

Krzysztof Rębała, Zofia Szczerkowska

Identyfikacja bardzo krótkiego allela YCAII w populacji północnej Polski

Identification of a very short YCAII allele in the northern Polish population

Z Katedry i Zakładu Medycyny Sądowej AM w Gdańsku
Kierownik: dr hab. Z. Szczerkowska - profesor AM

System YCAII należy do markerów STR chromosomu Y o dwunukleotydowej sekwencji repetytywnej, którego amplifikacja z uwagi na duplikację na chromosomie dostarcza dwóch polimorficznych produktów. Z uwagi na wysoki stopień polimorfizmu włączono go do Referencyjnej Bazy Danych Haplotypów Y-STR (<http://www.ystr.org>), co istotnie podniosło siłę dyskryminacji gromadzonych haplotypów. W trakcie badań populacyjnych nad tym systemem w populacji północnej Polski stwierdzono obecność bardzo krótkiego allela o stosunkowo niskim stosunku artefaktu powstającego na skutek poślizgu polimerazy DNA (ang. stutter) do właściwego allela. Sekwencjonowanie wykazało obecność 11 dwunukleotydowych jednostek repetytywnych charakterystycznych dla tego locus. Wśród 3998 rozszerzonych haplotypów zebranych w bazie danych allel ten stwierdzono dotychczas jedynie w populacji południowo-zachodnich Niemiec. Podobieństwa między haplotypami posiadającymi allel YCAIH1 mogą świadczyć o ich wspólnym pochodzeniu.

The YCAII system is one of the Y-chromosomal STR markers with dinucleotide repetitive units and due to the duplication on the chromosome, its amplification provides two polymorphic products. Due to the high polymorphism, YCAII was included in the Y-STR Haplotype Reference Database (<http://www.ystr.org>) and significantly increased discrimination power of the collected haplotypes. During studies on the Y-STR polymorphism in the Polish population, a very short YCAII allele with relatively low stutter was identified. Sequencing revealed the presence of 11 dinucleotide repetitive units characteristic of this system. Among 3998 extended haplotypes collected in the database, this allele was found only in the population of south-western Germany. Similarities between the haplotypes possessing the YCAIH1 allele suggest their common origin.

Słowa kluczowe: ludzki chromosom Y, sekwencje STR, YCAII, rzadki allel

Key words: human Y chromosome, short tandem repeats, YCAII, rare allele

WSTĘP

Badanie polimorfizmu loci STR ludzkiego chromosomu Y staje się rutynowo stosowaną analizą w genetyce sądowej. Szczególnie przydatne wydaje się w dochodzeniu spornego ojcostwa, kiedy domniemany ojciec jest nieosiągalny, zaś dostępny jest materiał biologiczny jego członków rodziny z linii męskiej, oraz w badaniu śladów biologicznych, kiedy podejrzany o przestępstwo jest mężczyzną (5). Dotychczas na ludzkim chromosomie Y zidentyfikowano kilkadziesiąt polimorficznych loci mikrosatelitarnych (9) i liczba ta stale rośnie.

Jednym z lepiej poznanych markerów genetycznych o wysokim polimorfizmie, sprzężonym z chromosomem Y, jest system YCAII. Zlokalizowany jest on na długim ramieniu tegoż chromosomu, zaś jego allele różnią się między sobą wielokrotnością dwunukletydowej sekwencji (CA)_n (11). Jednostki repetytywne tego systemu na skutek duplikacji wraz z sekwencjami flankującymi umożliwiają przy stosowaniu zaledwie jednej pary starterów jednoczesną amplifikację dwóch niezależnych loci (6). Jako że zakresy wielkości produktów amplifikacji obu sekwencji pokrywają się, nie jest możliwe rozróżnienie obu loci, dlatego genotypowanie systemu YCAII dostarcza dane bezpośrednio w postaci haplotypu dwóch alleli. Z uwagi na wysoki polimorfizm tego systemu włączono go do Referencyjnej Bazy Danych Haplotypów Y-STR (<http://www.ystr.org>), gdzie haplotypy zbadane pod względem ośmiu systemów Y-STR (DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393 i DYS385) określane są jako „minimalne haplotypy”, które wraz z systemem YCAII tworzą o wiele bardziej polimorficzne „rozszerzone haplotypy” (10).

