

POLIMORFISM OF THE VNTR LOCI: D7S22, D5S433 AND D16S309 IN THE POMERANIA-KUJAWY REGION OF POLAND

Danuta MIŚCICKA-ŚLIWKA, Jakub CZARNY

Chair and Department of Forensic Medicine, Ludwik Rydygier Academy of Medicine, Bydgoszcz

ABSTRACT: The paper presents the results of a population study of 3 VNTR loci: D7S22, D5S43 and D16S309. The allele distributions were in accordance with Hardy-Weinberg expectations. Values of heterozygosity, polymorphic information contents, power of discrimination, mean exclusion chance and typical paternity index demonstrate that these systems are valuable tools for paternity testing.

KEY WORDS: VNTR; D7S22; D5S43; D16S309; Population genetics; Pomerania-Kujawy.

*Z Zagadnień Nauk Sądowych, z. XLVIII, 2001, 65–73
Received 19 November 2001; accepted 14 December 2001*

INTRODUCTION

Among the 3 billions base pairs constituting the human genome are sequences revealing high interindividual variability. Some of them are so-called repetitive sequences, whose variability derives from the different number of repetitions of a repetitive element. An example of repetitive sequences is minisatellite sequences, known also as VNTR's (variable number of tandem repeat) characterised by a repetitive element of length from 9 to 80 base pairs and total length from hundreds to over 20 thousand base pairs [12]. Analysis of VNTR loci polymorphism has been applied to disputed paternity testing [6], relationship analysis performed for immigration purposes [5], individual identification [4] and linkage mapping of the human genome [9].

Among VNTR type sequences at present widely used, mostly in disputed paternity testing, are loci: D7S22, D5S43 [13] and D16S309 [10]. Locus D7S22 has been mapped in the terminal section of the long arm of human chromosome 7 (7q36-qter); the repetitive element consists of 37 base pairs. This sequence can be detected by hybridisation with the G3 probe (Cellmark Diagnostics). Locus D5S43 is situated in the terminal section of the long arm of human chromosome 5 (5q35-qter). In this locus two types of repetitive ele-

ments were ascertained: the first of length 30 base pairs and the second of length 29 base pairs, which can be detected using hybridisation probe MS 8 (Cellmark Diagnostics). Locus D16S309 has been mapped in the short arm of human chromosome 16 (p16). For this locus a few different sorts of repetitive sequences were identified of length from 45 to 54 base pairs, which can be detected by hybridisation with the MS205 probe (Cellmark Diagnostics).

Before any new marker is routinely applied to individual identification studies and disputed paternity testing, accurate assessment of its polymorphism in a given population is necessary. In this work, the nature of the polymorphism of VNTR loci D7S22, D5S43 and D16S309 in the Pomerania-Kujawy population is presented.

MATERIALS AND METHODS

The studied specimens consisted of peripheral blood taken from the ulnar vein of 400–429 (depending on the analysed locus) healthy, unrelated adult individuals, both male and female, living in the Pomerania-Kujawy region.

DNA was extracted from 5 ml of peripheral blood using a technique known as “the organic method” [11]. The concentration and purity of the extracted DNA was assessed with the spectrophotometric method, using the GeneQuant DNA/RNA Calculator (Pharmacia-Biotech). 5 µg of DNA was subjected to digestion by endonuclease Hinf I (Promega Co.) according to the manufacturer’s instructions. The digestion products were separated in 0.9% agarose gel (Promega) with separation length 25 cm in buffer 1 × TBE for 42 hours, applied voltage 1 V/cm and constant buffer circulation. After separation DNA was subjected to fragmentation with 0.25 M HCl followed by denaturation (0.5 M NaOH; 1.5 M NaCl) and neutralisation (0.5 M Tris-HCl, 1.5 M NaCl, pH = 7.4). DNA was transferred to a neutral nylon membrane (Biodyne A, GibcoBRL) using the capillary method with 10 × SSC buffer for 16 hours and then bound to the membrane by UV radiation for 3 minutes. Hybridisation was performed using molecular probes: Nice Oligo DNA Probe G3 (Cellmark Diagnostics) for locus D7S22 analysis, Nice Oligo DNA Probe MS8 (Cellmark Diagnostics) for locus D5S43 analysis and Nice Oligo DNA Probe MS205 (Cellmark Diagnostics) for locus D16S309 analysis. Hybridisation was performed according to the manufacturer’s directions. Chemiluminescent detection was carried out according to the manufacturer’s instructions using LumiPhos 480 as a substrate, and reaction products were made visible on an X-ray film (Foton, Warsaw). Restriction fragment sizes were determined according to DNA length marker Nice DNA Analysis Ladder (GibcoBRL), made visible by hybridisation with the Nice Oligo DNA Probe MW 100 (Cellmark Diagnostics), applying a technique of

computer image analysis using a video scanner with BKA Fingerprint v. 2.1 software (Biotec-Fisher GmbH).

