Int J Legal Med (2000) 113 : 263–267

ORIGINAL ARTICLE

C. Tomàs · A. Picornell · J. A. Castro · M. M. Ramon
L. Gusmão · M. V. Lareu · A. Carracedo

Genetic variability at nine STR loci in the Chueta (Majorcan Jews) and the Balearic populations investigated by a single multiplex reaction

Received: 12 April 1999 / Received in revised form: 11 August 1999

Abstract A study of the genetic variability in the Chueta (Majorcan Jews) and the Balearic (Majorca and Minorca Islands) populations was carried out using a multiplex system containing the nine tetrameric STRs D3S1358, vWA, FGA, D8S1179, D21S11, D18S51, D5S818, D13S317 and D7S820. The Chueta population has remained isolated because intermarriage with non-Jews did not take place until the middle of this century, which has resulted in it being a small inbred community. The results indicate the existence of HW equilibrium for the Chueta and Balearic populations. No pair-wise correlation was observed between the nine markers. Consequently, they seem to comprise a suitable group of markers for population genetics purposes and for paternity and forensic testing.

Key words STRs · Chuetas · Majorcan Jews · Balearic Islands · Multiplex reaction · Profiler Plus

Introduction

Various types of genetic markers have been used in order to understand the historical process of human differentiation better. Among these, short tandem repeats (STRs) are a well-known group of highly polymorphic markers which are abundant and widespread in the human genome. Since their first description in the early 1990s (Edwards et al. 1991; Polymeropoulos et al. 1991) they have been widely used as population, forensic and clinical markers, especially those based on a 4 bp motif. The fact that tetranucleotide STRs give widely spaced bands for different alleles in a gel and they have a low tendency for slippage, facilitates an unambiguous designation of the alleles, which allows their easy multiplexing and automation (Kimpton et al. 1993; Pérez-Lezaun et al. 1997). Moreover, previous studies have shown that multiple STR loci can be amplified in a single multiplex reaction (Morral and Estivill 1992; Lygo et al. 1994; Entrala et al. 1998). This is especially interesting in order to reduce both the quantity of sample needed and the time of determination.

Data on genotype frequencies in different populations is a basic prerequisite for forensic application but at present few data on local populations are available for STR allele frequency variation.

Here we present a population analysis of the Chueta community (Majorcan Jews, Balearic Islands, Spain) and the Balearic population (from Majorca and Minorca Islands). The characteristics of the Chueta and Balearic populations can be seen in depth elsewhere (Massanet et al. 1997; Picornell et al. 1997). Briefly, the present population of the Balearic Islands can be considered as a fusion of Mediterranean cultures, one of which, the Jewish culture has had an historical continuity on the island of Majorca and they are the so-called Chuetas. In spite of their official conversion to Christianity they kept their traditions and beliefs, although the Inquisition persecuted them until the seventeenth century. They have remained isolated because intermarriage with non-Jews did not take place until the middle of this century and consequently this community has remained a small and inbred population. Several genetic studies with classical markers and mtDNA RFLPs (Picornell et al. 1997; Castro et al. 1998) carried out on this population, showed significant differences between Chuetas and non-Chuetas. Comparisons with other Jewish populations confirmed the Jewish origin and also a certain degree of admixture between the Chuetas and their Gentile neighbours (population of Majorca).

C. Tomàs · A. Picornell · J. A. Castro · M. M. Ramon ✉)
Laboratori de Genètica. Departament de Biologia,
Facultat de Ciències, Universitat de les Illes Balears,
E-07071 Palma de Mallorca, Balearic Islands, Spain
e-mail: dbsmrj0@ps.uib.es

L. Gusmão
Instituto de Patologia e Inmunologia Molecular
da Universidade do Porto, (IPATIMUP) E-4200 Porto, Portugal

M. V. Lareu · A. Carracedo
Instituto de Medicina Legal. Facultade de Medicina,
E-15705 Santiago de Compostela, Galicia, Spain

The purpose of this work was to study the genetic variability of nine tetrameric STRs in the Chueta and the Balearic populations by means of a multiplex system.