W trakcie badań populacyjnych nad polimorfizmem loci Y-STR stwierdzono obecność bardzo krótkiego allela YCAII, który różnił się długością od kolejnego najkrótszego allela w populacji polskiej obejmującej 154 niespokrewnionych mężczyzn aż o 14 par zasad. Allel ten poddano sekwencjonowaniu i porównano jego występowanie w innych populacjach.

MATERIAŁ I METODY

Amplifikacja DNA

Reakcję PCR prowadzono w termocyklerze Mastercycler Gradient (Eppendorf) w całkowitej objętości 10 µl, zawierającej 1,75 mM MgCl₂, 300 µM dNTP, 5% (v/v) glicerol, 1 µl polimerazy AmpliTaq Gold (Applied Biosystems), 10 ng genomowego DNA mężczyzny o nieznanym allelu YCAII oraz startery o stężeniu równym 0,9 µM i sekwencjach zaczerpniętych z piśmiennictwa (3), przy czym koniec 5' pierwszego startera wyznakowano barwnikiem TET. Profil termiczny obejmował 10 min wstępnej denaturacji w 95 °C, 32 cykle amplifikacji (1 min w 94 °C, 1 min w 55 °C, 1 min w 72 °C) oraz 45 min końcowego wydłużania w 60 °C. Szybkość zmiany temperatury pomiędzy poszczególnymi etapami ustalono na 1 °C/s.

Elektroforeza kapilarna produktów amplifikacji

0,5 µl rozcieńczonego 10-krotnie w buforze TE (pH 8,0) amplifikatu zmieszano z 9,3 µl dejonizowanego formamidu i 0,15 µl wewnętrznego standardu wielkości GeneScan-500 TAMRA (Applied Biosystems), po czym przeprowadzono 5-minutową denaturację w 95 °C i szybko schłodzono w lodzie. Po 5-sekundowym nastrzyku przy 3 kV produkty reakcji rozdzielano przez 25 min w 60 °C przy 15 kV metodą elektroforezy kapilarnej na automatycznym sekwenatorze ABI Prism 310 (Applied Biosystems) przy zastosowaniu polimeru POP-4 (Applied Biosystems) i zestawu filtrów C. Wyniki analizowano programem komputerowym GeneScan 3.7 (Applied Biosystems) w stosunku do użytego wewnętrznego standardu wielkości.

Elektroforeza poliakrylamidowa

W celu rozdzielenia alleli w badanej próbce przeprowadzono elektroforezę w 7,2% natywnym żelu poliakrylamidowym o grubości 0,45 mm, usieciowanym w 3% 1,4-diakryloilopiperazyną (ang. piperazine diacrylamide). Na żel naniesiono 5 µl amplifikatu. Rozdzielone allele poddano barwieniu srebrem, po czym prążek odpowiadający krótszemu allelowi wycięto z żelu i umieszczono w 200 µl 5% zawiesiny żywicy Chelex. W celu elucji DNA z żelu całość inkubowano przez 1 h w 56 °C, wortexowano i wirowano przy 13000 rpm przez 3 min. Badany allel poddano ponownej amplifikacji. Do mieszaniny reakcyjnej pobrano 1 µl supernatantu, zaś całkowitą objętość zwiększono do 20 µl.

Przygotowanie produktu amplifikacji do sekwencjonowania

Produkt amplifikacji oczyszczono ze starterów i nukleotydów na kolumie Microcon YM-100 (Millipore). Do całości amplifikatu dodano 480 µl wody i poddano zagęszczaniu. Uzyskany koncentrat uzupełniono 450 µl wody i ponownie zagęszczono. Stężenie oczyszczonego produktu PCR wyznaczono z użyciem odczynnika PicoGreen (Molecular Probes) na automatycznym fluorymetrze Fluoroskan Ascent FL (Thermo Labsystems) w stosunku do krzywej wzorcowej, dodając po 1 µl próbki o badanym stężeniu do 99 µl 400-krotnie rozcieńczonego w wodzie odczynnika PicoGreen (12).