The fragments were grouped according to their sizes, using the “fixed-bin” method [3], linking them in groups (bins) corresponding to intervals of a marker of DNA size (Nice DNA Analysis Ladder). The obtained results enabled assessment of the following parameters for the analysed loci: expected heterozygosity (H_e) and observed heterozygosity (H_o) using the computer program Genetic Data Analysis [8] and standard error (SE). Polymorphic information content (PIC) [1], power of discrimination (PD) [7], mean exclusion chance (MEC) and typical paternity index [2] were calculated using the computer program PowerStats (available on the Internet: www.promega.com). The concordance of the observed genotype frequencies in particular loci with Hardy-Weinberg expectations were tested by calculation of the probability of concordance on the basis of the Fisher exact test (“exact p”) using computer program Genetic Data Analysis.

RESULTS AND DISCUSSION

Frequencies and number of observed restriction fragments, in the range of adopted size intervals, for loci D7S22, D5S43 and D16S309 and also parameters: expected heterozygosity (H_e), observed heterozygosity (H_o), standard error (SE), polymorphic information content (PIC), power of discrimination (PD), mean exclusion chance (MEC), typical paternity index (TPI) and probability of concordance of the observed genotypes frequencies distribution with the distribution of genotypes frequencies calculated on the basis of allelic frequencies according to the Hardy-Weinberg equilibrium on the basis of the Fisher exact test (“exact p”) are presented in Table I.

For the analysed loci observed genotypes frequencies distributions are in good concordance ($p > 0.05$) with distributions assessed on the basis of allelic frequencies according to the Hardy-Weinberg equilibrium. For locus D5S43, the value of “exact p” is close to the limit value (0.052188). This may be caused by genotypes identified for locus D5S43, whose observed frequency is much higher than the expected frequency. Such a situation mostly concerns very rare genotypes. In the case of locus D5S43 higher values of “exact p” can be obtained by so-called “rebinning”, which means linking bins in such a way that the number of alleles observed for them exceeds 5. In such a situation “exact p” equals 0.350313.

Locus D7S22 is characterised by the highest heterozygosity and polymorphic information content among the analysed loci. Moreover, indexes assessing usefulness of the locus in individual identification (power of discrimination) and disputed paternity testing (mean exclusion chance and typical

TABLE I. POLYMORPHISM OF THE VNTR LOCI D7S22, D5S43 AND D16S309 IN THE POPULATION OF THE POMERANIA-KUJAWY

Length of DNA fragment (bin)	D7S22 N = 400		D5S43 N = 421		D16S309 N = 429	
	Number	Frequency	Number	Frequency	Number	Frequency
1–653	–	–	–	–	1	0.001
654–784	–	–	–	–	2	0.002
785–910	–	–	–	–	2	0.002
911–993	–	–	–	–	1	0.001
142994–1176	–	–	–	–	8	0.009
1177–1287	–	–	–	–	–	–
1288–1431	1	0.001	–	–	3	0.003
1432–1568	–	–	–	–	19	0.022
1569–1672	78	0.098	1	0.001	12	0.014
1673–1861	20	0.025	2	0.002	6	0.007
1862–2015	1	0.001	1	0.001	34	0.04
2016–2213	7	0.009	1	0.001	99	0.115
2214–2433	3	0.004	76	0.090	97	0.113
2434–2650	–	–	–	–	107	0.125
2651–2876	26	0.033	39	0.046	134	0.158
2877–3101	47	0.059	1	0.001	106	0.125
3102–3397	50	0.063	2	0.002	101	0.118
3398–3812	21	0.026	2	0.002	73	0.085
3813–4333	17	0.021	1	0.001	45	0.052
4334–4716	14	0.018	63	0.075	3	0.003
4717–5415	29	0.036	232	0.276	3	0.003
5416–5861	47	0.059	15	0.018	1	0.001
5862–6442	76	0.095	52	0.062	1	0.001
6443–7421	114	0.141	324	0.386	–	–
7422–8271	93	0.116	3	0.004	–	–
8272–9416	56	0.070	27	0.032	–	–
9417–11919	81	0.101	–	–	–	–
11920–15004	18	0.023	–	–	–	–
15005–22621	1	0.001	–	–	–	–
> 22621	–	–	–	–	–	–
Σ	800	1	842	1	858	1
<i>He</i>	0.917643		0.755822		0.893758	
<i>Ho</i>	0.921348		0.764846		0.898017	
<i>SE</i>	0.00262		0.006381		0.003012	
<i>PIC</i>	0.91		0.72		0.88	
<i>PD</i>	0.985		0.900		0.976	
<i>MEC</i>	0.852		0.536		0.781	
<i>TPI</i>	6.90		2.13		4.66	
“exact p”	0.837500		0.052188		0.386875	

