Int J Legal Med (2004) 119 : 10–15
DOI 10.1007/s00414-004-0439-2

ORIGINAL ARTICLE

M. Brion · B. Sobrino · A. Blanco-Verea · M. V. Lareu
A. Carracedo

Hierarchical analysis of 30 Y-chromosome SNPs in European populations

Received: 11 December 2003 / Accepted: 2 March 2004 / Published online: 17 April 2004

Abstract Analysis of Y-chromosome haplogroups defined by binary polymorphisms, has became a standard approach for studying the origin of modern human populations and for measuring the variability between them. Furthermore, the simplicity and population specificity of binary polymorphisms allows inferences to be drawn about the population origin of any male sample of interest for forensic purposes. From the 245 binary polymorphisms that can be analysed by PCR described in the Y Chromosome Consortium tree, we have selected 30 markers. The set of 30 has been grouped into 4 multiplexes in order to determine the most frequent haplogroups in Europe, using only 1 or 2 multiplexes. In this way, we avoid typing unnecessary SNPs to define the final haplogroup saving effort and cost, since we only need to type 9 SNPs in the best case and in the worst case, no more than 17 SNPs to define the haplogroup. The selected method for allele discrimination was a single base extension reaction using the SNaPshot multiplex kit. A total of 292 samples from 8 different districts of Galicia (northwest Spain) were analysed with this strategy. No significant differences were detected among the different districts, except for the population from Mariña Lucense, which showed a distant haplogroup frequency but not higher Φ_{st} values.

Keywords Y chromosome · Single nucleotide polymorphisms · SNPs · Single base extension · SNaPshot reaction · Europe

Introduction

Single nucleotide polymorphisms (SNPs, including small insertions and deletions) are the markers of choice for many applications, including the location and identification of disease susceptibility genes, pharmacogenomics and pharmacogenetics, studying the origin of modern human populations or measuring the variability between them (Jorde et al. 2001; Zhao et al. 2003).

Markers located on the Y chromosome have specific interest as forensic tools, because most of the chromosome

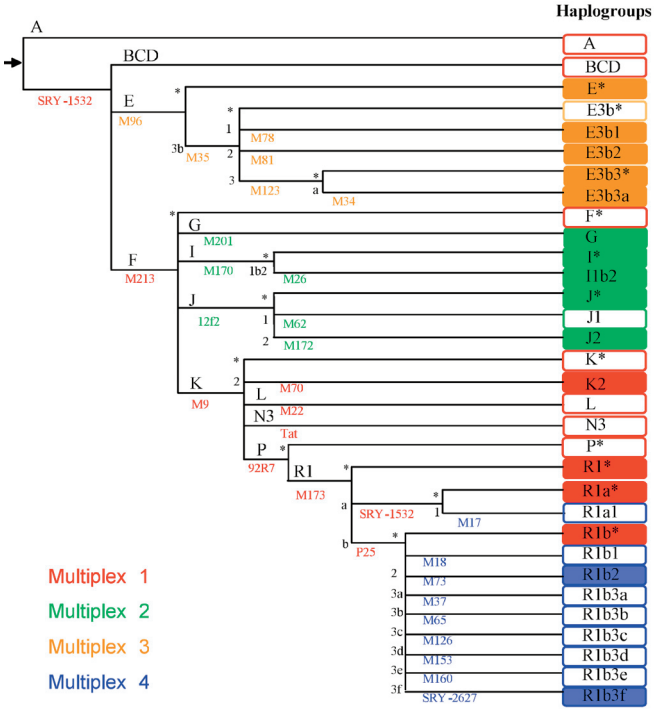

Fig. 1 Phylogenetic tree defined with the binary Y-chromosomal polymorphisms analysed. Marker names are indicated below the lines and lineage names are shown above the lines, but the length of each branch has no significance. Colours represent multiplex groups, and the coloured area represents the presence in Galicia

M. Brion (✉) · B. Sobrino · A. Blanco-Verea · M. V. Lareu
A. Carracedo
Institute of Legal Medicine,
University of Santiago de Compostela, San Francisco s/n,
15782 Santiago de Compostela, Spain
Tel.: +34-981582327, Fax: +34-981580336,
e-mail: brioniml@usc.es

does not undergo recombination. In particular, Y-chromosome SNPs, because of their abundance, simplicity and low mutation rate, are becoming an extensively used marker set. Forensic laboratories are starting to implement Y chromosome SNP analysis, in order to investigate the forensic usefulness of these markers (Sanchez et al. 2003; Borsting et al. 2004).

