PCR typing of the locus D17S30 (YNZ22 VNTR) in an Italian population sample

Int J Leg Med (1994) 106:200–204

International Journal of
Legal Medicine

PCR typing of the locus D17S30 (YNZ22 VNTR) in an Italian population sample

L. Buscemi[1], N. Cucurachi[1], R. Mencarelli[2], B. Sisti[1], A. Tagliabracci[2], S. D. Ferrara[1]

[1] Istituto di Medicina Legale dell'Università di Parma, Ospedale Maggiore, I-43100 Parma, Italy
[2] Istituto di Medicina Legale dell'Università di Ancona, Ospedale Regionale, I-60020 Torrette, Ancona, Italy

Received May 21, 1993 / Received in revised form September 2, 1993

Summary. A sample of 202 subjects living in 2 Italian provinces (Ancona and Parma) was tested for YNZ22 polymorphism by the polymerase chain reaction (PCR). After amplification, the phenotypes were separated by agarose gel electrophoresis, stained with ethidium bromide and identified by comparison with a molecular weight marker. No heterogeneity was found between the 2 populations. Alleles, pooled in 4 groups to calculate the Hardy-Weinberg (H-W) equilibrium, showed good accordance between observed and expected values. The power of discrimination (PD) was 0.95 and the chance of exclusion was 0.69. The allele comparison with previous studies on Caucasians showed no significant difference.

Key words: DNA polymorphism – PCR – Population genetics – YNZ22 – D17S30 locus

Zusammenfassung. Eine Populationsstichprobe von 202 Personen, welche in zwei italienischen Provinzen (Ancona und Parma) leben, wurde mit Hilfe der Polymerase-Kettenreaktion (PCR) auf den Polymorphismus YNZ22 untersucht. Nach Amplifikationen wurden die Phänotypen mit Hilfe der Agarose-Gel-Elektrophorese getrennt und mit Ethidiumbromid angefärbt und durch den Vergleich mit einem Molekulargewichts-Marker identifiziert. Zwischen beiden Populationen bestand keine Heterogenität. Die Allele, welche in 4 Gruppen zusammengefaßt wurden, um das Hardy-Weinberg-Gleichgewicht zu überprüfen, zeigten eine gute Übereinstimmung zwischen beobachteten und erwarteten Werten. Die Diskriminationskraft war 0,95 und die Ausschließungschance war 0,69. Der Allel-Vergleich mit vorhergehenden Untersuchungen an Kaukasiern zeigte keinen signifikanten Unterschied.

Schlüsselwörter: DNA-Polymorphismus – PCR – Populationsgenetik – YNZ22 – D17S30 locus

Correspondence to: S. D. Ferrara, Istituto di Medicina Legale, via Falloppio 50, I-35121 Padova, Italy

Introduction

Polymorphism at the D17S30 locus (YNZ22 system) was first discovered by Wolff et al. (1988) and then further investigated by Horn et al. (1989) using the PCR technique (Saiki et al. 1985). The polymorphism is due to a core sequence of 70 bp tandemly repeated a variable number of times. The variation at this locus has been studied by both Southern Blotting and PCR techniques. Odelberg et al. (1989) identified 14 different alleles as a result of RFLP (restriction fragment length polymorphism) analysis. AMP-FLP (amplifiable fragment length polymorphism) analysis by PCR revealed up to 12 alleles (Batanian et al. 1990) and more recently Deka et al. (1992), using both techniques, found 15 alleles. Because of its high polymorphism this system is a helpful tool in identity testing for forensic purposes. The power of discrimination (PD) in a sample of 88 Caucasians was calculated to be 0.93 and the exclusion rate for paternity 0.67 (Rand et al. 1992).

However further population studies are necessary before this conclusion can be extended to the general population and a database formed. To our knowledge no population data on this system are available in Italy. For this reason we investigated a representative population sample from 2 areas in North and Central Italy with the following aims: 1) to increase the number of populations examined to form a data base, 2) to compare the allelic frequencies found with those reported for other countries, checking possible differences between different samples.

Materials and methods

Fresh blood samples were collected from healthy unrelated donors living in Parma ($n = 100$) and Ancona ($n = 102$), 2 cities located in North and Central Italy respectively.

Both laboratories followed the same procedures for DNA extraction, amplification, separation and visualisation of phenotypes.