N – number of analysed genotypes, *He* – expected heterozygosity, *Ho* – observed heterozygosity, *SE* – standard error, *PIC* – polymorphic information content, *PD* – power of discrimination, *MEC* – mean exclusion chance, *TPI* – typical paternity index, “exact p” – concordance probability of the observed genotypes frequencies distribution with the Hardy-Weinberg rule.

paternity index) are highest for locus D7S22. Corresponding values for locus D16S309 are a bit lower. Locus D5S43, among analysed loci, is characterised by the lowest values for all the calculated parameters.

The total usefulness of the analysed loci for individual identification is expressed by power of discrimination equal to 0.999964. The total mean exclusion chance calculated for the analysed loci equals 0.984961 and the typical paternity index 68.49. The obtained results indicate that analysis of D7S22, D5S43 and D16S309 loci polymorphism can be a useful tool, especially in disputed paternity cases.

References:

1. Botstein D., White R. L., Skolnick M. [et al.], Construction of a genetic linkage map in man using restriction fragment length polymorphisms, *American Journal of Human Genetics* 1980, vol. 32, pp. 314–331.
2. Brenner C., Morris J., Paternity index calculations in single locus hyper-variable DNA probes: validation and other studies, [in:] Proceedings for the International Symposium on Human Identification 1989, Promega Corporation, Madison 1990, pp. 21–53.
3. Budowle B., Giusti A. M., Wayne J. S. [et al.], Fixed-bin analysis for statistical evaluation of continuous distributions of allelic data from VNTR loci, for use in forensic comparison, *American Journal of Human Genetics* 1991, vol. 48, pp. 841–855.
4. Gill P., Jeffreys A. J., Werrett D. J., Forensic application of DNA “fingerprints”, *Nature* 1985, vol. 318, pp. 577–579.
5. Jeffreys A. J., Brookfield J. F. Y., Semeonoff R., Positive identification of an immigration test-case using human DNA fingerprints, *Nature* 1985, vol. 317, pp. 818–819.
6. Jeffreys A. J., Wilson V., Thein S. L., Individual-specific “fingerprints” of human DNA, *Nature* 1985, vol. 316, pp. 76–79.
7. Jones D. A., Blood samples: Probability of discrimination, *Journal of Forensic Sciences Society* 1972, vol. 12, pp. 355–359.
8. Lewis P. O., Zaykin D., Genetic Data Analysis: Computer program for the analysis of allelic data. Version 1.0. Free program distributed by the authors over the Internet from the GDA home page at <http://chee.unm.edu/gda/>.
9. Nakamura Y., Leppert M., O’Connell P. [et al.], Variable Number of Tandem Repeat (VNTR) markers for human gene mapping, *Science* 1987, vol. 235, pp. 1616–1622.
10. Royle N. J., Armour J. A., Webb M. [et al.], A hypervariable locus D16S309 located at the distal end of 16p, *Nucleic Acid Research* 1992, vol. 20, p. 1164.
11. Sambrook J., Fritsch E. F., Maniatis T., Isolation of high-molecular-weight DNA from mammalian cells, [in:] Analysis and cloning of eucaryotic genomic DNA. Molecular cloning a laboratory manual, Cold Spring Harbor Laboratory Press, New York 1989, pp. 16–22.

12. Weir B. S., Independence of VNTR alleles defined as fixed bins, *Genetics* 1992, vol. 130, pp. 873–887.
13. Wong Z., Wilson V., Patel I. [et al.], Characterization of a panel of highly variable minisatellites cloned from human DNA, *Annals of Human Genetics* 1987, vol. 51, pp. 269–288.