In the case of autosomal SNPs, the required number of SNPs to give comparable levels of information to STRs (Gill 2001), the most widespread markers used in the forensic field has already been determined. However, in the case of the Y chromosome, the selection of SNPs is complicated by the far more extensive genetic differentiation exhibited. This differentiation across geographical distance (Seielstad et al. 1998) results in markedly different haplogroup profiles, depending on the region of the world studied (Karafet et al. 1999; Rosser et al. 2000).

At the time of writing this manuscript, the Ensembl database lists 36,449 Y-SNPs, however most of them could be paralogous sequence variants (Jobling and Tyler-Smith 2003; Sanchez et al. 2004), highlighted by comparing true Y-chromosomal sequences with similar sequences elsewhere. Currently more than 240 Y-chromosome SNPs are available and well characterised. They define a highly resolved tree of binary haplogroups with a unified nomenclature proposed by the Y Chromosome Consortium (YCC 2002; Jobling and Tyler-Smith 2003).

Checking the literature, an extensive search has been performed looking for the allele frequencies of each SNP in European populations (Rosser et al. 2000; Semino et al. 2000; Brion et al. 2003). As a result of this search, a set of 30 SNPs was selected in order to determine 32 of the most frequent haplogroups present in European populations (Fig. 1). The inclusion of a male sample in one of these haplogroups might be useful in the identification of the regional origin within Europe. In combination with Y-chromosome STR variation, the regional identification of any male sample involved in a forensic case could be possible in the future.

A large number of SNP genotyping methods are now available (Chen and Sullivan 2003), and usually the choice of the appropriate method depends on the number of SNPs and the number of individuals that need to be typed. As this study typed 30 SNPs in 292 samples, the genotyping method selected was a multiplex PCR followed by the single base extension reaction using the SNaPshot multiplex kit (Applied Biosystems, Foster City, CA).

Table 1 SNP primer sequences and PCR concentrations of the SNPs used in this study, grouped by multiplexes

SNP	Primer (5'→3')		Size (bp)	Conc. (μM)	
	Forward	Reverse			
92R7	TGCATGAACACAAAAGACGTA	GCATTGTTAAATATGACCAGC	55	0.20	Multiplex1
M70	TCATAGCCCACTATACTTTGGAC	CTGAGGGCTGGACTATAGGG	81	0.20	
M22	GCTGATAGTCCTGGTTTCCCTA	TGAGCATGCCTACAGCAGAC	106	0.20	
Tat	GACTCTGAGTGTAGACTTGTGA	GAAGGTGCCGTAAAAGTGTGAA	112	0.20	
P25	GGACCATCACCTGGGTAAAGT	AGTGCTTGTCCAAGGCAGTA	121	0.20	
SRY1532	TCCTTAGCAACCATTAATCTGG	AAATAGCAAAAACTGACACAAGGC	167	0.20	
M173	GCACAGTACTCACTTTAGGTTTGC	GCAGTTTTCCCAGATCCTGA	172	0.20	
M213	GGCCATATAAAAACGCAGCA	TGAATGGCAAATTGATTCCA	208	0.30	
M9	GCAGCATATAAAACTTTCAGG	AAAACCTAACTTTGCTCAAGC	340	0.35	
12f2	CACTGACTGATCAAAATGCTTACAGAT	GGATCCCTTCCTTACACCTTATACA	90	0.25	Multiplex2
M201	TCAAATTGTGACACTGCAATAGTT	CATCCAACACTAAGTACCTATTACGAA	144	0.25	
M26	AGCAGAAGAGACCAAGACAGC	GACGAAATCTGCAGCAAAAA	147	0.25	
M170	TGCAGCTCTTATTAAGTTATGTTTTCA	CCAATTACTTTCAACATTTAAGACC	158	0.30	
M172	TCCTCATTCACCTGCCTCTC	TCCATGTTGGTTTGGAACAG	187	0.25	
M62	ACTAAAACACCATTAGAAACAAAGG	CTGAGCAACATAGTGACCCC	309	0.25	
M96	GTGATGTGTAACTTGGAAAACAGG	GGACCATATATTTTGCCATAGGTT	88	0.25	Multiplex3
M34	CACAGTGTTTTCTCATGTTAATGC	GGGGACCCCAATAATCATAA	92	0.25	
M81	TTATAGTTTCAATCCCTCAGTAATTTT	TGTTTCTTCTTGGTTTGTGTGAGTA	176	0.25	
M35	GCATGGTCCCTTTCTATGGAT	GAGAATGAATAGGCATGGGTTC	198	0.25	
M123	CACAGAGCAAGTGACTCTCAAAG	TCTTTCCCTCAACATAGTTATCTCA	248	0.25	
M78	CTTCAGGCATTATTTTTTTTGGT	ATAGTGTTCCTTCACCTTTCCTT	301	0.25	
M65	AAGGCTACCCATTCCCAAAT	AAGTCTGGCATCTGCAAAATC	71	0.15	Multiplex4
M126	GTGCTTGAAACCGAGTTTGT	TCGGGAAACACAATTAAGCA	83	0.15	
M73-M160	AAAACAATAGTTCCAAAAACTTCTGA	CCTTTGTGATTCCTCTGAACG	98	0.5	
M37	ATGGAGCAAGGAACACAGAA	AAGAAAGGAGATTGTTTTCAATTTT	124	0.3	
M167	GAGGCTGGGCCAAGTTAAGG	CTTCCTCGGAACCACTACCA	130	0.15	
M17-M18	CTGGTCATAACACTGGAAATC	AGCTGACCACAAACTGATGTAGA	171	0.10	
M153	TCTGACTTGGAAAGGGGAAA	TTTTCTCCTCATTATTTGTCTTCA	239	0.5	