Materials and methods

Blood samples from 215 unrelated individuals from the Balearic Islands (102 from the Chueta community and 113 from autochthonous individuals from Majorca and Minorca Islands) were collected and frozen at -20 °C. DNA was organically extracted by a standard protocol (Sambrook et al. 1989) and quantified spectrophotometrically. The co-amplification of the nine STR loci was performed using 5 ng of genomic DNA in a total reaction volume of 10 µl. The AmpF*l*STR Profiler Plus PCR amplification kit (PE Applied Biosystems) was used for the amplification and analysis of the STRs following the manufacturer's instructions. The STRs studied are shown in Table 1

Amplified DNA was analysed on an ABI Prism 377 DNA Sequencer (PE Applied Biosystems). Samples (0.5 µl) were mixed with formamide, loading buffer and the internal standard size (GS-500 Rox), denatured at 97 °C for 5 min and run in a 4.6% denaturing gel. The fragment size of the different alleles was determined with the allelic ladders provided with the kit and GeneScan 2.1 Analysis software was used for interpretation.

The sequence analysis of the novel alleles was carried out by isolating the amplified alleles from heterozygous individuals using a 6% polyacrylamide gel and eluting using the "crush and soak" method (Sambrook et al. 1989). Isolated fragments were reamplified and purified by using the QIAquick PCR purification kit (Qiagen). The sequence reaction was carried out using the BigDye Terminator Cycle Sequencing Ready Reaction kit (PE Applied Biosystems). Sequences were analysed on an ABI Prism 310 DNA Sequencer (PE Applied Biosystems). Sequencing Analysis version 3.0 was used for the automatic analysis of the sequenced data.

Table 1 STR loci used in this study

Locus	Chromosome location	Reference
D3S1358	3p	Li et al. 1993
vWA	12p12-pter	Kimpton et al. 1992
FGA	4q28	Mills et al. 1992
D8S1179	8	Oldroyd et al. 1995
D21S11	21	Sharma and Litt 1992
D18S51	18q21.3	Urquhart et al. 1994
D5S818	5q21–31	Hudson et al. 1995
D13S317	13q22–31	Hudson et al. 1995
D7S820	7q	Green et al. 1991

Statistical analyses

Allele frequencies were estimated by the gene counting method. The Hardy-Weinberg (HW) equilibrium was tested by calculating

Table 2 Allele frequencies of the nine STRs studied in the Chueta population (Majorcan Jews)

Allele	STR								
	D3S1358	vWA	FGA	D8S1179	D21S11	D18S51	D5S818	D13S317	D7S820
7	–	–	–	0.005	–	–	–	–	–
8	–	–	–	0.025	–	–	0.020	0.044	0.113
9	–	–	–	0.005	–	–	0.103	0.054	0.172
10	–	–	–	0.029	–	–	0.108	0.054	0.255[a]
11	–	–	–	0.103	–	0.015	0.309[a]	0.255	0.230
12	–	–	–	0.132	–	0.093	0.299	0.436[a]	0.230
13	0.034	–	–	0.221[a]	–	0.206[a]	0.147	0.074	–
14	0.044	0.123	–	0.186	–	0.132	0.005	0.083	–
15	0.309[a]	0.157	–	0.216[a]	–	0.137	0.010	–	–
16	0.265	0.230	–	0.078	–	0.152[a]	–	–	–
17	0.157	0.235[a]	–	–	–	0.069	–	–	–
18	0.181[a]	0.157	0.034	–	–	0.098	–	–	–
19	0.005	0.074	0.029	–	–	0.074	–	–	–
20	0.005	0.025	0.108	–	–	0.025	–	–	–
21	–	–	0.181[a]	–	–	–	–	–	–
21.2	–	–	0.015	–	–	–	–	–	–
22	–	–	0.118[a]	–	–	–	–	–	–
22.2	–	–	0.005	–	–	–	–	–	–
23	–	–	0.162[a]	–	–	–	–	–	–
23.2	–	–	0.005	–	–	–	–	–	–
24	–	–	0.186[a]	–	–	–	–	–	–
25	–	–	0.152	–	–	–	–	–	–
27	–	–	0.005	–	0.015	–	–	–	–
27.2	–	–	–	–	0.010	–	–	–	–
28	–	–	–	–	0.270[a]	–	–	–	–
29	–	–	–	–	0.108	–	–	–	–
30	–	–	–	–	0.250[a]	–	–	–	–
30.2	–	–	–	–	0.010	–	–	–	–
31	–	–	–	–	0.069	–	–	–	–
31.2	–	–	–	–	0.054	–	–	–	–
32	–	–	–	–	0.015	–	–	–	–
32.2	–	–	–	–	0.132[a]	–	–	–	–
33.2	–	–	–	–	0.069	–	–	–	–

