

ALLELE FREQUENCIES OF D1S80, LDLR, GYPA, D7S8, GC, HBGG AND SE 33 IN POLISH POPULATION SAMPLE

Miścicka-Śliwka D., Śliwka K., Syroczyńska A., Grzybowski T., Baranowska B., Berent J.A.

Forensic Medicine Institute, The Ludwik Rydygier's University School of Medical Sciences in Bydgoszcz, ul. M. Skłodowskiej-Curie 9, 85-094 Bydgoszcz, Poland

INTRODUCTION

The application of any new genetic system in forensic casework requires a sufficiently large data base of population. In the present study, we report the allele frequency distribution for seven different loci in the Polish population of Pomerania-Kujawy Region. Five of these markers are coamplified simultaneously using the Amplitype PM kit (Perkin-Elmer): low density lipoprotein receptor (LDLR), glycoprotein A (GYPA), hemoglobin G gammaglobin (HBGG), D7S8 linked to cystic fibrosis locus and Group specific Component (GC). The other two loci are VNTR D1S80 (Nakamura 1988) and STR ACTBP2 (SE33) (Polymeropoulos 1992). In addition, we have sequenced selected SE 33 alleles to determine possible sequence structure variations which may cause the allele designation difficult.

MATERIALS AND METHODS

DNA EXTRACTION

Blood samples were obtained from unrelated individuals. Sample sizes were 306 for D1S80, 187 for PM and 31 for SE 33 analysis. DNA was isolated by Proteinase K/Phenol/Chloroform extraction method followed by absolute ethanol precipitation.

PCR AMPLIFICATION

PCR reactions for D1S80 and PM were performed with commercial kits and procedure. SE 33 amplification was performed using 5 ng template in a 10 µl reaction volume comprising 1x Promega buffer, 1.5 mM MgCl₂, 200 µM of each dNTP (Perkin-Elmer), 0.7 µM of each primer and 0.5 U of Taq DNA Polymerase (Promega). Primer sequences were used according to Polymeropoulos (1992). Samples were amplified for 10 cycles of 10 s at 94° C, 66° C and 72° C and then the same denaturing and extension conditions, but an annealing temperature of 64° C for an additional 20 cycles on a 9600 thermal cycler.

PCR PRODUCTS VISUALISATION

The analysis of PM alleles was based on reverse dot blot reactions with oligonucleotide probes. D1S80 alleles were separated by electrophoresis on ultrathin layer horizontal polyacrylamide gels 12% T, 3% C with DATD as a crosslinker. Gels were run at constant temperature of 10° C and constant current 15 mA. SE 33 alleles were diagnosed by vertical electrophoresis in 0.8 mm thick 8% Tris-sulfate polyacrylamide gels run at a constant 500 V. Typing was performed by comparison with a sequenced allelic ladder kindly donated by Prof. B. Brinkmann (Institute of Legal Medicine, Munster). Bands were visualised by silver staining (Allen 1989) and recorded by computer imaging.

CYCLE SEQUENCING

For sequencing reactions, DNA (100 ng) was amplified in a volume 100 µL consisting of 1x Promega buffer, 0.7 mM MgCl₂, 200 µM of each dNTP, 0.5 µM of each primer and 2 U Taq DNA Polymerase. To increase the size of PCR products, new primer sequences were designed: F: 5' GAG CCG AGG TCA TGC CAT TG 3' ; R: 5' GAC GAC GAG CGC GGT GAT AG 3'. After the Hot Start at 94° C for 2 minutes samples were amplified for 29 cycles of 30 s at 94° C, 30 s at 64° C and 90 s at 72° C (Perkin-Elmer 9600). Cycle sequencing reactions were carried out using the AutoCycle Sequencing Kit (Pharmacia), 160-180 ng PCR template and 0.5 µM internal (nested) sequencing primer (ACTBP2-primer 1: 5' AAT CTG GGC GAC AAG AGT GA 3'), fluorescently labelled at 3' end (Markiewicz 1992). Amplification conditions were: Hot Start at 94° C-2 min, 94° C-15 s, 64° C-15 s, 72° C-40 s; 36 cycles, followed by final elongation at 72° C-10 min. The sequencing products were separated in a 6% 0.35 thick denaturing gel which was run for 200 min at 1500 V, 40 mA and 40 W on the A.L.F. DNA Sequencer (Pharmacia).

