.224	0.250	0.231	0.181	0.320
12	0.041	0.041	0.056	0.051	0.021	0
13	0	0	0	0	0	0.010
Het	0.634	0.642	0.710	0.646	0.679	0.698
PIC	0.580	0.596	0.669	0.601	0.636	0.646
POD	0.812	0.826	0.875	0.829	0.854	0.857
CE	0.386	0.405	0.483	0.410	0.445	0.448
CE2	0.224	0.234	0.305	0.239	0.269	0.279

Table 13 Allele frequencies, heterozygosity (Het), PIC, POD, CE and CE2 for the STR VWA in six populations from the Iberian Peninsula and North Africa

VWA (2n)	Bas 100	Cat 98	And 70	Por 78	Ara 94	Ber 98
13	0	0	0	0.013	0	0
14	0.080	0.061	0.114	0.051	0.106	0.133
15	0.120	0.163	0.057	0.141	0.117	0.102
16	0.250	0.204	0.329	0.244	0.330	0.306
17	0.320	0.235	0.300	0.256	0.170	0.245
18	0.160	0.204	0.157	0.205	0.223	0.112
19	0.050	0.092	0.043	0.077	0.032	0.082
20	0.020	0.041	0	0.013	0.011	0.020
21	0	0	0	0	0.011	0
Het	0.786	0.821	0.759	0.804	0.786	0.799
PIC	0.755	0.796	0.722	0.776	0.756	0.771
POD	0.923	0.943	0.905	0.933	0.924	0.932
CE	0.587	0.642	0.542	0.613	0.588	0.609
CE2	0.408	0.468	0.363	0.436	0.409	0.432

Table 14 Mean heterozygosity (Het), POD, CE and CE2 values for all the markers and the six populations tested

Population	Het	1-POD	1-CE	1-CE2
BAS	0.769	1.0×10^{-13}	3.6×10^{-5}	2.3×10^{-3}
CAT	0.778	3.5×10^{-14}	2.1×10^{-5}	1.6×10^{-3}
AND	0.779	4.1×10^{-14}	2.4×10^{-5}	1.7×10^{-3}
POR	0.782	2.2×10^{-14}	1.7×10^{-5}	1.3×10^{-3}
POR[a]	0.783	1.6×10^{-15}	7.0×10^{-6}	7.7×10^{-4}
ARA	0.780	2.5×10^{-14}	1.8×10^{-5}	1.4×10^{-3}
ARA[a]	0.780	1.9×10^{-15}	7.6×10^{-6}	8.2×10^{-4}
BER	0.789	1.3×10^{-14}	1.4×10^{-5}	1.1×10^{-3}

[a] Including marker D16S539

were not found to be in allelic association in any of the six populations tested ($p > 0.05$). It should be noted that under the hypothesis of no allele association, 5% of the tests (or 21 out of 420) are expected to appear as significant by chance. Therefore, we have disregarded allele association when estimating combined a priori statistics (Table 14).

The exact test of population differentiation (Raymond and Rousset 1995) among all six groups showed that all populations are significantly heterogeneous ($p < 0.0001$).

When each locus was analysed separately, six loci, namely D18S51, D13S317, D21S11, D7S820, TH01 and D3S1358 showed statistically significant differences between groups ($p < 0.001$, except for D3S1358, $p = 0.016$) when all the populations are included, whereas the remaining seven loci showed uniform allele frequencies in the six populations studied. When North African populations were excluded from the comparison for each loci, five of the six previous loci still showed statistically significant differences between populations (D18S51, D13S317, D21S11, D7S820, and D31358) indicating that the differences in allele frequencies for those loci were confined within the Iberian Peninsula. Pairwise comparisons showed that at three loci (D13S317, D21S11 and D7S820), all population pairs with statistically significant

differences in allele frequencies included the Basques, which seemed to be responsible for the overall differentiation in the Iberian Peninsula for those loci.

The *a priori* statistical power of this marker set is remarkable. Even in a difficult setting, such as paternity testing when only one parent and the child are typed, the combined *a priori* chance of exclusion is greater than 0.9977 in all of the populations tested.

Acknowledgements This research was supported by Dirección General de Investigación Científica Técnica (Spain) grant PB95–0267-C02–01, by the Direcció General de Recerca, Generalitat de Catalunya (1998SGR00009) and by Institut d'Estudis Catalans. We would like to thank B. Gutiérrez, L. Fañanás and E. Pintado for kindly providing us with the Andalusian samples. This work was done in collaboration with PE Biosystems.