[a] Modal alleles

the exact *p*-values proposed by Guo and Thompson (1992) using the Markov chain method by means of the updated version of GENEPOP (Raymond and Rousset 1995). Furthermore, a standard χ^2 goodness-of-fit test was calculated to assess HW expectations by means of the BIOSYS-1 package (Swofford and Selander 1989).

Gene and genotype differentiation between the Chueta and the Balearic populations was estimated using the GENEPOP package.

Independence among loci was estimated by means of genotype disequilibrium and the *p*-value for each locus pair using Fisher's method (Raymond and Rousset 1995). When appropriate, the Bonferroni procedure (Weir 1996) was used to correct multiple analyses.

We also calculated the power of discrimination (PD) following Fisher's method (Fisher 1951), the *a priori* chance of exclusion (CE) (Ohno et al. 1982) and the heterozygosity value (h) (Nei and Roychoudhury 1974). These three parameters allowed us to determine the utility of the markers studied in our populations from a forensic point of view.

Results and discussion

Allele frequencies of the nine STRs studied in the Chueta and Balearic populations are shown in Tables 2 and 3. Alleles were designated according to the number of repeats of the tetranucleotide motif. Some intermediate alleles were detected at two of the loci studied: alleles 21.2, 22.2 and 23.2 at the FGA locus and alleles 24.2, 27.2, 31.2, 32.2 and 33.2 at the D21S11 locus. Moreover, two new alleles, not yet described in the literature, were found in the Chueta population: one allele at the D8S1179 locus with an electrophoretic mobility corresponding to an allele 7, and an allele at the D21S11 locus that would correspond to allele 27.2 of the allelic ladder. Sequencing data has been made in order to characterise these new alleles and confirm that the allele 7 of the D8S1179 locus has seven units of the repetitive motif $(TCTA)_7$. The new allele of the D21S11 locus was designated as allele 27.2. Sequencing analysis of this allele revealed a 14 bp deletion of $(TCTA)_3$ TA at the beginning of the constant region (Brinkmann et al. 1998), as described for allele 25.2 in a Belgian Caucasian individual (Xiao et al. 1998). The complete structure of the allele is as follows: $(TCTA)_5$ $(TCTG)_6$ $(TCTA)_3$ TCA $(TCTA)_2$ TCCATA $(TCTA)_{12}$.

