

## ALLELE FREQUENCIES OF HIGHLY POLYMORPHIC LOCUS D8S1132 IN A POPULATION SAMPLE FROM NORTH POLAND

Ryszard PAWŁOWSKI<sup>1,2</sup>, Marta ŻYWICKA<sup>3</sup>

<sup>1</sup>*Chair and Department of Forensic Medicine, Medical Academy, Gdańsk*

<sup>2</sup>*Institute of Forensic Research, Cracow*

<sup>3</sup>*Intercollegiate Faculty of Biotechnology, Medical Academy, Gdańsk, and University of Gdańsk*

**ABSTRACT:** This paper presents the alleles frequency distribution for the D8S1132 locus in a population sample of 158 unrelated individuals of both sexes living in north Poland. PCR products were identified on a horizontal polyacrylamide gel stained with silver. The presence of the D8S1132\*15 allele, which had not been observed till now, was ascertained using capillary electrophoresis on an ABI310 DNA sequencer. In the analysed population sample, the presence of 10 D8S1132 alleles was observed. Owing to the very high discrimination power (0.9622) of the analysed locus, it is highly useful for forensic identification purposes.

**KEY WORDS:** D8S1132 locus; STR; PCR; Population genetics; North Poland.

*Z Zagadnień Nauk Sądowych, z. XLVIII, 2001, 74–82*

*Received 6 August 2001; accepted 14 November 2001*

### INTRODUCTION

Among the many VNTR-STR loci described till now, those characterised by high discrimination power [1] and the presence of short alleles are considered especially useful for forensic genetics. These two features make them particularly useful for DNA analysis based on amplification of extremely degraded biological material. In a recent paper Wiegand et al. [11] gave a detailed description of the degree of polymorphism of alleles and presented sequencing data for alleles of the new locus D8S1132 (8pter-8qpter). This locus is characterised by a composite repetitive sequence of the type (TCTA)<sub>n</sub>TCA (TCTA)<sub>n</sub>. Up till now ten alleles (alleles 16 to 25) in the length range from 134–170 bp have been described. The analysed STR system is very sensitive, giving the possibility of gaining an amplification signal from as little as 20–50 pg of template DNA [10].

The main goal of this work was to research allele frequencies of locus D8S1132 in a population sample from northern Poland, and to compare ob-

tained population data with literature data for German and north-eastern Polish populations.

## MATERIALS AND METHODS

The material used for the population study was blood samples and buccal swabs which had been collected from individuals living in northern Poland as reference material for identification studies of biological stains. 158 samples originating from unrelated individuals of both sexes were subjected to analysis. DNA extraction was carried out according to a method described earlier [5]. The amount of extracted DNA was determined fluorometrically or by a method based on DNA hybridisation with Primate specific probe (QuantiBlot, Applera, USA) followed by chemiluminescent detection.

Amplification of the alleles of locus D8S1132 was carried out on ABI877 PCR robotic station (Applera, USA) using primer sequences (primer D8 I: 5' – GGCTAGGAAAGGTTAGTCGC – 3'; primer D8 II: 5' – CCCTCTCTTTTCGAGCAAT – 3') and PCR conditions described by Wiegand et al. [11]. PCR products were resolved and detected on horizontal polyacrylamide gels according to previously described methods [6]. The allelic ladder used in order to assess lengths of observed alleles was kindly supplied by Dr W. Pepiński from the Department of Legal Medicine of the Medical Academy in Białystok. In some cases, obtained D8S1132 alleles were verified using the capillary electrophoresis method on automatic sequencer ABI 310 (Applera, USA), with one of the primers (primer D8 I) being fluorescently labelled on the 5' end by TET. Sample preparation and parameters of capillary electrophoresis were performed according to conditions described previously [3, 4, 5]. As an internal length standard we used GS500 marked with ROX (PE) or ILS600 marked with CXR (Promega).

Concordance of the data with the Hardy-Weinberg equation was checked by the "exact" and LR tests [2, 9]. Values for observed heterozygosity ( $H$ ), power of discrimination ( $PD$ ), mean exclusion chance ( $MEC$ ) and polymorphism information content ( $PIC$ ) were calculated as described previously [3, 5, 6]. Homogeneity of the allelic distribution for locus D8S1132 between studied population sample and other populations was tested with Carmody's computer program.