Instead of typing the whole set of polymorphisms in each sample, our strategy was to group the SNPs in a hierarchical way following the YCC tree. Four multiplex PCR/primer extension reactions were developed, allowing the assignment of a sample to 1 of 32 possible haplogroups using only 1 or 2 multiplexes.

Material and methods

DNA samples

A total of 292 male subjects, belonging to 8 different districts of Galicia (northwest Spain) were analysed. Appropriate informed consent was obtained from all individuals. Blood was collected by venous puncture using EDTA as anticoagulant. Genomic DNA was extracted using a phenol-chloroform method.

Multiplex PCRs

PCR multiplexes were performed in 25 µl final volume, with 1× buffer, 300 µM of dNTPs, 2 mM of $MgCl_2$, 2 U of AmpliTaq Gold polymerase (Applied Biosystems) and 10 ng of genomic DNA. The cycling conditions were 95°C for 10 min then 32 cycles of 94°C for 30 s, 59°C for 30 s, 70°C for 30 s, and a final extension at 65°C for 15 min. Despite the fact that primer designs for PCR amplification of these SNPs are available in the literature, most of them were redesigned using the Primer3 software (http://www-genome.wi.mit.edu/cgi-bin/primer/primer3_www.cgi) and checked for possible secondary structures with the Oligonucleotide properties calculator software v. 3.02 (http://www.basic.nwu.edu/biotools/oligocalc.html). Primer sequences and detailed individual concentrations are shown in Table 1.

Amplicon sizes have been designed to be different enough to allow checking the amplification by electrophoresis. Conventional polyacrylamide electrophoresis (T=9, C=5) with silver stain detection was used for checking the amplification products.

Multiplex single base extensions

Before single base extension (SBE), 1 µl of the PCR product was cleaned up with 0.5 µl of ExoSAP-IT (Amersham Biosciences) and incubated at 37°C for 15 min followed by 85°C for 15 min to inactivate the enzyme.

Multiplex single base extension reactions were performed in a 5 µl final volume, combining 2 µl of SNaPshot ready reaction mix (Applied Biosystems), 1.5 µl of cleaned PCR product and extension primers. The cycling conditions were 96°C for 10 s, 50°C for 5 s and 60°C for 30 s, for 25 cycles. The same primer design software used to develop PCR primers helped to select the SBE primers. However, in this case each primer had varying lengths of poly (dC) non-homologous tails attached at the 5' end. All SBE primer sequences and concentrations are shown in Table 2.