Reakcja sekwencjonowania

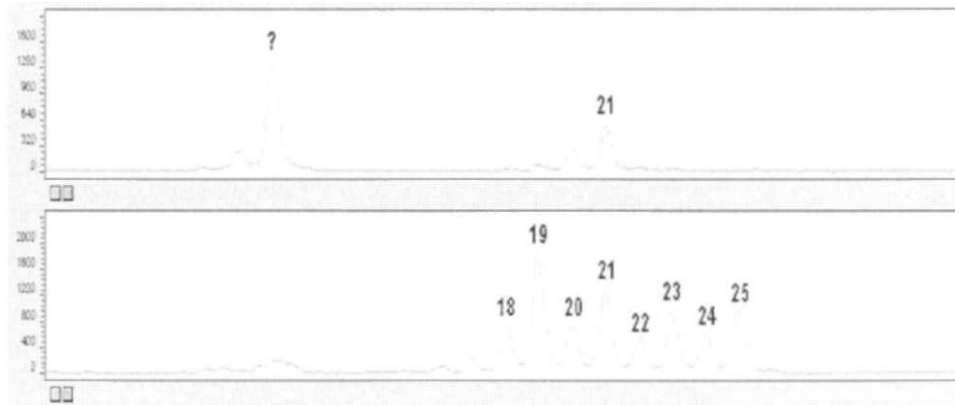
Reakcję sekwencjonowania prowadzono przy użyciu zestawu ABI Prism BigDye Terminator v1.1 Cycle Sequencing Kit we wspomnianym termocyklerze. Do sekwencjonowania użyto 5 ng produktu PCR oraz 1,6 pmol nieznakowanego fluorescencyjnie startera w łącznej objętości 10 µl. Profil termiczny obejmował 2 min wstępnej denaturacji w 96 °C oraz 30 cykle reakcji wydłużania startera (30 s w 96 °C, 15 s w 50 °C, 4 min w 60 °C). Szybkość zmiany temperatury pomiędzy poszczególnymi etapami ustalono na 1 °C/s. Produkty sekwencjonowania oczyszczono poprzez precypitację etanolem w obecności octanu sodu (1). Osad wysuszyli w 90 °C przez 1 min i zawieszono w 10 µl dejonizowanego formamidu.

Elektroforeza kapilarna produktów sekwencjonowania

Po 2-minutowej denaturacji w 95°C próbkę szybko schłodzono w lodzie. Produkty reakcji rozdzielano metodą elektroforezy kapilarnej na wspomnianym sekwenatorze po 8-sekundowym nastrzyku przy 2 kV przy zastosowaniu polimeru POP-6 (Applied Biosystems), modułu szybkiego sekwencjonowania i zestawu filtrów E. Rozdział prowadzono przez 28 min w 50 °C przy 15 kV. Do analizy wyników posłużył program komputerowy Sequencing Analysis 3.7 (Applied Biosystems).

WYNIKI I DISKUSJA

Uzyskany obraz elektroforegramu produktów amplifikacji przedstawia rycina 1,



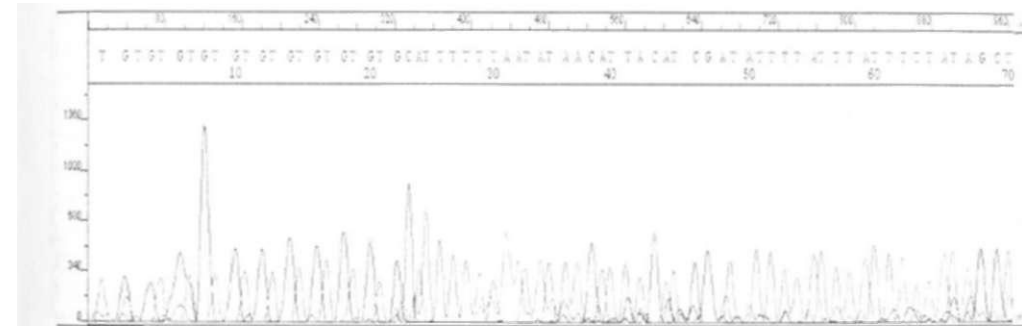
Ryc. 1. Elektroforegram produktów amplifikacji systemu YCAII dla próbek o nieznanym krótkim allelu oraz drabiny alleli złożonej z ośmiu znanych alleli wykazanych w populacji 154 niespokrewnionych mężczyzn.

Fig. 1. Electrophoregram of amplification products of the YCAII system for a sample with an unknown short allele and an allelic ladder consisting of eight known YCAII alleles found in a population of 154 unrelated males.