## RESULTS AND DISCUSSION

Completely satisfactory amplification products were obtained for all 158 analysed DNA samples. Furthermore, the observation made by Wie-

gand et al. [11] concerning high sensitivity of amplification of locus D8S1132 was confirmed by our studies. The short length of the PCR fragments may be taken as one of the reason for the very effective amplification of the analysed locus. Phenotype and allele frequencies observed in the population sample of the 158 individuals (males and females) collected from northern Poland are presented in Tables I and II.

TABLE I. OBSERVED PHENOTYPE FREQUENCIES OF LOCUS D8S1132 IN A POPULATION SAMPLE OF 158 INDIVIDUALS FROM NORTH POLAND

Pheno-type	Observed number	[%]	Pheno-type	Observed number	[%]	Pheno-type	Observed number	[%]
15,23	1	0.633	18,22	11	6.962	20,22	6	3.797
17,18	8	5.063	18,23	7	4.43	20,23	4	2.532
17,19	4	2.532	18,24	3	1.899	20,24	2	1.266
17,20	3	1.899	18,25	1	0.633	21,21	3	1.899
17,21	7	4.43	19,19	3	1.899	21,22	5	3.165
17,22	4	2.532	19,20	3	1.899	21,23	5	3.165
17,23	5	3.165	19,21	3	1.899	21,24	1	0.633
17,24	2	1.266	19,22	7	4.43	22,22	2	1.266
18,18	5	3.165	19,23	7	4.43	22,23	5	3.165
18,19	8	5.063	19,24	4	2.532	22,24	1	0.633
18,20	7	4.43	20,20	2	1.266	23,23	2	1.266
18,21	5	3.165	20,21	11	6.962	23,24	1	0.633

TABLE II. ALLELES FREQUENCY DISTRIBUTION FOR LOCUS D8S1132 IN A NORTH POLISH POPULATION

Allel	15	17	18	19	20	21	22	23	24	25
Number	1	33	60	42	40	43	43	39	14	1
Frequency	0.0032	0.1044	0.1898	0.1329	0.1266	0.1361	0.1361	0.1234	0.0443	0.0032

In the analysed population sample we determined the presence of 10 alleles, 9 of which had already been observed in the Polish population [8] and in three different German population samples [11], and one of which had not been observed till now. The new allele, migrating in gel below the lowest allele present in the applied allelic ladder (D8S1132\*16) was analysed on automatic DNA sequencer ABI 310 in order to be more accurately characterised. It was established that this allele is 4 bp shorter than allele 16, which supports the fact that this population study revealed the presence of a new allele, allele 15, in locus D8S1132. Full confirmation of this discovery, however, requires sequencing of the new allele. This is presently being undertaken in our laboratory and will be the subject of a separate paper.

The distribution of the allele frequencies of locus D8S1132 in the population of northern Poland, similarly to four other analysed populations, is unimodal with 7 frequent and 3 rare alleles (Table I).

The most frequently observed allele in the analysed population sample was allele 18 ( $f = 0.1899$ ) and the rarest were alleles 25 and 15 with frequency  $f = 0.0032$ . Allele 18 was also most frequently observed in the population of north-eastern Poland [8] and the studied German population samples [11]. In the analysed population sample from the region of northern Poland the presence of allele 16 was not ascertained. This allele appeared in the population of north-eastern Poland with a frequency = 0.004 [8] and in the population samples from Germany with frequencies from 0.004 to 0.015 [11].

Among 158 analysed DNA samples the presence of 36 phenotypes out of 55 theoretically possible for the observed number of alleles, was found. The most frequent ones are 18,22 and 20,21 observed with a frequency of around 7% (Table II). Locus D8S1132 in the analysed population is characterised by a very high value of observed heterozygosity ( $H_{obs} = 0.89$ ), which is very similar to the calculated value for expected heterozygosity ( $0.87 \pm 0.01$ ). Further statistical characterisation of the locus shows its high polymorphism ( $PIC = 0.8498$ ), significant usefulness in the analysis of disputed paternity (mean exclusion chance  $MEC = 0.7255$ ) and potential usefulness in the process of identification of biological traces (power of discrimination  $PD = 0.9622$ ).  $PD$  Coefficient calculated for D8S1132 has one of the highest values among popular STR loci. Its value is higher than that of all loci present in the Profiler Plus [5] kit, and, comparing to other frequently used STR systems, only locus ACTBP2 has a higher value of  $PD$  coefficient (0.995) [7]. Taking into account the presence of significantly shorter alleles in the case of locus D8S1132 (130–170 bp) than in the case of locus ACTBP2 and the very efficient amplification of the system, one can consider locus D8S1132 more useful for the identification process of biological traces than the ACTBP2 locus.