To remove the unincorporated ddNTPs, the final product was incubated with 1 U of shrimp alkaline phosphatase (Amersham

Table 2 SBE primer sequences and concentrations of the SNPs used in this study, grouped by multiplexes

SNP		Minisequencing primer (5'→3')	Size (bp)	Conc (µM)	
M22	For	CCGCCATTCCTGGTGGCTCT	20	0.10	
P25	For	CCCCCCCTCTGCCTGAAACCTGCCTG	26	0.15	
92R7	Rev	CCCCGCATGAACACAAAAGACGTAGAAG	28	0.20	
SRY1532	For	CCCCCCTTGTATCTGACTTTTTCACACAGT	30	0.20	Multiplex 1
M70	Rev	CCCCCCCCTAGGGATTCTGTTGTGGTAGTCTTAG	34	0.15	
M173	For	CCCCCCCCCCTTACAATTCAAGGGCATTTAGAAC	34	0.20	
Tat	Rev	CCCCCCCCCCCCCCCCCCTCTGAAATATTAAATTAAAACAAC	42	0.20	
M213	Rev	CCCCCCCCCCCCCCCCCCCCCTCAGAACTTAAAACATCTCGTTAC	45	0.25	
M9	For	CCCCCCCCCCCCCCCCCCCCCCCCCGAAACGGCCTAAGATGGTTGAAT	48	0.20	
M170	Rev	ACACAACCCACACTGAAAAAAA	22	0.45	
M62	Rev	CCCCCCCCAATGTTTGTTGGCCATGGA	27	0.50	
M172	For	CCCCCCCCCCCCCCCAAACCCATTTTGATGCTT	32	0.10	Multiplex 2
M26	Rev	CCCCCCCCCCCCCCCCATAGGCCATTCAGTGTTCTCTG	37	0.25	
M201	For	CCCCCCCCCCCCCCCGATCTAATAATCCAGTATCAACTGAGG	42	0.05	
M34	Rev	TTGCAGACACACCACATGTG	20	0.15	
M81	For	CCCCCCTAAATTTTGTCCTTTTTTGAA	27	0.20	
M78	For	CCCCCCCCCCCACACTTAACAAAGATACTTCTTTC	34	0.35	
M35	Rev	CCCCCCCCCCCCCCCCCCCCCCAGTCTCTGCCTGTGTC	36	0.03	Multiplex 3
M96	For	CCCCCCCCCCCCGTAACTTGGAAAACAGGTCTCTCATAATA	40	0.05	
M123	Rev	CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCTTCTAGGTATTCAGGCGATG	51	0.35	
M167	For	CCCAAGCCCCACAGGGTGC	19	0.45	
M153	For	AAAGCTCAAAGGGTATGTGAACA	23	0.30	
M17	For	CCAAAATTCACTTAAAAAAACCC	23	0.20	
M18	For	CCCCAGTTTGTGGTTGCTGGTTGTTA	26	0.15	
M126	For	CCCCGCTTGAAACCGAGTTTGTACTTAATA	30	0.05	Multiplex 4
M37	For	CCGGAACACAGAAAATAAAATCTATGTGTG	30	0.35	
M73	Rev	CCCCCCCCCCCCGATTCCTCTGAACGTCTAACCA	33	0.30	
M65	Rev	CCCCCCCCCCCCCCCCCCCCCCCCCCACCCGCGGTAAG	36	0.05	
M160	For	CCCCCCCCTTACAAGTTTAATACATACAACTTCAATTTTC	40	0.20	

Biosciences) at 37°C for 1 h, and at 85°C for 15 min to inactivate the enzyme.

Electrophoretic detection

The products of the single base extension reactions were run on an ABI 3100 Genetic Analyser (Applied Biosystems). Analysis of electropherograms was performed using the GeneScan 3.7 software (Applied Biosystems), determining the size of the fragments based on GeneScan-120 LIZ size standards.

Statistical analysis

Binary marker haplogroup frequencies were calculated and Arlequin 2.0 software (Schneider et al. 2000) was used to test the hypothesis of a random distribution of the individuals between pairs of populations, through an exact test of population differentiation, and to calculate genetic distances, as pair-wise values of Φ_{ST}. A multidimensional scaling (MDS) analysis was performed using SPSS version 11.5 software package, with the genetic distances.