The modal distribution was analysed following the definition given by Shriver et al. (1993), where a mode was defined as any allele that presents a frequency higher than the next adjacent alleles. In the Balearic population a unimodal pattern was detected at three of the nine markers

Table 3 Allele frequencies of the nine STRs studied in the Balearic population

Allele	STR								
	D3S1358	vWA	FGA	D8S1179	D21S11	D18S51	D5S818	D13S317	D7S820
7	–	–	–	–	–	–	–	–	0.027
8	–	–	–	0.022	–	–	–	0.116[a]	0.186[a]
9	–	–	–	–	–	–	0.022	0.054	0.119
10	–	–	–	0.103	–	0.004	0.128	0.063	0.261[a]
11	0.004	–	–	0.112	–	0.018	0.323	0.268	0.204
12	–	–	–	0.138	–	0.124	0.398[a]	0.353[a]	0.164
13	–	0.004	–	0.308[a]	–	0.137	0.111	0.116	0.031
14	0.124	0.129[a]	–	0.170	–	0.164[a]	0.018	0.031	0.009
15	0.177	0.112	–	0.112	–	0.124	–	–	–
16	0.327[a]	0.237	–	0.027	–	0.150[a]	–	–	–
17	0.168	0.241[a]	–	0.009	–	0.115	–	–	–
18	0.190[a]	0.183	0.013	–	–	0.075	–	–	–
19	0.009	0.085	0.066	–	–	0.058	–	–	–
20	–	0.009	0.150	–	–	0.013	–	–	–
21	–	–	0.177	–	–	0.013	–	–	–
22	–	–	0.208[a]	–	–	0.004	–	–	–
22.2	–	–	0.027	–	–	–	–	–	–
23	–	–	0.088	–	–	–	–	–	–
24	–	–	0.119[a]	–	–	–	–	–	–
24.2	–	–	–	–	0.004	–	–	–	–
25	–	–	0.097	–	–	–	–	–	–
26	–	–	0.044	–	0.009	–	–	–	–
27	–	–	0.004	–	0.022	–	–	–	–
28	–	–	0.004	–	0.146	–	–	–	–
29	–	–	–	–	0.239	–	–	–	–
30	–	–	–	–	0.243[a]	–	–	–	–
30.2	–	–	–	–	0.053	–	–	–	–
31	–	–	–	–	0.058	–	–	–	–
31.2	–	–	–	–	0.080	–	–	–	–
32	–	–	–	–	0.018	–	–	–	–
32.2	–	–	–	–	0.088	–	–	–	–
33.2	–	–	–	–	0.031	–	–	–	–
34.2	–	–	–	–	0.009	–	–	–	–

[a] Modal alleles

Table 4 Statistical data for the nine genetic systems in the Chueta (Majorcan Jews) and Balearic populations

Population	System	Power of discrimination	Chance of exclusion	Allelic diversity h obs./h exp.	Number of alleles	χ^2 test[a]	*P* exact test[a]	Deficiency of heterozygotes[a]
Chueta	D3S1358	0.913	0.561	0.745/0.778	8	0.843	0.784	0.296
	vWA	0.943	0.643	0.833/0.825	7	0.235	0.156	0.342
	FGA	0.962	0.708	0.922/0.860	12	0.890	0.450	0.977
	D8S1179	0.951	0.669	0.853/0.838	10	0.410	0.061	0.428
	D21S11	0.947	0.655	0.716/0.827	11	0.015	0.005	0.022
	D18S51	0.969	0.736	0.882/0.873	10	0.068	0.056	0.599
	D5S818	0.913	0.560	0.755/0.775	8	0.794	0.713	0.396
	D13S317	0.889	0.512	0.686/0.728	7	0.583	0.393	0.038
	D7S820	0.922	0.575	0.843/0.791	5	0.281	0.320	0.957
Balearic	D3S1358	0.919	0.575	0.788/0.785	7	0.380	0.223	0.392
	vWA	0.940	0.633	0.768/0.820	8	0.481	0.203	0.104
	FGA	0.967	0.728	0.867/0.868	12	0.000	0.008	0.099
	D8S1179	0.946	0.651	0.821/0.824	9	0.504	0.340	0.400
	D21S11	0.956	0.687	0.885/0.844	13	0.206	0.204	0.855
	D18S51	0.972	0.753	0.867/0.882	13	0.990	0.962	0.426
	D5S818	0.865	0.465	0.646/0.711	6	0.252	0.055	0.092
	D13S317	0.914	0.565	0.759/0.773	7	0.667	0.579	0.365
	D7S820	0.938	0.627	0.823/0.817	8	0.487	0.171	0.722

[a] These are probability values

analysed (D8S1179, D21S11 and D5S818) whereas a bimodal distribution was observed in the rest of the systems. In the Chueta population a unimodal pattern was detected in four of the loci studied (vWA, D5S818, D13S317 and D7S820) whereas a multimodal pattern was observed in the other systems.