Application of appropriate statistical tests [2, 9] confirmed concordance with the Hardy-Weinberg equation. Coefficient  $P$  calculated by “exact” and G-tests gave very high values: 0.94 and 0.93 respectively. Table III presents a comparison of the homogeneity distribution of alleles observed in locus D8S1132 between the population of northern Poland and the population of north-eastern Poland [8] and three German population samples [11]. Comparative analyses did not reveal statistically significant differences ( $P > 0.05$ ). The least significant concordance, however, was obtained when comparing the studied population with the population sample from Wiesbaden and the most significant when comparing with the population sample from Münster [11].

TABLE III PAIRWISE COMPARISON OF ALLELE FREQUENCIES DISTRIBUTION BETWEEN NORTH POLISH POPULATION AND OTHER CAUCASIAN POPULATION SAMPLES

Compared populations	$\chi^2$	$P$	G-statistic	$P$
NPL $\times$ NEPL [8]	8.65	0.590	8.018	0.581
NPL $\times$ Halle [11]	12.74	0.236	13.86	0.239
NPL $\times$ Münster [11]	7.86	0.684	8.81	0.672
NPL $\times$ Wiesbaden [11]	16.09	0.078	18.02	0.069

NPL – north Poland; NEPL – north-east Poland.

The results of this work, and also observations made by other researchers during analysis of biological traces [10] strongly confirm the extreme usefulness of locus D8S1132 in forensic genetics. This locus has been applied to forensic casework performed in the Department of Legal Medicine of the Medical Academy in Gdańsk. The authors of this paper are carrying out work concerning optimisation of multiplex amplification of locus D8S1132 together with other STR loci.

#### References:

1. Brinkmann B., The STR approach, *Advances of Forensic Haemogenetics* 1996, no. 6, pp. 41–51.
2. Guo S. W., Thompson E. A., Performing the exact test of Hardy-Weinberg proportion for multiple alleles, *Biometrics* 1992, vol. 48, pp. 361–372.
3. Pawłowski R., HUMFIBRA allele distribution in Northern Poland using capillary electrophoresis, *International Journal of Legal Medicine* 1999, vol. 112, pp. 139–141.
4. Pawłowski R., Branicki W., Kupiec T., Y-chromosomal polymorphic loci DYS19, DYS390, DYS393 in a population sample from northern Poland, *Electrophoresis* 1999, vol. 20, pp. 1702–1706.
5. Pawłowski R., Maciejewska A., The forensic validation of multiplex containing nine STRs – population genetics in Northern Poland, *International Journal of Legal Medicine* 2000, vol. 114, pp. 45–49.
6. Pawłowski R., Maciejewska A., Paszkowska R. [et al.], Frequencies of the five short tandem repeat systems (STR) in a population from north Poland, *International Journal of Legal Medicine* 1997, vol. 110, pp. 10–13.
7. Pawłowski R., Maciejewska A., Paszkowska R. [et al.], The current status of forensic DNA analysis in Poland, Proceedings from the First European Symposium on Human Identification, Promega Corporation 1997, pp. 116–125.
8. Pepiński W., Niemcunowicz-Janica A., Janica J. [et al.], Population data for the STR systems D8S1132, CD4, VWA and TH01 in the region of Podlasie (north-eastern Poland), *Medical Science Monitor* 2001, vol. 7, pp. 130–133.
9. Weir B. S., Independence of VNTR alleles defined by fixed bins, *Genetics* 1992, vol. 130, pp. 873–878.

10. Wiegand P., Kleiber M., DNA typing of epithelial cells after strangulation, *International Journal of Legal Medicine* 1997, vol. 110, pp. 181–183.
11. Wiegand P., Schneider H. R., Schurenkamp M. [et al.], Tetranucleotide STR system D8S1132: sequencing data and population genetic comparisons, *International Journal of Legal Medicine* 1998, vol. 111, pp. 180–182.

## CZĘSTOŚCI ALLELI WYSOCE POLIMORFICZNEGO LOKUS D8S1132 W PRÓBCE POPULACYJNEJ Z OBSZARU POLSKI PÓŁNOCNEJ

Ryszard PAWŁOWSKI, Marta ŻYWICKA

### WSTĘP

Wśród bardzo wielu opisanych dotychczas *loci* typu VNTR-STR do szczególnie przydatnych z punktu widzenia genetyki sądowej należą te, które charakteryzują się wysoką siłą dyskryminacji [1] oraz obecnością krótkich alleli. Te obie cechy czynią je szczególnie przydatnymi do profilowania DNA powielanego ze zdegradowanego materiału biologicznego. Ostatnio Wiegand i in. [11] szczegółowo opisali polimorfizm oraz zsekwencjonowali allele lokus D8S1132 (8pter-8qpter). Lokus ten charakteryzuje się złożoną sekwencją repetytywną typu  $(TCTA)_n TCA (TCTA)_n$ . Do tej pory opisano 10 alleli (od 16 do 25) w zakresie wielkości od 134–170 pz. Badany system należy do bardzo czułych, dając możliwość uzyskania sygnału amplifikacyjnego z 20–50 pg matrycowego DNA [10].

