

Monika Reichert, Agnieszka Maciejewska, Krzysztof Talarczyk, Regina Paszkowska, Joanna Jakubowska, Anita Dettlaff-Kąkol, Izabela Maciejewska¹, Ryszard Pawłowski

Polimorfizm locus SE33 w populacji polskiej

Polymorphism in the SE33 locus in the Polish population

Katedra i Zakład Medycyny Sądowej Gdańskiego Uniwersytetu Medycznego

Kierownik: dr hab. med. Z. Jankowski

¹ Katedra i Zakład Protetyki Stomatologicznej

Kierownik: prof. dr hab. Z. Bereznowski

Locus ACTBP2 (human beta-actin related pseudogene; SE33) jest najbardziej polimorficznym ze wszystkich, dotychczas stosowanych markerów w genetyce sądowej, zarówno do celów badania pokrewieństwa jak i identyfikacji śladów biologicznych. Ze względu na bardzo liczne allele i stosunkowo małą tendencję do tworzenia stutterów, jest prawie idealnym kandydatem do zastosowań z zakresu genetyki sądowej [1]. Do wad tego systemu należą stosunkowo długie amplikony, które mogą sprawiać trudności w przypadku badania zdegradowanego DNA. Duża długość amplikonów SE33 jest oczywiście związana z ogromną liczbą alleli obserwowanych dla tego locus. Hiperzmienny locus SE33 zlokalizowany jest na 6 chromosomie człowieka (GeneBank Accession: V00481) i wykazuje szczególnie złożony polimorfizm, na który składa się flankujący obszar 5', centralna sekwencja repetytywna i flankujący obszar 3' [2]. W zależności od rodzaju komercyjnego zestawu użytego do badania polimorfizmu SE33, a tym samym rodzajem starterów użytych do powielania tego locus, obserwowane wielkości amplikonów mieszczą się w zakresie od 197 do 381pz dla zestawu ABI SEfiler, od 258 do 442pz dla ESX17 Promegi oraz od 300 do 484pz dla zestawu ES17 Promegi (zakres alleli od 3-49). Dotychczas opisano aż 178 alleli w zakresie tego locus [3]. Ze względu między innymi na bardzo wysoką siłę dyskryminacji, locus ten znalazł się wśród ośmiu wybranych loci do Niemieckiej Bazy DNA (D3S1358, FGA, D8S1179, D18S51, D21S11, TH01, VWA, SE33, amelogenina) [4, 5].

Poza genetyką sądową locus SE33 znalazł również zastosowanie między innymi w badaniu chimerizmu po przeszczepach szpiku kostnego [6, 7].

Celem niniejszej pracy jest ponowna analiza polimorfizmu tego locus przeprowadzona dla populacji Polski północnej w liczącej ponad 1000 osobników próbie populacyjnej.

MATERIAŁ I METODY

Badaniom populacyjnym poddano materiał biologiczny pochodzący od 1007 osobników obu płci zamieszkujących głównie obszar Polski północnej. Badaniom poddano tylko próbki pochodzące od niespokrewnionych osób, przesyłanych jako materiał referencyjny do spraw badanych w Pracowni Biologii i Genetyki Sądowej ZMS GUMed. DNA izolowano metodą organiczną fenol-chloroform. Ilość wyizolowanego DNA oznaczano techniką RT PCR z zastosowaniem zestawu Quantifiler™ na urządzeniu 7900HT RealTime Fast PCR System (Applied Biosystems), jak również techniką fluorymetrii z fluorochromem PicoGreen™ (Fluoroskan Ascent, ThermoLabsystem, Finlandia).

Do namnażania DNA użyto komercyjnych zestawów PowerPlex ESX17 firmy Promega (USA) oraz SEfiler firmy ABI (USA). Reakcję PCR prowadzono w termocyklerze ABI 9700 firmy Applied Biosystems, zgodnie z zaleceniami producenta. Produkty PCR identyfikowano metodą elektroforezy kapilarnej na sekwenatorze ABI 3130 Avant firmy Applied Biosystems (USA) wykorzystując oprogramowanie Data Collection v. 3.0 i Gene Mapper v. 3.2. Jako standardu wewnętrznego wielkości DNA użyto markera CC5 ILS500 firmy Promega (USA) dla zestawu PowerPlex ESX17 oraz LIZ500 w przypadku zestawu SEfiler. Otrzymane fragmenty DNA porównywano z drabiną alleli (Promega, USA), która zawiera

rafa 37 alleli: 4.2, 6.3, 8-20, 20.2, 21, 21.2, 22, 22.2, 23.2, 24.2, 25.2, 26.2, 27.2, 28.2, 29.2, 30.2, 31.2, 32.2, 33.2, 34.2 35-37, 39 i 42 lub z drabiną dla zestawu SEfiler zawierającą łącznie 50 alleli: 4.2, 6.3, 8-37, 21.1, 21.2, 22.2, 23.2, 24.2, 25.2, 26.2, 27.2, 28.2, 29.2, 30.2, 31.2, 32.2, 33.2, 34.2, 35.2, 36, 37.