Other statistical parameters of genetic variability, such as heterozygosity and the number of alleles, were calculated (Table 4). The heterozygosity ranged from 0.686 to 0.922 in the Chueta population, with a mean value of 0.804, and from 0.646 to 0.885 in the Balearic population with a mean value of 0.803. The number of alleles observed ranged from 5 to 12 for the Chueta and from 6 to 13 for the Balearic populations. Some forensic parameters of interest (PD and CE) are also shown in Table 4. The usefulness of the nine systems can be judged by their combined PD, which corresponded to 1 in 80,000 million individuals (in the Chueta community) and 1 in 135,000 million individuals (in the Balearic population).

When the Hardy-Weinberg equilibrium was tested (Table 4) only one marker was significant in each population ($P < 0.05$), D21S1 in the Chueta population and FGA in the Balearic population. Nevertheless, if the Bonferroni procedure is applied no system is statistically significantly different from HW equilibrium expectations. These results are consistent with other studies in the literature, which in most cases show only slight deviations, or none at all, from HWE. The endogamy of the Chueta community has been demonstrated by means of classical marker polymorphisms (Picornell et al. 1997). Nevertheless, only two loci out of the nine analysed showed a significant deficiency of heterozygotes (D21S11 and D13S317 loci).

When gene and genotypic differentiation was estimated between the two populations studied, a highly significant p value was obtained for the overall analysis. In particular, all but two markers (vWA and D18S51) used in the present study showed significant differences between the Chueta and the Balearic populations.

Analyses were also performed to determine whether there were any detectable associations between STR loci. An inter-class correlation test analysis demonstrated that there was little evidence for correlation between the alleles at any of the pairs of loci, only 6 out of 36 pair-wise comparisons showed a deviation from the expected result ($p < 0.05$). When a Bonferroni procedure was used as a correction for multiple tests performed on a population sample, there was no evidence of deviation for the pair-wise comparisons, with the exception of the FGA – D7S820 and D8S1179 – D7S820 pairs.

Although the STRs studied in this work can be typed individually or by means of different multiplex systems, the results obtained revealed that this method is highly effective. Moreover, the nine loci studied are very informative, so they constitute a group of good markers that have general applications in population genetics, paternity testing and for forensic purposes.

Acknowledgements This work was supported by grant PM97–0041 from the Dirección General de Enseñanza Superior (Spain). C. Tomàs is a recipient of a F.P.I. fellowship from the Ministerio de Educación y Cultura (Spain).

References

Brinkmann B, Klintschar M, Neuhuber F, Hühne J, Rolf B (1998) Mutation rate in human microsatellites: influence of the structure and length of the tandem repeat. Am J Hum Genet 62: 1408–1415