Głównym celem niniejszej pracy było zbadanie rozkładu częstości alleli dla lokus D8S1132 w próbie populacyjnej pochodzącej od osobników zamieszkujących Polskę północną oraz porównanie ich rozkładu z opublikowanymi danymi pochodzącymi z populacji niemieckiej oraz z Polski północno-wschodniej.

### MATERIAŁY I METODY

Materiał do badań populacyjnych stanowiły próbki krwi lub wymazy z jamy ustnej pobierane jako materiał porównawczy do badań identyfikacyjnych śladów biologicznych pochodzące od osób zamieszkujących Polskę północną. Badaniom poddano 158 próbek pochodzących od niespokrewnionych osobników obu płci. Ekstrakcję DNA przeprowadzano metodą opisaną wcześniej [5]. Ilość DNA oznaczano metodą hybrydyzacji z sondą swoistą dla DNA naczelnych (QuantiBlot, Applera, Stany Zjednoczone) z chemiluminescencyjną metodą detekcji lub pomiarem fluorymetrycznym.

Amplifikację alleli lokus D8S1132 z zastosowaniem odpowiednich starterów (starter D8 I: 5' – GGCTAGGAAAGGTTAGTCGC – 3'; starter D8 II: 5' – CCCTCTCTCTTCGAGCAAT – 3') prowadzono przy pomocy automatycznego urządzenia do amplifikacji PCR – ABI877 firmy Applera (Stany Zjednoczone), stosując warunki podane przez Wieganda i in. [11]. Rozdziału i detekcji produktów PCR dokonano na poziomych żelach poliakrylamidowych metodami jak opisano uprzednio [6]. Jako standardu wielkości alleli użyto drabiny alleli udostępnionej do reamplifikacji dzięki uprzejmości Pana dra W. Pepińskiego z Katedry i Zakładu Medycyny Sądowej AM w Białymstoku. W niektórych przypadkach fenotypy D8S1132 weryfikowano metodą elektroforezy kapilarnej na automatycznym sekwenatorze DNA ABI310 firmy Applera (Stany Zjednoczone), stosując podczas reakcji PCR jeden ze starterów (starter D8 I) znakowany na końcu 5' znacznikiem fluorescencyjnym

TET. Przygotowanie próbek do tej analizy i sam ich rozdział prowadzono w warunkach jak opisano uprzednio [3, 4, 5]. Jako standardu wewnętrznego wielkości DNA używano markera GS500 znakowanego ROX (PE) lub ILS600 znakowanego CXR (Promega).

Weryfikację zgodności z równaniem Hardy'ego-Weinberga prowadzono testami „exact” i LR [2, 9]. Obserwowaną heterozygotyczność ( $H$ ), siłę dyskryminacji ( $PD$ ), średnią szansę wykluczenia ( $MEC$ ) i zawartość informacji polimorficznej ( $PIC$ ) obliczano jak uprzednio [3, 5, 6]. Homogenność rozkładu alleli D8S1132 pomiędzy badaną próbką populacyjną a innymi populacjami testowano programem Carmody'ego.

#### WYNIKI I DYSKUSJA

W przypadku wszystkich 158 badanych próbek DNA uzyskano w pełni zadawalającą amplifikację, a jednocześnie potwierdzono obserwację Wieganda i in. [11], iż locus D8S1132 charakteryzuje się wysoką czułością powielania. Jednym z powodów bardzo wydajnej amplifikacji może być niewielka długość amplifikowanych fragmentów. Tabele I i II przedstawiają częstości fenotypów i alleli zaobserwowane w próbce populacyjnej pochodzącej od 158 osób obu płci z obszaru Polski północnej.