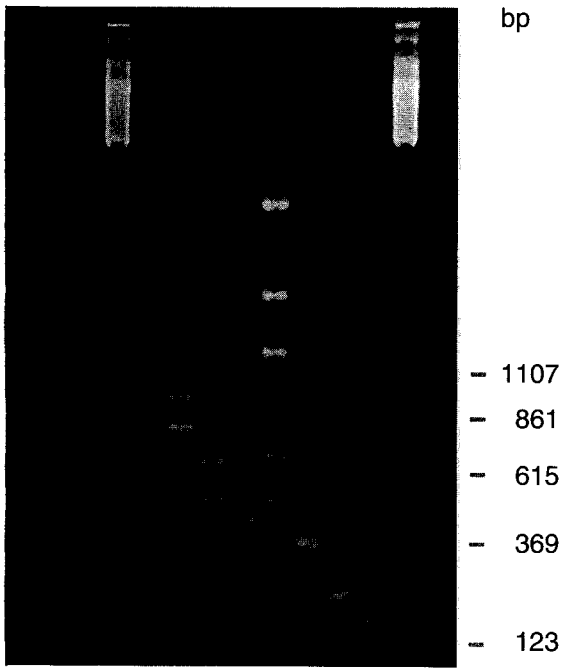

Fig. 1. YNZ22 phenotypes. Separation of 10 µl of amplified products in 1.5% agarose gel in TAE at 50 V for 16 h. Staining with ethidium bromide. From right to left: 123 bp weight marker BRL, 1–3, 2–11, 4–9, pGEM weight marker, 5–7, 6–8, 10–12, 3–13, 123 bp weight marker BRL

DNA was extracted using the phenol-isoamyl alcohol method developed by the FBI (Budowle and Baechtel 1990).

Amplification was carried out in a thermal cycler PTC 100–60 (MJ Research, USA) with the following conditions:

denaturation at 94°C for 1 min
annealing at 63°C for 1 min
extension at 72°C for 1 min

for 27 cycles with an initial denaturation at 94°C for 5 min and final extension at 63°C for 11 min.

The primers used were those proposed by Budowle (Rand et al. 1992):

5′-AAACTGCAGAGAGAAAGGTCGAAGAGTGAAGTG-3′
5′-AAAGGATCCCCCACATCCGCTCCCCAAGTT-3′

The amplification mix contained 50 ng DNA, 10 m*M* Tris-HCl pH 8.3, 50 m*M* KCl, 1.5 m*M* $MgCl_2$, 0.001% Triton X-100, 2 U (0.4 µl) Taq DNA-polymerase (Promega), 1 µ*M* each primer, 200 µ*M* each dNTP made up to a total volume of 50 µl and overlayed with 30 µl oil.

Electrophoresis of 10 µl of amplified product was performed in 1.5% agarose (Ultra-pure DNA grade agarose, Bio-Rad) gels in 1X TAE buffer at 50 V for 16 h. The bands were visualized directly on the gel by adding ethidium bromide at a concentration of 1 mg/ml. The phenotypes were identified by comparison with the 123 bp BRL ladder (Fig. 1).

The comparison between the observed genotypes in the 2 Italian cities was performed by a computer program kindly supplied by G. Carmody (Carleton University, Ottawa, Canada). The same program was used to compare allelic observations between Italians and other populations.

As suggested by Rand et al. (1992), the alleles were pooled into 4 groups to calculate the chi-square between observed and expected genotype frequencies assuming the Hardy-Weinberg law