Castro JA, Picornell A, Ramon MM (1998) Mitochondrial DNA: a tool for populational genetics studies. Int Microbiol 1:327–332
Edwards A, Citivello A, Hammond HA, Caskey CT (1991) DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. Am J Hum Genet 49:746–756
Entrala C, Lorente M, Lorente JA, Alvarez JC, Moretti T, Budowle B, Villanueva E (1998) Fluorescent multiplex analysis of nine STR loci: Spanish population data. Forensic Sci Int 98(3):179–183
Fisher RA (1951) Standard calculations for evaluating a blood group system. Heredity 5:95–102
Green DE, 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 SW, Thompson EA (1992) Performing the exact test of Hardy-Weinberg proportions for multiple alleles. Biometrics 48:361–372
Hudson TJ, Stein LD, Gerety SS, et al. (1995) An STS-based map of the human genome. Science 270(5244):1945–1954
Kimpton CP, Walton A, Gill P (1992) A further tetranucleotide repeat polymorphism in the vWF gene. Hum Mol Genet 1:28
Kimpton CP, Gill P, Walton A, Urquhart A, Millican ES, Adams M (1993) Automated DNA profiling employing multiplex amplification of short tandem repeat loci. PCR Methods Appl 3:13–22
Li H, Schmidt L, Wei M, Hustad T, Lerman, MI, Zbar B, Tory K (1993) Three tetranucleotide polymorphisms for loci: D3S1352; D3S1358; D3S1359. Hum Mol Genet 2:1327
Lygo JE, Johnson PE, Holdaway DJ, Woodroffe S, Whitaker JP, Clayton TM, Kimpton CP, Gill P (1994) The validation of short tandem repeat (STR) loci for use in forensic casework. Int J Legal Med 107:77–89
Massanet MF, Castro JA, Picornell A, Ramon MM (1997) Study of the populations of the Balearic Islands (Spain) using mtDNA RFLPs. Hum Biol 69:483–498
Mills KA, Even D, Murray JC (1992) Tetranucleotide repeat polymorphism at the human alpha fibrinogen locus (FGA). Hum Mol Genet 1:1327
Morral N, Estivill X (1992) Multiplex PCR amplifications of three microsatellites within the CFTR gene. Genomics 13:1362–1364
Nei M, Roychoudhury AK (1974) Sampling variances of heterozygosity and genetic distance. Genetics 76:379–390
Ohno Y, Sebetan IM, Akaishi S (1982) A simple method for calculating the probability of excluding paternity with any number of codominant alleles. Forensic Sci Int 19:93–98
Oldroyd NJ, Urquhart A, Kimpton CP, Downs TJ, Millican ES, Watson SK, Gill PD (1995) A highly discriminating octoplex short tandem repeat polymerase chain reaction system suitable for human individual identification. Electrophoresis 16:334–337
Pérez-Lezaun A, Calafell F, Mateu E, Comas D, Ruiz-Pacheco R, Bertranpetit J (1997) Microsatellite variation and the differentiation of modern humans. Hum Genet 99:1–7
Picornell A, Castro JA, Ramon MM (1997) Genetics of the chuetas (Majorcan Jews): a comparative study. Hum Biol 69:313–328
Polymeropoulos MH, Rath DS, Xiao H, Merril CR (1991) Tetranucleotide repeat polymorphism at the human tyrosine hydroxylase gene (TH). Nucleic Acid S Res 19(13):3753
Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. J Heredity 86:248–249
Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor
Sharma V, Litt M (1992) Tetranucleotide repeat polymorphism at the D21S11 locus. Hum Mol Genet 1:67
Shriver MD, Jin L, Chakraborty R, Boerwinkle E (1993) VNTR allele frequency distributions under the stepwise mutation model: a computer simulation approach. Genetics 134:983–993
Swofford DL, Selander RB (1989) BIOSYS-1. A computer program for the analysis of allelic variation in population genetics and biochemical systematics. Release 1.7. University of Illinois, Urbana, Il
Urquhart A, Kimpton CP, Downs TJ, Gill P (1994) Variation in short tandem repeat sequences – a survey of 12 microsatellite loci for use as forensic identification markers. Int J Legal Med 107:13–20
Weir BS (1996) Genetic data analysis II. Sinauer, Sunderland, Mass
Xiao FX, Gilissen A, Cassiman JJ, Decorte R (1998) Quadruplex fluorescent STR typing system (HUMVWA, HUMTH01, D21S11 and HPRT) with sequence-defined allelic ladders identification of a new allele at D21S11. Forensic Sci Int 94:39–46