W badanej próbce stwierdzono obecność 2 alleli: YCAIT21 oraz allela krótszego od poprzedniego o 20 par zasad. Identyczny układ alleli YCAII stwierdzono u biologicznego syna badanego mężczyzny, co potwierdziło dziedziczenie badanego allela sprzężone z chromosomem Y. Obraz amplifikacji tego allela różnił się wyraźnie od obrazu obserwowanego dla innych, dłuższych alleli tego systemu. Podobnie jak dla innych loci o powtórzeniach sekwencji dwunukleotydowej (8) stwierdzono charakterystyczną obecność krótszego o dwie pary zasad artefaktu będącego wynikiem poślizgu polimerazy DNA w trakcie reakcji PCR. Jednakże sygnał amplifikacji tego artefaktu był wyraźnie słabszy, zaś produktu krótszego o cztery pary zasad praktycznie nie zaobserwowano. Stosunek

powierzchni piku odpowiadającego produktowi poślizgu n-2 do powierzchni piku odpowiadającego właściwemu allelowi wyniósł w przypadku tego allela zaledwie 22,6%, podczas gdy dla allela YCAIT21 stosunek ten sięgnął 53,4%. Wyniki te są zgodne z wcześniejszymi obserwacjami dla loci o powtórzeniach czteronukleotydowych, dla których stosunek ilości artefaktu do ilości właściwego allela rośnie wraz z liczbą jednakowych jednostek repetytywnych (7, 13).

Aby oznaczyć badany allel zgodnie z zaleceniami Międzynarodowego Towarzystwa Genetyki Sądowej (ISFG) (4) przeprowadzono jego sekwencjonowanie (rycina 2). Jako że sekwencja rozpoznawana przez drugi starter leży bezpośrednio obok sekwencji repetytywnej systemu YCAII (rycina 3), uzyskana sekwencja rozpoczyna się 11 powtórzeniami sekwencji TG, co odpowiada sekwencji komplementarnej (CA)_n. Nie stwierdzono różnic w regionie flankującym w stosunku do sekwencji referencyjnej. Zgodnie z nową nomenklaturą alleli YCAII, uwzględniającą liczbę powtórzeń jednostki repetytywnej (11), badany allel oznaczono jako YCAII*11.



Ryc. 2. Wynik sekwencjonowania nieznanego allela YCAII z 11 powtórzeniami jednostki (TG)₁₁ na początku sekwencji.

Fig. 2. Result of sequencing of an unknown YCAII allele with 11 repetitive units (TG)₁₁ at the beginning of the sequence.

1	<u>TGTCAAAATT</u>	<u>TAACCCACAA</u>	TCATAAAAAATATACTTAATA
41	<u>AGCTATAAAA</u>	<u>ATAAATAAAATATCG</u>	ATGTAATGTTATATT
81	<u>AAAAAATGCA</u>	<u>CACACACACA</u>	CACACACACACACACACACA
121	<u>CACACACACA</u>	<u>CACAGCTAACCTTATGGTGA</u>	<u>AAGACTG</u>

Ryc. 3. Sekwencja DNA amplifikowana przez zastosowane w pracy startery (nukleotydy 79865-80021 sekwencji AC015978, dostępnej w bazie danych NCBI pod adresem <http://www.ncbi.nlm.nih.gov>). Podkreślono sekwencje rozpoznawane przez startery. Kolorem szarym wyróżniono 23 powtórzenia sekwencji (CA)₂₃ allela YCAIT23.

Fig. 3. DNA sequence amplified with the primers used in the present work (nucleotides 79865-80021 of the sequence AC015978 available in the NCBI database at <http://www.ncbi.nlm.nih.gov>). Sequences recognised by primers are underlined. 23 repeat motifs (CA)₂₃ of the YCAII*23 allele are highlighted in grey.

Korzystając z Referencyjnej Bazy Danych Haplotypów Chromosomu Y (<http://www.yhrd.org>, ostatnie uaktualnienie z 29.12.2003.) zbadano występowanie allela YCAIT11 w różnych populacjach na świecie. Na 3998 rozszerzonych haplotypów z 27 różnych populacji (w tym 3313 rozszerzonych haplotypów z 17 populacji europejskich), zebranych w bazie danych, allel YCAIT11 wystąpił zaledwie trzykrotnie. Allel ten był obecny wyłącznie w populacji południowo-zachodnich Niemiec (rejon Freiburga) w postaci identycznego do wykazanego w badanej próbce haplotypu YCAII*11,21. Zestawienie wszystkich czterech haplotypów posiadających allel YCAIT11 wykazało dalsze podobieństwa (tabela I). Wszystkie te haplotypy są identyczne lub prawie identyczne nie tylko pod względem bardzo rzadkiego haplotypu YCAII, ale również pod względem alleli DYS392, DYS393 i DYS385, co może sugerować ich wspólne pochodzenie.