Results

The 30 SNPs were divided into 4 multiplex PCR/SBEs (Table 1), according to their location on the Y Chromosome Consortium tree. Multiplex 1 allows the detection of the more frequent major clades in Europe (Rosser et al. 2000; Semino et al. 2000; Brion et al. 2003), multiplex 2 determines haplogroups G, I and J, multiplex 3 subdivides haplogroup E, and multiplex 4 subdivides haplogroup R1b.

The products of the PCR multiplex reactions were designed to give different fragment sizes, allowing checking of the results in polyacrylamide gels. The sizes of the amplicons are represented in Table 1, and all fragments were unambiguously identified, even those fragments most similar in size, with only 3 bp differences (Fig. 2). None of the 292 samples analysed failed PCR amplification, therefore the multiplexes described seem to be a robust methodology for Y chromosome SNP typing.

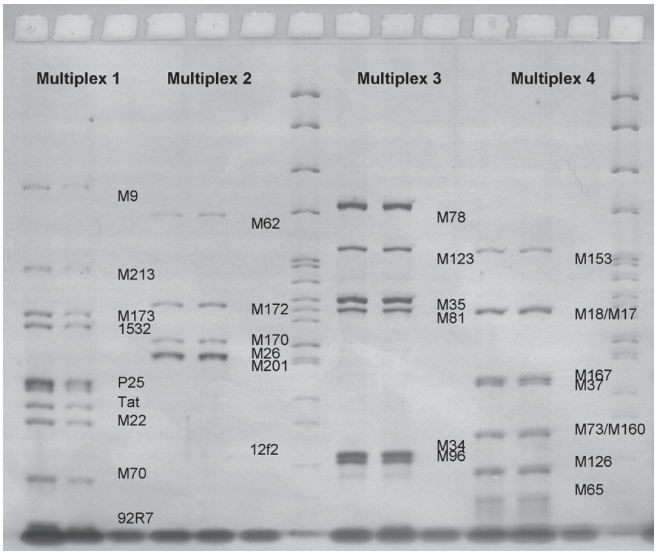

Fig. 2 Multiplex PCR products separated in a polyacrylamide gel and silver stained. Two samples and a negative control were run for each multiplex, in this order. *Lane* 7 and 14 show a pBR322 DNA-Msp I digest molecular weight standard (New England BioLabs). Multiplex 2 includes the 12f2 deletion, which was typed by presence/absence in the PCR amplification

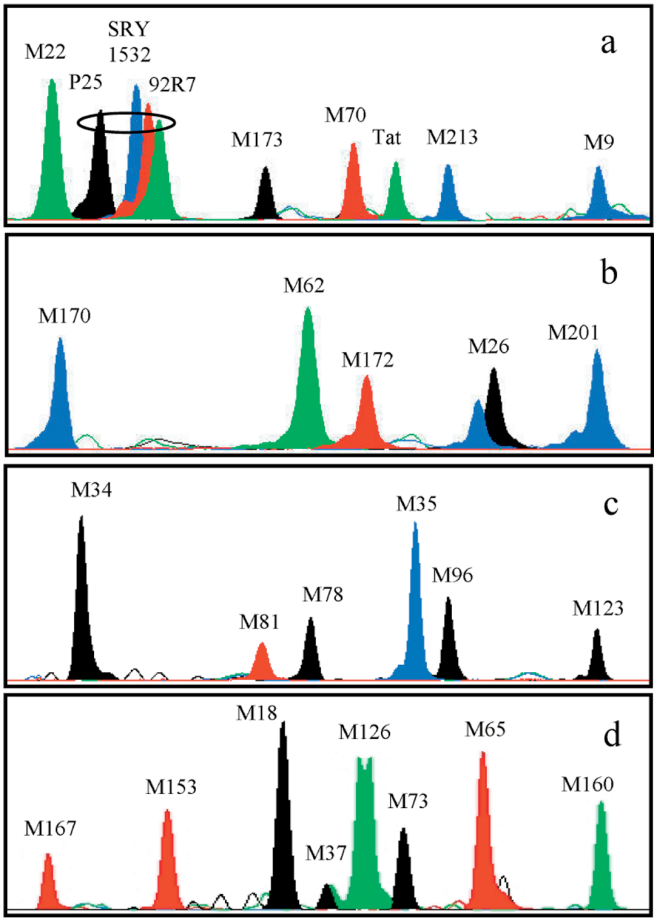

Fig. 3a–d Four SNaPshot multiplexes from different samples. **a** Multiplex 1 from sample assigned to haplogroup R1b, the P25 shows a duplicated pattern, **b** multiplex 2 from sample assigned to Hg I*(xI1b2), M26 always shows an artefactual blue peak, **c** multiplex 3 from sample assigned to Hg E3b2, **d**, multiplex 4 from sample assigned to Hg R1b3f