RESULTS AND DISCUSSION

POPULATION DATA

Allele frequency distribution of PM loci is given in Fig. 1. The most common alleles were LDLR B ($f = 0.623$) and D7S8 A ($f = 0.591$). Our results have shown that in locus LDLR, AB and AA genotypes are the most and the least frequent, respectively. In GYPA, HBGG and D7S8 markers the most frequent genotype was AB. It clearly corresponds to earlier data obtained for Swiss, U.S. Caucasian and U.S. Hispanic populations (Hochmeister 1994; kit leaflet by Roche Molecular Systems). In locus GC however, we have found AC genotype being the most frequent. It correlates with the data obtained for Swiss and U.S. Caucasian population samples, but differs from U.S. Hispanic frequencies. What is more interesting, results obtained for Portugal population sample ($n = 115$) (Pinheiro 1994) also indicate, as in the case of our data, AC as the most frequent genotype in GC locus. We hope that further interlaboratory comparisons and validation experiments will explain this striking difference. Heterozygosity Index (HI) for these markers was in the range of 0.51-0.79 and the combined power of discrimination (PD) value was approximately 0.94. The genotype distribution was in Hardy-Weinberg equilibrium for all the five loci. These findings, if followed by estimation of mutation rates and level of sensitivity of the system in the cases when heavily degraded DNA is available, seem to validate PM loci for paternity and forensic casework.

In our D1S80 population study 21 alleles were observed. Fig. 2 shows their frequency distribution. The most common were 18 ($f = 0.209$) and 24 ($f = 0.322$). It clearly corresponds to our previous studies. The comparison of our allele frequencies to population data from Western Europe shows certain similarities. Some existing differences are probably caused by the size of the samples, but it is confirmed that 18 and 24 alleles are the most frequent in D1S80 system. In addition to expected PCR products, the "interalleles" (19^m, 20^m, 22^m, 24^m) were found. The distribution of observed and expected D1S80 phenotypes is in Hardy-Weinberg equilibrium. The observed heterozygosity was 0.80 (expected 0.97) and power of discrimination value was 0.95. The high level of polymorphism and high HI and PD values make this marker very useful in forensic blood identification and alleged paternity examination.

SE 33 allele frequencies are given in Fig. 3. Seventeen different alleles were observed. The most common were 19 ($f = 0.145$), 26 and 28 ($f = 0.113$). In our population sample, 28 genotypes were observed with the genotypes 23/28 and 26/28 being the most frequent ($f = 0.065$). Heterozygosity Index was 0.94 in comparison with expected heterozygote frequency $h = 0.966$. The population data clearly corresponds to those obtained by Polymeropoulos (1992) for small U.S. population ($n = 39$) but show a striking difference to data from German sample (Wiegand 1993). Different sizes in both studies ($n = 31$ and $n = 180$) probably could explain the differences observed. Power of discrimination for our sample was 0.96. It confirmed the value of this STR system as highly polymorphic and sensitive marker for human identification.

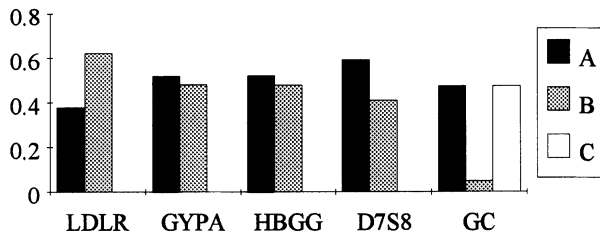


Fig. 1. Frequencies of PM alleles in population of Pomerania-Kujawy region of Poland.

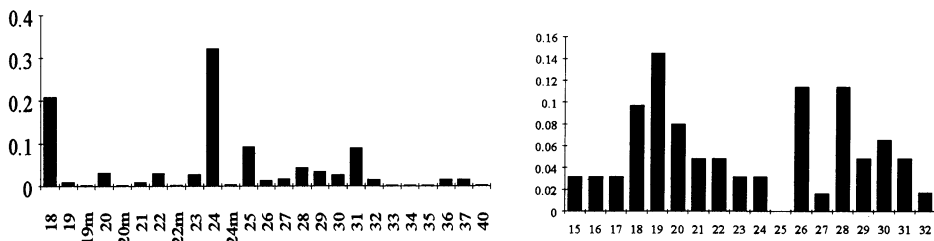


Fig. 2. (left) Frequencies of D1S80 alleles in population of Pomerania-Kujawy region of Poland.

Fig. 3. (right) Frequencies of SE33 alleles in population of Pomerania-Kujawy region of Poland.