References

Alonso S, Castro A, Fernández I, Gómez de Cedrón M, García-Orad A, Meyer E, Martínez de Pancorbo M (1995) Population study of 3 STR loci in the Basque Country (Northern Spain). Int J Legal Med 107:239–245

Anker R, Steinbruek T, Donis-Keller H (1992) Tetranucleotide repeat polymorphism at the human thyroid peroxidase (hTPO) locus. Hum Mol Genet 1:137

Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphism. Am J Hum Genet 32:314–331

Cabrero C, Díez A, Valverde E, Carracedo A, Alemany J (1995) Allele frequency distribution of four PCR-amplified loci in the Spanish population. Forensic Sci Int 71:153–164

Calafell F, Bertranpetit J (1994) Principal component analysis of gene frequencies and the origin of Basques. Am J Phys Anthropol 93:201–215

Chakraborty R, Jin L (1993) Determination of relatedness between individuals using DNA fingerprinting. Hum Biol 65:875–895

Comas D, Mateu E, Calafell F, Pérez-Lezaun A, Bosch E, Martínez-Arias R, Bertranpetit J (1998) HLA class I and class II DNA typing and the origin of Basques. Tissue Antigens 51:30–40

Edwards A, Hammond HA, Jin L, Caskey CT, Chakraborty R (1992) Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups. Genomics 12:241–253

Fisher R (1951) Standard calculations for evaluating a blood system. Heredity 5:95–102

Garofano L, Pizzamiglio M, Vecchio C, Lago G, Floris T, D'Errico G, Brembilla G, Romano A, Budowle B (1998) Italian population data on thirteen short tandem repeat loci: HUMTH01, D21S11, D18S51, HUMVWFA31, HUMFIBRA, D8S1179, HUMTPOX, HUMCSF1PO, D16S539, D7S820, D13S317, D5S818, D3S11358. Forensic Sci Int 97:53–60

Green ED, Mohr RM, Idol JR, Jones M, Buckingham JM, Deaven LL, Moyzis RK, Olson MV (1991) Systematic generation of sequence-tagged sites for physical mapping of human chromosomes: application to the mapping of human chromosome 7 using yeast artificial chromosomes. Genomics 11:548–564

Guo S, Thompson E (1992) Performing the exact test of Hardy-Weinberg proportion for multiple alleles. Biometrics 48:361–372

Hammond HA, Jin L, Zhong Y, Caskey CT, Chakraborty R (1994) Evaluation of 13 short tandem repeat loci for use in personal identification applications. Am J Hum Genet 55:175–189

Hudson TJ, Stein LD, Gerety SS, Ma J, Castle AB, Silva J, Slonim DK, Baptista R, Kruglyak L, Xu SH (1995) An STS-based map of the human genome. Science 270:1945–1954

Iriondo M, de la Rúa C, Barbero MC, Aguirre A, Manzano C (1999) Analysis of 6 short tandem repeat loci in Navarra (Northern Spain). Hum Biol 71:43–54

Kimpton C, Walton A, Gill P (1992) A further tetranucleotide repeat polymorphism in the vWF gene. Hum Mol Genet 1:287

Li H, Schmidt L, Wei MH, Hustad T, Lerman MI, Zbar B, Tory K (1993) Three tetranucleotide polymorphisms for loci: D3S1352; D3S1358; D3S1359. Hum Mol Genet 2:1327

Mills KA, Even D, Murray JC (1992) Tetranucleotide repeat polymorphism at the human alpha fibrinogen locus (FGA). Hum Mol Genet 1:779

Oldroyd NJ, Urquhart A, Kimpton CP, Millican ES, Watson SK, Downes T, Gill PD (1995) A highly discriminating octoplex short tandem repeat polymerase chain reaction system suitable for human individual identification. Electrophoresis 16:334–337

Pancorbo NM, Castro A, Fernández-Fernández I, García-Orad A (1996) Population genetics and forensic applications using multiplex PCR (CSF1PO, TPOX and TH01) loci in the Basque Country. J Forensic Sci 43:1181–1187

Pestoni C, Lareu MV, Rodríguez MS, Muñoz I, Barros F, Carracedo A (1995) The use of the STRs HUMTHO1, HUMVWA31/A, HUMF13A1, HUMFES/FPS, HUMLPL in forensic application: validation studies and population data for Galicia (NW Spain). Int J Legal Med 107:283–290

Raymond M, Rousset F (1995) An exact test for population differentiation. Evolution 49:1280–1283

Schneider S, Kueffer JM, Roessli D, Excoffier L (1997) Arlequin (ver.1.0) a software environment for the analysis of population genetics data. Genetics and Biometry Lab, University of Geneva, Switzerland

Sharma V, Litt M (1992) Tetranucleotide repeat polymorphism at the D21S11 locus. Hum Mol Genet 1:67

Slatkin M, Excoffier L (1996) Testing for allele association in genotypic data using the EM algorithm. Heredity 76:377–383

Smouse PE, Chakraborty R (1986) The use of restriction fragment length polymorphisms in paternity analysis. Am J Hum Genet 38:918–939

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 15:636–641

Torroni A, Bandelt HG, D'Urbano L, Lahermo P, Moral P, Sellito D, Rengo C, Forster P, Savontaus ML, Bonne-Tamir B, Scozzari R (1998) MtDNA analysis reveals a major late Paleolithic population expansion from southwestern to northeastern Europe. Am J Hum Genet 62:1137–1152

Urquhart A, Oldroyd NJ, Kimpton CP, Gill P (1995) Highly discriminating heptaplex short tandem repeat PCR system for forensic identification. Biotechniques 18:116–121