W badanej próbce populacyjnej stwierdzono obecność 10 alleli, z których 9 obserwowano już w populacji polskiej [8] i trzech niemieckich [11], a jednego do tej pory w ogóle nie obserwowano. Allel ten, wędrujący poniżej najniższego allela w stosowanej drabinie (D8S1132\*16), został poddany analizie na sekwenatorze DNA ABI310 w celu dokładniejszego scharakteryzowania. Ustalono, iż jest on krótszy o 4 pz od allela 16, co przemawia za faktem, iż podczas badań populacyjnych stwierdzono nie obserwowany dotychczas allel 15 w locus D8S1132. Pełna weryfikacja tego faktu wymaga jednak przeprowadzenia sekwencjonowania i tym sposobem ostatecznego potwierdzenia. Stosowne badania są wykonywane i staną się przedmiotem osobnej publikacji.

Rozkład częstości alleli D8S1132 w populacji Polski północnej, podobnie jak w innych czterech do tej pory badanych populacjach, jest jednomodalny z 7 powszechnymi i 3 rzadkimi allelami (tabela I).

Najczęstszym obserwowanym allelem w badanej próbce był allel 18 ( $f = 0,1899$ ), a najrzadszym 25, który, podobnie jak allel 15, wystąpił z częstością  $f = 0,0032$ . Allel 18 jest też najczęściej występującym allelem zarówno w populacji Polski północno-wschodniej [8], jak i populacjach niemieckich [11]. W badanej próbce populacyjnej z obszaru Polski północnej nie obserwowano obecności allela 16, który pojawił się w populacji Polski północno-wschodniej z częstością  $f = 0,004$  [8], a w populacjach niemieckich z częstościami  $f$  od 0,004 do 0,015 [11].

Wśród 158 badanych próbek DNA stwierdzono istnienie 36 fenotypów z 55 teoretycznie możliwych dla tej liczby alleli. Najczęstsze z nich to 18,22 i 20,21 występujące z częstością około 7% (tabela II). Locus D8S1132 w badanej populacji charakteryzuje wysoka wartość obserwowanej heterozygotyczności ( $H_{obs} = 0,89$ ), która zbliżona jest do obliczonej wartości heterozygotyczności oczekiwanej ( $0,87 \pm 0,01$ ). Dalsza charakterystyka statystyczna locus pokazuje jego wysoki polimorfizm (współczynnik informacji o polimorfizmie  $PIC = 0,8498$ ), wysoką przydatność do dochodzenia spornego ojcostwa (średnia szansa wykluczenia ojcostwa  $MEC = 0,7255$ ), jak i potencjal-

nie do identyfikacji śladów biologicznych (moc dyskryminacyjna  $PD = 0,9622$ ). Współczynnik  $PD$  jest jednym z najwyższych wśród powszechnie stosowanych *loci* typu STR. Jego wartość jest wyższa np. od wszystkich *loci* obecnych w zestawie Profiler Plus [5], a z innych, powszechnie stosowanych, ustępuje jedynie lokus ACTBP2, który charakteryzuje współczynnik  $PD$  bliski 1 (0,995) [7]. Zważywszy jednak na fakt obecności znacznie krótszych alleli (130–170 pz) niż w przypadku ACTBP2 i bardzo wydajnej amplifikacji, system ten może okazać się znacznie bardziej przydatny do identyfikacji śladów biologicznych niż wspomniany lokus ACTBP2.

W wyniku zastosowania odpowiednich testów [2, 9] stwierdzono, że badany lokus znajduje się w równowadze opisanej równaniem Hardy'ego-Weinberga. Współczynnik  $P$  uzyskany w testach „exact” i  $G$  dał bardzo wysokie wartości, odpowiednio 0,94 i 0,93.

Tabela III przedstawia porównanie homogenności rozkładu alleli lokus D8S1132 pomiędzy populacją Polski północnej a populacją Polski północno-wschodniej [8] oraz trzema populacjami niemieckimi [11]. Dla wszystkich przeprowadzonych porównań nie zaobserwowano różnic statystycznych ( $P > 0,05$ ). Najmniejszą zgodność otrzymano, porównując badaną populację z populacją Wiesbaden, a najwyższą zgodność z populacją z Münster [11].

Wszystkie opisane w niniejszej pracy badania i analizy oraz obserwacje poczynione przez innych przy badaniu śladów biologicznych [10] w pełni potwierdzają bardzo wysoką przydatność lokus D8S1132 do badań z zakresu genetyki sądowej. Lokus ten został wprowadzony do badań z zakresu genetyki sądowej w Zakładzie Medycyny Sądowej AM w Gdańsku. Autorzy prowadzą również prace zmierzające do jego amplifikacji z innymi *loci* w reakcji kompleksowego PCR.