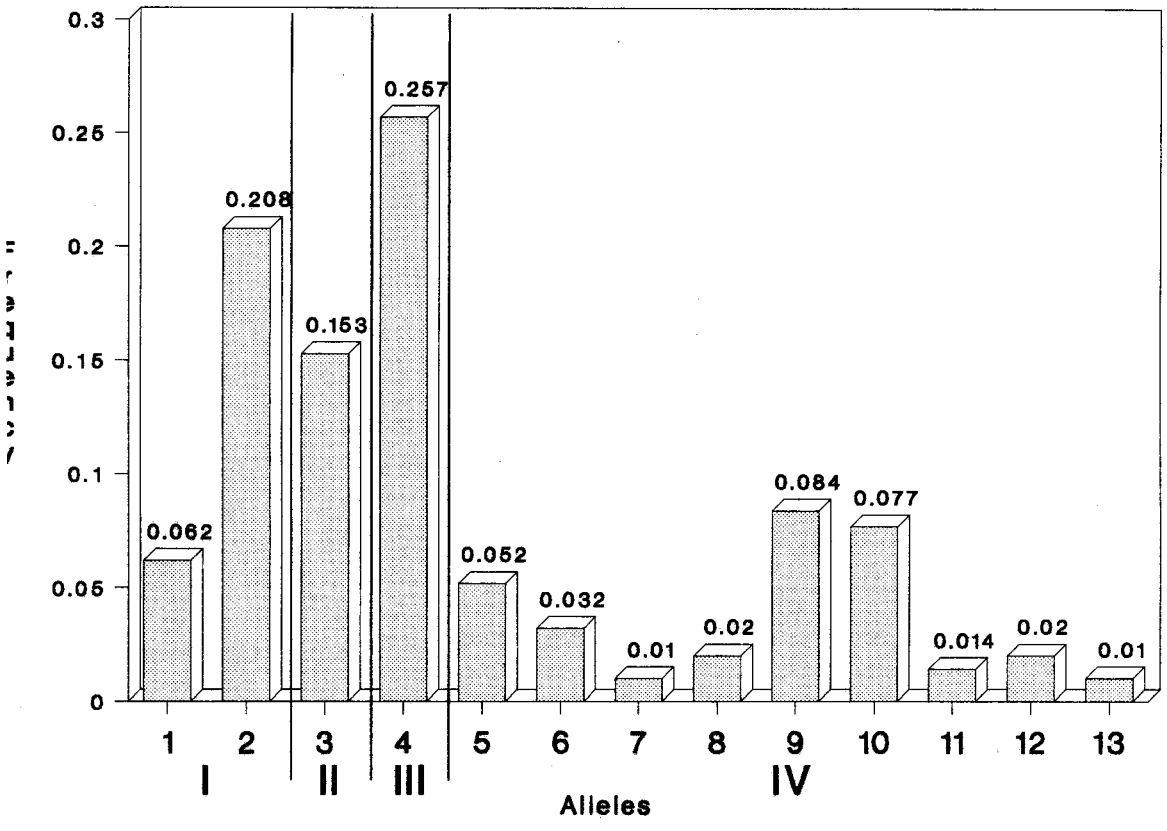

Fig. 2. Frequency distribution of YNZ22 system alleles in a population sample of 202 unrelated Italians

(Fig. 2). The results were also verified by comparing the observed and expected heterozygosity frequencies. The expected value is equivalent to allelic diversity, calculated as follows:

$H = [1 - \Sigma X_i^2]\ [n/n - 1]$, where X_i is the allelic frequency and n the number of alleles (Nei and Roychoudhury 1974) and $\sqrt{[h(1-h)/N]}$ is the formula to compute the standard error for H, were h is the expected heterozygote frequency and N the number of subjects examined. The power of discrimination (PD) was calculated using the equation proposed by Fischer (1951): $1 - \Sigma(Pi)^2$, where Pi represents the frequency of each genotype. The exclusion chance was calculated from allele frequencies (Garber and Morris 1983).

Results

The allelic frequencies of the YNZ22 system from 2 populations living in 2 different areas of Italy are shown in Table 1. The 2 populations were tested for heterogeneity with an R × C contingency table and the results indicated that the genotype frequencies were very similar ($P = 0.7030 \pm 0.0322$). So from this moment on the 2 samples will be considered as a single Italian population sample.

A total of 54 genotypes corresponding to 13 alleles were found in the 202 subjects examined (Table 2). The most frequent alleles were, in decreasing order, YNZ22*4, YNZ22*2, YNZ22*3, YNZ22*9,YNZ22*10, YNZ22*1 and YNZ22*5, while the remaining alleles were much less common. In particular, the frequency of YNZ22*7 and YNZ22*13 was below 1% (Fig. 2). A large allele, temporarily named >13 and still under investigation, was also found.