Tabela I. Porównanie znanych haplotypów posiadających allel YCAIT11.

Table I. Comparison of known haplotypes possessing YCAII*11 allele.

populacja population	DYS 19	DYS 389I	DYS 389II	DYS 390	DYS 391	DYS 392	DYS 393	DYS 385	YCAII
Freiburg	17	13	28	23	10	11	13	11,12	11,21
Freiburg	16	13	28	22	11	11	13	12,12	11,21
Freiburg	15	14	30	24	9	11	13	12,12	11,21
Gdańsk	16	14	29	24	10	11	13	12,12	11,21

Amplifikacja metodą PCR loci mikrosatelitarnych prowadzi do jednoczesnej syntezy artefaktów krótszych o wielokrotność jednostki repetytywnej, których obecność może utrudniać interpretację wyników w przypadku mieszanin, jako że wspomniane artefakty mają rozmiary odpowiadające właściwym allelom. Stosunek artefaktu do właściwego allela jest tym wyższy, im krótsza jest jednostka repetytywna mikrosatelity i im więcej powtórzeń jednostki posiada dany allel (2). System YCAII należy do loci STR chromosomu Y o powtórzeniach sekwencji dwunukleotydowej, stąd udział tego typu artefaktów jest szczególnie wysoki. Jednakże przy analizie DNA pochodzącego od jednej osoby, z jakim ma się do czynienia w ogromnej większości przypadków w badaniach mających na celu ustalenie pokrewieństwa, interpretacja wyników amplifikacji loci STR o powtórzeniach dwunukleotydowych nie budzi wątpliwości. System YCAII włączony do Referencyjnej Bazy Danych Haplotypów Y-STR (<http://www.ystr.org>, ostatnie uaktualnienie z 30.12.2003.) zwiększył udział różnych haplotypów zdeponowanych w bazie danych z 41% do 65%. Wysoki stopień polimorfizmu YCAII (6), dostosowanie nazewnictwa alleli do zaleceń ISFG (11) oraz włączenie do ogólnie dostępnych baz danych umożliwiających szacowanie częstości haplotypów w różnych populacjach (10) decydują o dużej wartości tego systemu dla badań w genetyce sądowej.

PIŚMIENNICTWO

- Applied Biosystems (2000): ABI Prism BigDye Terminator Cycle Sequencing Ready Reaction Kits, original and version 2.0, Foster City, Kalifornia, USA. - 2. Bacher J., Schumm J.W. (1998): Development of highly polymorphic pentanucleotide tandem repeat loci with low stutter, *Profiles DNA*, 2 (2), 3-6. - 3. Butler J.M., Schoske R., Vallone P.M., Kline M.C., Redd A.J., Hammer M.F. (2002): A novel multiplex for simultaneous amplification of 20 Y chromosome STR markers, *Forensic Sci. Int.* 129, 10-24. - 4. Gili P., Brenner C., Brinkmann B., Budowle B., Carracedo A., Jobling M.A., de Knijff P., Kayser M., Krawczak M., Mayr W.R., Morling N., Olaisen B., Pascali V., Prinz M., Roewer L., Schneider P.M., Sajantila A., Tyler-Smith C. (2001): DNA Commission of the International Society of Forensic Genetics: recommendations on forensic analysis using Y-chromosome STRs, *Int. J. Legal Med.* 114: 305-309. - 5. Jobling M.A., Pandya A., Tyler-Smith C. (1997): The Y chromosome in forensic analysis and paternity testing, *Int. J. Legal Med.* 110, 118-124. - 6. Kayser M., Caglia A., Córach D., Fretwell N., Gehrig C., Graziosi G., Heidorn F., Herrmann S., Herzog B., Hidding M., Honda K., Jobling M., Krawczak M., Leim K., Meuser S., Meyer E., Oesterreich W., Pandya A., Parson W., Penacino G., Perez-Lezaun A., Piccinini A., Prinz M., Schmitt C., Schneider P.M., Szibor R., Teifel-Greding J., Weichhold G., de Knijff P., Roewer L. (1997): Evaluation of Y-chromosomal STRs: a multicenter study, *Int. J. Legal Med.* 110, 125-133. - 7. Klitsch M., Wiegand P. (2003): Polymerase slippage in relation to the uniformity of tetrameric repeat stretches, *Forensic Sci. Int.* 135, 163-166. - 8. Murray V., Monchawin C., England P.R. (1993): The determination of the sequences present in the shadow bands of a dinucleotide repeat PCR, *Nucleic Acids Res.* 21, 2395-2398. - 9. Redd A.J., Agellon A.B., Kearney V.A., Contreras V.A., Karafet T., Park H., de Knijff P., Butler J.M., Hammer M.F. (2002): Forensic value of 14 novel STRs on the human Y chromosome, *Forensic Sci. Int.* 130, 97-111. - 10. Roewer L., Krawczak M., Willuweit S., Nagy M., Alves C., Amorim A., Anslinger K., Augustin C., Betz A., Bosch E., Caglia A., Carracedo A., Córach D., Dekairelle A-F., Dobosz T., Dupuy B.M., Furedi S., Gehrig C., Gusmao L., Henke J., Henke L., Hidding M., Hohoff C., Hoste B., Jobling M.A., Kargel H.J., de Knijff P., Lessig R., Liebeherr E., Lorante M., Martinez-Jarreta B., Nievas P., Nowak M., Parson W., Pascali V.L., Penacino G., Ploski R., Rolf B., Sala A., Schmidt U., Schmitt C., Schneider P.M., Szibor R., Teifel-Greding J., Kayser M. (2001): Online reference database of European Y-chromosomal short tandem repeat (STR) haplotypes, *Forensic Sci. Int.* 118, 106-113.