In the primer extension reactions most of the samples gave a full profile (Fig. 3), however, some failed to give a detectable peak for certain SNPs or gave a very weak signal. In all these cases the SNPs which failed primer extension, also showed a weak signal in the PCR amplification. This problem was always resolved by repeating the PCR to produce a better amplification signal for the SNPs.

Multiplex 1 is a combination of 9 SNPs (Fig. 3a), however in the SNaPshot results more than 9 different primer extension peaks could be seen. The reason is because two SNPs showing paralogous sequence variants (PSV), P25 and 92R7 (Sanchez et al. 2004) form part of the multiplex. SBE multiplex 2 comprises 5 SNPs (Fig. 3b), however the results always produced 6 primer extension peaks. De-

Table 3 Haplogroup frequencies and diversity in 8 different districts of Galicia

Haplotype	Galicia	Noroeste	Golfo Artabro	Mariña Lucense	Lugo	Santiago	Ourense	Rias Baixas	Montes Baixo Miño
E*(E3b)	0.0034	0	0	0	0.0164	0	0	0	0
E3b1	0.0205	0	0.0385	0	0.0164	0.0435	0	0	0.0714
E3b2	0.0411	0.0345	0	0	0.0820	0.0435	0.0541	0	0.0714
E3b3*(xE3b3a)	0.0034	0	0	0	0	0	0	0	0.0357
E3b3a	0.0103	0	0	0	0	0.0435	0	0.0323	0
G	0.0308	0.0690	0.0385	0.0588	0.0164	0.0435	0	0	0.0357
I*(xI1b2)	0.0959	0.1379	0.1154	0	0.0984	0.0870	0.0811	0.1290	0.1429
I1b2	0.0171	0.0345	0	0.0294	0	0	0.0270	0.0645	0
J*(xJ1,2)	0.0445	0.0345	0	0.1765	0.0164	0.0217	0.0541	0	0.0714
J2	0.1301	0.1724	0.1154	0.1471	0.1148	0.0870	0.1892	0.1290	0.1071
K2	0.0240	0	0	0.0294	0.0328	0.0435	0.0270	0	0.0357
R1*(xR1ab)	0.0103	0	0	0	0.0328	0.0217	0	0	0
R1a	0.0137	0	0.0385	0	0	0.0217	0.0270	0.0323	0
R1b*(xR1b1,2,3a-3f)	0.5377	0.5172	0.6538	0.5588	0.5738	0.5217	0.5405	0.5161	0.3929
R1b2	0.0034	0	0	0	0	0	0	0.0323	0
R1b3f	0.0137	0	0	0	0	0.0217	0	0.0645	0.0357
n	292	29	26	34	61	46	37	31	28
HgD	0.6806	0.6995	0.5631	0.6488	0.6486	0.7169	0.6757	0.7118	0.8228

tailed analysis indicated the SNP M26 has an artefactual blue peak a few bases smaller than the true SBE peak. Because M26 is a G to A transition and we are analysing the reverse strand, the artefact does not affect the interpretation of results in this case. Multiplex 3 and 4 comprise 6 and 8 SNPs, respectively (Fig. 3c,d), and none of them exhibited artefactual results.

In order to check the reproducibility of the four PCR/SBE multiplexes and to know the haplogroup composition of the Galician population (northwestern Spain), 292 samples, taken from locations scattered throughout the whole region, were analysed. The strategy adopted was to perform multiplex 1 in all of the samples, and depending on the results obtained, to continue with the appropriate additional multiplex to define the haplogroup more precisely.