The limited size of our population sample compared to the high number of alleles and expected genotypes can cause some problems when checking whether the Hardy-Weinberg law is respected, as discussed by Rand et al. (1992). To overcome such difficulties, the chi-square test was performed on alleles binned into 4 classes, as suggested by these authors. Following this approach the chi-

Table 1. Allelic frequencies in two Italian populations

YNZ22 Allele	Ancona 102 individuals	Parma 100 individuals
1	0.088	0.035
2	0.206	0.210
3	0.132	0.175
4	0.240	0.275
5	0.064	0.040
6	0.054	0.010
7	0.005	0.015
8	0.024	0.015
9	0.064	0.105
10	0.083	0.070
11	0.024	0.005
12	0.000	0.040
13	0.015	0.005

Heterogeneity test (observed genotypes): $\chi^2 = 49.7066$, $P = 0.7079 \pm 0.0320$; G statistic 64.4868, $P = 0.7030 \pm 0.0322$

Table 2. Observed YNZ22 genotypes

Genotype	n.	Frequency
1–1	2	0.0099
1–2	6	0.0297
1–3	3	0.0149
1–4	7	0.0347
1–5	1	0.0050
1–7	1	0.0050
1–8	1	0.0050
1–10	1	0.0050
1–13	1	0.0050
2–2	9	0.0446
2–3	13	0.0644
2–4	23	0.1139
2–5	5	0.0248
2–6	2	0.0099
2–8	1	0.0050
2–9	8	0.0396
2–10	6	0.0297
2–11	2	0.0099
3–3	9	0.0446
3–4	11	0.0545
3–5	2	0.0099
3–6	1	0.0050
3–9	3	0.0149
3–10	5	0.0248
3–11	1	0.0050
3–12	4	0.0198
3–13	1	0.0050
4–4	17	0.0841
4–5	6	0.0297
4–6	3	0.0149
4–8	2	0.0099
4–9	11	0.0545
4–10	3	0.0149
4–11	2	0.0099
4–12	1	0.0050
4–13	1	0.0050
5–6	1	0.0050
5–7	1	0.0050
5–9	2	0.0099
5–10	3	0.0149
6–6	1	0.0050
6–7	1	0.0050
6–8	1	0.0050
6–10	1	0.0050
6–13	1	0.0050
7–10	1	0.0050
8–8	1	0.0050
8–9	1	0.0050
9–9	2	0.0099
9–10	4	0.0198
9–12	1	0.0050
10–10	2	0.0099
10–11	1	0.0050
10–12	2	0.0099

Table 3. Single alleles binned together to calculate the chi square between observed and expected genotypes under Hardy-Weinberg law expectations

Grouping	Observed	Expected	Chi square
I–I	17	14.70	0.358
I–II	16	16.73	0.031
I–III	30	28.06	0.134
I–IV	29	34.80	0.968
II–II	9	4.75	3.783
II–III	11	15.96	1.541
II–IV	17	19.80	0.395
III–III	17	13.39	0.975
III–IV	29	33.21	0.533
IV–IV	27	20.60	1.991
$\chi^2 = 10.713$	$0.2 < P < 0.3$	d.f. = 9	

square test carried out on these bins shows good agreement between observed and expected values ($0.20 < P < 0.30$) (Table 3). The heterozygosity rate was 0.787 while the allelic diversity was 0.8473 ± 0.025.

The chance of exclusion was 0.69 and the PD was 0.95. Our data were compared with those reported for a sample of 254 Germans (S. Rand pers. com.) for population sample homogeneity using a 2-way RXC contingency table. No heterogeneity was found between allele frequencies ($\chi^2 = 18.1176$, $P = 0.0998 \pm 0.0099$; G statistic = 18.5501, $P = 0.1086 \pm 0.0103$).

Discussion

YNZ22 is an AMP-FLP system which displays a high number of alleles in the few populations so far investigated by PCR techniques. Agarose gel electrophoresis with ethidium bromide staining, as performed in this study, is an efficient method to identify alleles because they differ from each other by multiples of a 70 bp repeat unit and therefore their identification by comparison with a molecular weight marker of suitable size does not encounter any major problem.

In this study 202 subjects from 2 Italian regions were investigated with the same procedures and the 13 known alleles were identified in the range 168–1008 bp. Another larger allele, temporarily designated >13, is under investigation. The same 13 alleles were found by S. Rand (pers. com.), while Deka et al. (1992) also detected a further allele of 1078 bp, corresponding to allele 14, and a large allele composed of 24 or 25 repeats. Since the latter allele, named >14, does not correspond to the >13 we are investigating, the total number of alleles identified up to now is therefore sixteen.