POLIMORFIZM *LOCI* VNTR: D7S22, D5S43 I D16S309 W POPULACJI KUJAWSKO-POMORSKIEJ

Danuta MIŚCICKA-ŚLIWKA, Jakub CZARNY

WSTĘP

Wśród 3 miliardów par zasad tworzących genom człowieka znajdują się sekwencje o dużej zmienności międzypersonicznej. Część z nich tworzy tzw. sekwencje powtarzające się, których zmienność wynika z różnej liczby powtórzeń jednostki repetytywnej. Przykładem sekwencji powtarzających się są sekwencje minisatelitarne, zwane także VNTR (*variable number of tandem repeat*) charakteryzujące się jednostką repetytywną wielkości od 9 do 80 par zasad i całkowitą długością od kilkuset do ponad 20 tysięcy par zasad [12]. Analiza polimorfizmu *loci* VNTR znalazła zastosowanie w badaniach spornego ojcostwa [6], analizie pokrewieństwa wykonywanej dla celów imigracyjnych [5], identyfikacji osobniczej [4] oraz mapowaniu sprzężeń genomu człowieka [9].

Do grona sekwencji typu VNTR, wykorzystywanych obecnie powszechnie przede wszystkim w dochodzeniu spornego ojcostwa, należą między innymi *loci*: D7S22, D5S43 [13] i D16S309 [10]. Lokus D7S22 mapowane jest w terminalnym odcinku długiego ramienia chromosomu 7 człowieka (7q36-qter); powtarzający się motyw zbudowany jest z 37 par zasad. Sekwencję tę można wykrywać na drodze hybrydyzacji z sondą G3 (Cellmark Diagnostics). Lokus D5S43 leży w terminalnym odcinku długiego ramienia chromosomu 5 człowieka (5q35-qter). W lokus tym stwierdzono dwa typy jednostek powtarzających się: jedną o długości 30 par zasad i jedną o długości 29 par zasad, które można wykrywać dzięki hybrydyzacji z sondą MS 8 (Cellmark Diagnostics). Lokus D16S309 mapowane jest w krótkim ramieniu chromosomu 16 człowieka (p16). Stwierdzono w nim występowanie kilku rodzajów sekwencji powtarzających się długości od 45 do 54 par zasad, które można wykrywać dzięki hybrydyzacji z sondą MS205 (Cellmark Diagnostics).

Przed wprowadzeniem markera do rutynowych badań wykonywanych dla potrzeb identyfikacji osobniczej i dochodzenia spornego ojcostwa, konieczna jest dokładna charakterystyka jego polimorfizmu w danej populacji. W niniejszej pracy przedstawiono charakterystykę polimorfizmu *loci* VNTR D7S22, D5S43 i D16S309 w populacji kujawsko-pomorskiej.

MATERIAŁ I METODA

Materiał do badań stanowiła krew obwodowa pobrana z żyły łokciowej od 400–429 (w zależności od lokus) zdrowych, niespokrewnionych dorosłych osób, mężczyzn i kobiet z regionu kujawsko-pomorskiego.

DNA izolowano z 5 ml krwi obwodowej tzw. „metodą organiczną” [11]. Stężenie i czystość DNA oceniano metodą spektrofotometryczną wykorzystując GeneQuant DNA/RNA Calculator (Pharmacia-Biotech). 5 µg DNA poddawano trawieniu endo-

nukleazą Hinf I (Promega Co.) według zaleceń producenta. Produkty trawienia rozdzielano w 0,9% żelu agarozowym (Promega Co.) o drodze rozdziału długości 25 cm w buforze $1 \times$ TBE przez 42 godziny przy napięciu 1 V/cm z recyrkulacją buforu. Po zakończeniu rozdziału DNA poddawano fragmentacji w 0,25 M HCl, następnie denaturowano (0,5 M NaOH; 1,5 M NaCl) i neutralizowano (0,5 M Tris-HCl, 1,5 M NaCl, pH = 7,4). DNA przenoszono na nylonową membranę neutralną (Bio-dyne A, GibcoBRL) metodą kapilarną w buforze $10 \times$ SSC przez 16 godzin, a następnie wiązano z membraną poprzez naświetlanie promieniami UV przez 3 minuty. Hybrydyzację prowadzono wykorzystując sondy molekularne: Nice Oligo DNA Probe G3 (Cellmark Diagnostics) do analizy lokus D7S22, Nice Oligo DNA Probe MS8 (Cellmark Diagnostics) do analizy lokus D5S43 i Nice Oligo DNA Probe MS205 (Cellmark Diagnostics) do analizy lokus D16S309. Hybrydyzację prowadzono również zgodnie z zaleceniami producenta. Detekcję chemiluminescencyjną prowadzono zgodnie z zaleceniami producenta, wykorzystując jako substrat LumiPhos 480, a produkty reakcji uwidacziano na błonie rentgenowskiej (Foton Warszawa). Wielkości fragmentów restrykcyjnych określano względem wzorca długości DNA Nice DNA Analysis Ladder (GibcoBRL) uwidocznionego w drodze hybrydyzacji z sondą Nice Oligo DNA Probe MW 100 (Cellmark Diagnostics) przy użyciu techniki komputerowej analizy obrazu z zastosowaniem wideoskanera wraz z oprogramowaniem BKA Fingerprint v. 2.1 firmy Biotec-Fisher GmbH.