All the results were completely consistent with the Y Chromosome Consortium tree (Y Chromosome Consortium 2002; Jobling and Tyler-Smith 2003). The 30 SNPs analysed describe a total of 32 haplogroups, however, for the 292 samples analysed only 16 of these haplogroups were detected (Fig. 1). The frequencies and haplogroup diversity values are represented in Table 3, and it can be seen that the highest diversity value was present in Montes Baixo Miño (0.823).

An exact test of sample differentiation based on haplogroup frequencies was also performed, and only Mariña Lucense was significantly differentiated from Lugo and Rias Baixas (P=0.0089±0.0021 and P=0.0065±0.0027, respectively).

We calculated genetic distances between populations as pairwise Φ_{ST} values. The distance matrix was represented in two-dimensional space using multidimensional scaling (Fig. 4). Golfo Artabro, Montes Baixo Miño and Rias Baixas appeared as the furthest outliers, while Lugo and Ourense appeared as the closest populations.

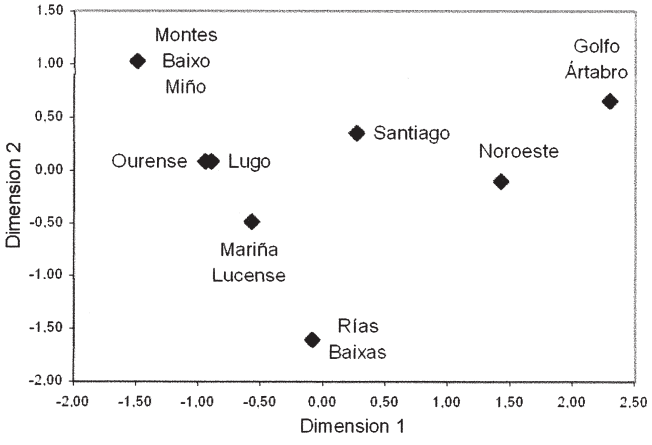

Fig. 4 Multidimensional scaling plot of Galician populations, from pairwise Φ_{st} distances based on Y-chromosome haplogroup frequencies. Stress value =0.003

Discussion

We have developed a strategy for Y-chromosome SNP typing, which allows the quick assignment of a sample to one of 32 haplogroups defined by 30 SNPs, after only 1 or 2 multiplex reactions.

Despite the presence of some ambiguous positions (Weale et al. 2003), the phylogenetic tree of binary Y-chromosomal haplogroups (Y Chromosome Consortium 2002; Jobling and Tyler Smith 2003) has a stable structure at this moment, allowing researchers to assign the haplogroup of any sample, through a hierarchical analysis following the branches of the tree. This avoids unduly extensive SNP typing, saving time, costs and target DNA. The only disadvantage is the possibility of missing occa-

sional recurrent mutations, which are rare, due to the low mutation rate of the SNPs.

From a technical point of view, in the case of PCR multiplexes no problems emerged. The development of the PCR reaction was performed according to the guidelines of Henegariu et al. (1997), trying to establish similar melting temperatures for all the primers and to use the same standard conditions, in order to allow the incorporation of new SNPs or the regrouping of the SNPs used. In the development of the primer extension reactions, crucial factors are the purity of the primers, (HPLC purification is highly recommended) and the cleaning of the PCR product, with the minimal possible residual primers and dNTPs from the previous PCR (Sanchez et al. 2003). In our experience there is a clear correlation between the PCR results and the primer extension results; allele designation was been possible for all the markers in all the samples. However, all the cases of weak PCR amplification resulted in primer extension results that were difficult to interpret.

The results always correlated with the Y Chromosome Consortium tree. Haplogroup composition and frequencies are in agreement with previous publications (Rosser et al. 2000; Semino et al: 2000; Brion et al. 2003) including European studies. Once more, the high degree of population homogeneity present in Europe has been confirmed, since more than half the samples belong to the same paragroup (157 individuals belong to paragroup P). In addition, when the diversity was checked among the different districts belonging to Galicia, no significant differences were detected, except for the population from Mariña Lucense, which showed a distant haplogroup frequency but Φ_{st} values comparable to the other regions (Fig. 4). It is clear that in a microgeographical study of binary haplogroups in a general population, no significant differentiation is expected. But this general pattern is not always true, and examples of clear differentiation at a local level have been shown in the literature (Brion et al. 2003). This can have an important impact in the forensic evaluation of the Y chromosome evidence and for this reason it is important to check for possible differentiation at a local level.