11. Schmidt U., Lutz S., Roewer L. (2003): A nomenclature for YCA II which is compatible with the ISFG guidelines for Y-STR analysis, *Int. Congr. Ser.* 1239, 481-485. - 12. Singer V.L., Jones L.J., Yue S.T., Haugland R.P. (1997): Characterization of PicoGreen reagent and development of a fluorescence-based solution assay for double-stranded DNA quantitation, *Anal. Biochem.* 249, 228-238. - 13. Walsh P.S., Fildes N.J., Reynolds R. (1996): Sequence analysis and characterization of stutter products at the tetranucleotide repeat locus vWA, *Nucleic Acids Res.* 24, 2807-2812.

Adres pierwszego autora:
Katedra i Zakład Medycyny Sądowej AM
ul. Marii Skłodowskiej-Curie 3A
80-210 Gdańsk

Renata Jacewicz, Jarosław Berent, Katarzyna Bąbol, Stefan Szram

Rozkład częstości alleli w 10 loci STR w regionie centralnej Polski¹

The allele distribution of the ten STR loci in Central Poland population

Z Katedry i Zakładu Medycyny Sądowej Uniwersytetu Medycznego w Łodzi
Kierownik: prof. dr hab. med. S. Szram

Przedstawiono bazę danych dla 10 loci STR: D3S1358, vWA, D16S539, D2S1338, D8S1179, D21S11, D18S51, D19S433, TH01, FGA opracowaną na podstawie analizy 350-osobowej próby populacyjnej z regionu Polski centralnej (region łódzki). Wykazano stan równowagi Hardyego i Weinberga (HWE) dla badanych układów. Przeprowadzono analizę porównawczą uzyskanych rozkładów z innymi populacjami.

Allele frequency data for ten STR loci: D3S1358, vWA, D16S539, D2S1338, D8S1179, D21S11, D18S51, D19S433, TH01, FGA were estimated in a sample of 350 unrelated individuals in Central Poland population (Łódź region). The accordance with the Hardy and Weinberg equilibrium (HWE) for all investigated loci was proved. The allele distributions were compared to the similar data for other populations.

Słowa kluczowe: STR multiplex, dane populacyjne, region łódzki, Polska centralna

Key words: multiplex STR, population data, Łódź region, Central Poland

WSTĘP

Jednoczesna amplifikacja wielu loci STR z różnych miejsc ludzkiego DNA tzw. STR multiplex jest podstawową techniką badawczą stosowaną w genetyce sądowej. Wykorzystując tę technikę opracowaliśmy rozkład częstości alleli dla

¹ Temat opracowany w ramach prac własnych uczelni nr 502-11-107.