The most frequent alleles found were the small ones: YNZ22*4, YNZ22*2 and YNZ22*3. Of the larger alleles, YNZ22*9 and YNZ22*10 were most often detected. This distribution is in line with those found in previous works on Caucasians and other ethnic groups (Batanian et al. 1990; Rand et al. 1992; Rand et al. pers com.; Deka et al. 1992).

Heterozygote deficiency from the Hardy-Weinberg equilibrium for VNTR systems may be due to inaccuracy of analytical methods, particularly for systems identified by Southern blotting, as well as the composition of the population examined. Furthermore, another major problem for YNZ22 is the preferential amplification of shorter alleles, also known as the dropout phenomenon (Kloosterman et al. 1992). Two biostatistical tests were used to verify this condition: the chi-square between observed and expected genotype frequencies calculated on the basis of Hardy-Weinberg equilibrium from the allele frequencies binned into 4 groups, and the unbiased estimate of the expected heterozygote frequency. The results showed that the Hardy-Weinberg law was satisfied, while there was a significant difference between expected and observed heterozygotes, suggesting that preferential amplification does in fact occur in the YNZ22 system. To attempt to eliminate this problem, homozygotes were systematically retested varying amplification conditions and PCR reaction mix, but without significant differences in the results (data not shown).

To perform a statistical comparison between the 2 Italian populations we used a program to calculate the probability of obtaining an observed 2-way contingency table even when some cells have less than 5 observed genotypes, as in our population sample. No heterogeneity was found in the distribution of observed genotypes from the 2 different Italian areas examined in this study. The same test was applied to compare Italians with Germans (Rand et al. pers. com.). In this case the comparison was made between allele frequencies, because the genotypes observations were not known. Italians appear to be statistically similar to Germans ($P > 0.05$).

In conclusion, the representative sample of Italians studied for the YNZ22 system shows allelic frequencies similar to the other Caucasian populations studied so far. The high power of discrimination and chance of exclusion make this system a powerful and useful tool for forensic identification and paternity analysis. However when analysing the results of such studies it is important to consider the possible influence of the phenomenon of preferential amplification which can bias the results.

References

Batanian JR, Ledbetter SA, Wolff RK, Nakamura Y, White R, Dobyns WB, Ledbetter DH (1990) Rapid diagnosis of Miller-Dieker syndrome and isolated lissencephaly sequence by the polymerase chain reaction. Hum Genet 85:555–559

Budowle B, Baechtel FS (1990) Modifications to improve effectiveness of restriction fragment length polymorphism typing. Appl Theor Electrophor 1:181–187

Deka R, De Croo S, Yu LM, Ferrell RE (1992) Variable number of tandem repeat (VNTR) polymorphism at locus D17S5 (YNZ22) in four ethnically defined human populations. Hum Genet 90:86–90

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

Garber RA, Morris JW (1983) General equations for the average power of exclusion for genetic system of n codominant alleles in one-parent and no-parent cases of disputed parentage. In:

Walker RH (ed) Inclusion probabilities in parentage testing. AABB, Arlington, VA, pp 277–280

Horn GT, Richards B, Klinger KW (1989) Amplification of a highly polymorphic VNTR segment by the polymerase chain reaction. Nucleic Acids Res 17:2140

Kloosterman A, Vossen R, Wust D, de Leeuw W, Uitterlinden A (1992) Detection of three different VNTR's by DNA-amplification. In: Rittner C, Schneider PM (eds) Advances in forensic haemogenetics 4, Springer, Berlin Heidelberg, pp 53–56

Nei M, Roychoudhury AK (1974) Sampling variances of heterozygosity and genetic distance. Genetics 76:379–390

Odelberg SJ, Plaetke R, Eldridge JR, Ballard L, O'Connell P, Nakamura Y, Leppert M, Lalouel JM, White R (1989) Characterization of eight VNTR loci by agarose gel electrophoresis. Genomics 5:915–924

Rand S, Puers C, Skowasch K, Wiegand P, Budowle B, Brinkmann B (1992) Population genetics and forensic efficiency data of 4 AMPFLP's. Int J Leg Med 104:329–333

Saiki RK, Scharf S, Faloona F, Mullis KB, Horn GT, Erlich HA, Arnheim N (1985) Enzymatic amplification of ß-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. Science 230:1350–1354

Wolff RK, Nakamura Y, White R (1988) Molecular characterization of a spontaneously generated new allele at a VNTR locus: no exchange of flanking DNA sequence. Genomics 3:347–351