With a hierarchical strategy adapted for European populations extensive typing of SNPs was avoided, and therefore the time and cost involved in the study was reduced. Primer extension reactions using the SNaPshot multiplex kit, allowed with previous multiplex amplification and with an automatic sequencer, the quick development of a sufficiently large study, without investment in new technologies.

Despite the fact that Y-chromosome SNP analysis is becoming increasingly accessible and the number of polymorphisms detected is being extended even further, the population-specific diversity exhibited by these polymorphisms should be interpreted with great caution. In comparison to SNPs, microsatellites, which are variable in all populations, provide a less biased measure of diversity. This must always be borne in mind in forensic analysis, when the haplogroup frequency of a random sample is estimated and interpreted according to anthropological data.

Acknowledgements The technical assistance of Meli Rodriguez and Raquel Calvo is highly appreciated. This work was supported by the grant from the "Ministerio de Ciencia y Tecnologia" (DGCYT.P4.BIO2000-09822).

References

Borsting C, Sanchez JJ, Morling N (2004) Multiplex PCR, amplicon size and hybridization efficiency on the NanoChip electronic microarray. Int J Legal Med (in press)

Brion M, Salas A, González-Neira A, Lareu MV, Carracedo A (2003) Insights over the Iberian population origin through the construction of highly informative Y-chromosome haplotypes using biallelic markers, STRs and the MSY1 minisatellite. Am J Phys Anthropol 122:147–161

Chen X, Sullivan PF (2003) Single nucleotide polymorphism genotyping: biochemistry, protocol, cost and throughput. Pharmacogenom J 3:77–96

Gill P (2001) An assessment of the utility of single nucleotide polymorphisms (SNPs) for forensic purposes. Int J Legal Med 114:204–210

Henegariu O, Heerema NA, Dlouhy SR, Vance GH, Vogt PH (1997) Multiplex PCR: critical paremeters and step-by-step protocol. Biotechniques 23:504–511

Jobling MA, Tyler-Smith C (2003) The human Y chromosome: an evolutionary marker comes of age. Nat Rev Genet 4:598–612

Jorde LB, Watkins WS, Bamshad MJ (2001) Population genomics: a bridge from evolutionary history to genetic medicine. Hum Mol Genet 10:2199–2207

Karafet TM, Zegura SL, Posukh O et al. (1999) Ancestral Asian source(s) of New World Y-chromosome founder haplotypes. Am J Hum Genet 64:817–831

Rosser Z, Zerjal T, Hurles MH et al. (2000) Y chromosomal diversity within Europe is clinal and influenced primarily by geography, rather than language. Am J Hum Genet 67:1526–1543

Sanchez JJ, Borsting C, Hallenberg C, Buchard A, Hernandez A, Morling N (2003) Multiplex PCR and minisequencing of SNPs – a model with 35 Y chromosome SNPs. Forensic Sci Int 137: 74–84

Sanchez JJ, Brion M, Parson W et al. (2004) Duplications of the Y-chromosome specific loci P25 and 92R7 and forensic implications. Forensic Sci Int 140:241–250

Schneider S, Roessli D, Excoffier L (2000) Arlequin ver. 2.000: a software for population genetics data analysis. Genetics and Biometry Laboratory, University of Geneva, Switzerland

Seielstad MT, Minch E, Cavalli-Sforza LL (1998) Genetic evidence for a higher female migration rate in humans. Nat Genet 20:278–280

Semino O, Passarino G, Oefner PJ et al. (2000) The genetic legacy of paleolithic Homo sapiens sapiens in extant Europeans: a Y chromosome perspective. Science 290:1155–1159

Weale ME, Shah T, Jones AL et al. (2003) Rare deep-rooting Y chromosome lineages in humans: lessons of phylogeography. Genetics 165:229–234

Y Chromosome Consortium (2002) A nomenclature system for the tree of human Y-chromosomal binary haplogroups. Genome Res 12:339–348

Zhao Z, Fu YX, Hewett-Emmett D, Boerwinkle E (2003) Investigating single nucleotide polymorphism (SNP) density in the human genome and its implications for molecular evolution. Gene 312:207–213