Fragmenty pogrupowano, stosując metodę *fixed-bin* [3], pod względem wielkości, łącząc je w grupy (biny) odpowiadające przedziałom markera wielkości DNA (Nice DNA Analysis Ladder). Na podstawie uzyskanych wyników wyznaczono dla analizowanych *loci*: heterozygotyczność oczekiwaną (*He*) i heterozygotyczność obserwowaną (*Ho*), wykorzystując program Genetic Data Analysis [8] oraz błąd standardowy (*SE*). Zawartość informacji polimorficznej (*PIC* – *polymorphic information content*) [1], siłę dyskryminacji (*PD* – *power of discrimination*) [7], teoretyczną szansę wykluczenia niesłusznie pozwanego mężczyzny (*MEC* – *mean exclusion chance*) i teoretyczną średnią szansę ojcostwa (*TPI* – *typical paternity index*) [2] wyznaczono wykorzystując program komputerowy PowerStats (program dostępny w sieci Internet: www.promega.com). Zgodność obserwowanych częstości genotypów w poszczególnych *loci* z prawem Hardy’ego-Weinberga testowano, obliczając prawdopodobieństwo zgodności na podstawie testu dokładnego Fischera („exact p”) w oparciu o program komputerowy Genetic Data Analysis.

WYNIKI I OMÓWIENIE

Częstości i liczbę obserwowanych fragmentów restrykcyjnych, w ramach przyjętych przedziałów wielkości, dla *loci* D7S22, D5S43 i D16S309 oraz parametry: heterozygotyczność oczekiwaną (*He*), heterozygotyczność obserwowaną (*Ho*), błąd standardowy (*SE*), zawartość informacji polimorficznej (*PIC*), siłę dyskryminacji (*PD*), teoretyczną szansę wykluczenia niesłusznie pozwanego mężczyzny (*MEC*) i teoretyczną średnią szansę ojcostwa (*TPI*) oraz prawdopodobieństwo zgodności obserwowanego rozkładu częstości genotypów z rozkładem częstości genotypów wyznaczonym w oparciu o częstości alleliczne według reguły Hardy’ego-Weinberga na podstawie testu dokładnego Fischera („exact p”) przedstawiono w tabeli I.

Obserwowane w analizowanych *loci* rozkłady częstości genotypów są zgodne ($p > 0.05$) z rozkładami wyznaczonymi w oparciu o częstości alleliczne według reguły Hardy'ego-Weinberga. Dla lokus D5S43 wartość „exact p” jest bliska wartości granicznej (0,052188). Przyczyną tej sytuacji może być stwierdzenie w lokus D5S43 genotypów, których częstość obserwowana jest znacznie wyższa od częstości oczekiwanej. Sytuacja taka dotyczy przede wszystkim bardzo rzadkich genotypów. W przypadku lokus D5S43 uzyskanie wyższych wartości „exact p” jest możliwe na drodze tzw. „rebinowania”, tzn. połączenia binów tak, aby liczba alleli w nich obserwowanych była większa od 5. W takiej sytuacji „exact p” wynosi 0,350313.

Najwyższą heterozygotycznością oraz zawartością informacji polimorficznej spośród badanych *loci* charakteryzuje się lokus D7S22. Również wskaźniki określające przydatność w identyfikacji osobniczej (siła dyskryminacji) i dochodzeniu spornego ojcostwa (teoretyczna szansa wykluczenia niesłusznie pozwanego mężczyzny oraz teoretyczna średnia szansa ojcostwa) są najwyższe w przypadku lokus D7S22. Odpowiednie wartości dla lokus D16S309 są nieznacznie niższe. Lokus D5S43, spośród badanych układów, charakteryzuje się najniższymi wartościami wszystkich wyznaczonych parametrów.

Łączna przydatność analizowanych *loci* dla potrzeb identyfikacji osobniczej wyraża się siłą dyskryminacji równą 0,999964. Łączna teoretyczna szansa wykluczenia niesłusznie pozwanego mężczyzny dla analizowanych *loci* wynosi 0,984961, a teoretyczna średnia szansa ojcostwa 68,49. Uzyskane wyniki wskazują, że analiza polimorfizmu *loci* D7S22, D5S43 i D16S309 może być cennym narzędziem przede wszystkim w dochodzeniu spornego ojcostwa.