Obliczenia statystyczne

Weryfikację zgodności z równaniem Hardy'ego-Weinberga przeprowadzono za pomocą programu komputerowego HWE wersja 1.20 [8]. Obserwowaną heterozygotyczność (H), siłę dyskryminacji (PD), średnią szansę wykluczenia (MEC) i zawartość informacji polimorficznej (PIC) obliczano jak uprzednio [9, 10, 11]. Homogenność rozkładu alleli SE33 pomiędzy badaną próbką populacyjną, a innymi populacjami prowadzono programem Conting v. 2.36 [12]. Zróznicowanie rozkładu alleli pomiędzy populacjami oceniano również za pomocą programu Arlequin v. 3.1, badając wartości F_{ST} dla poszczególnych par populacji [13].

WYNIKI I DISKUSJA

Tabela I przedstawia obserwowany rozkład częstości alleli SE33 w populacji Polski północnej A, dla 1007 osobników oraz dla dwóch innych, uprzednich badań populacyjnych Podlasia (N=220) [14] i Polski północnej B (N=255) [15].

W naszych badaniach obserwowaliśmy łącznie 46 alleli, w porównaniu do 35 w populacji Podlasia i 37 w populacji Polski północnej B, na co wyraźny wpływ miała wielkość próbki populacyjnej. Nasze poprzednie badania dotyczące polimorfizmu DNA dla locus SE33 opublikowane w 1996 roku dla próbki 176 osób, pozwoliły na zidentyfikowanie tylko 25 alleli [16]. Liczba ta wydaje się być zaniżona, ze względu na stosowaną metodę detekcji amplikonów za pomocą elektroforezy na ultracienkich, poziomych żelach poliakrylamidowych. Niektóre z obserwowanych alleli nie były obecne w komercyjnych drabinach alleli, a ich oznaczenie zostało oparte na precyzyjnym pomiarze wielkości fragmentów i nie było weryfikowane sekwencjonowaniem. Dotyczyło to alleli 12.2, 13.2, 14.2, 15.2, 15.3, 16.2, 19.2 i 30.3. Badania 1007 próbek DNA dla populacji Polski północnej A prowadzono w większości stosując zestaw PowerPlex ESX17

(785 próbek), pozostałe zestawem SEfiler. Żadna z próbek nie była weryfikowana ponownym badaniem innym zestawem, chociaż miałyby to swoje logiczne uzasadnienie zważywszy na fakt obserwowanych niezgodności przy zastosowaniu innych starterów [17, 18].

Najczęściej obserwowanym allelem dla próbki 2014 alleli był allel 17 ($f=0,077$), a obok niego 27.2 ($f=0,076$), 28.2 ($f=0,075$), 19 ($f=0,065$) i 29.2 ($f=0,061$). Dla populacji Polski północnej B, w której badano 510 alleli [15], najczęstszy był allel 28.2 ($f=0,092$), a obok niego 19 ($f=0,080$), 29.2 ($f=0,075$) i 17 wraz z 19 ($f=0,067$). Badanie porównawcze częstości alleli, zaobserwowanych w tych dwóch badaniach dla Polski północnej A i B, nie wykazały znaczących różnic statystycznych ($\chi^2=9,07$; 8df; $p=0,17$ oraz $F_{ST}=0,00005$; $P=0,40451$).

W populacji z Podlasia [14] najczęściej obserwowanym allelem były allele 28 i 28.2 ($f=0,084$), a ponadto 19 ($f=0,082$), 17 ($f=0,070$) oraz 30.2 ($f=0,063$). Rozkład alleli różnił się znamienne statystycznie przy porównaniu zarówno z populacją A z Polski północnej ($F_{ST}=0,00662$; $P=0,00000$, a dla testu $\chi^2=49,8$; 8df; $p=0,000$) jak i z populacją B z Polski północnej ($F_{ST}=0,00590$; $P=0,00000$). Zważywszy na raczej dużą jednorodność populacji polskiej obserwowanej dla różnych innych autosomalnych markerów genetycznych typu STR, te różnice wydają się być zaskakujące. Ponieważ przyczyna różnego rozkładu alleli pomiędzy populacją z Polski północnej, a Podlasiem nie jest jasna, wydaje się, że wskazane byłoby przy prowadzeniu obliczeń statystycznych w sprawach identyfikacyjnych czy badania pokrewieństwa stosowanie jednak lokalnych częstości alleli.

Badanie zgodności rozkładu częstości obserwowanych i oczekiwanych alleli w populacji, z prawem Hardy-Weinberga wykazało brak odstępstwa od tego prawa ($\chi^2=42,8$; 45df; $p=0,23$).

Locus ten, jak już wspomniano na wstępie, jest uważany za najbardziej polimorficzny z dotychczas stosowanych w badaniach z zakresu genetyki sądowej, zarówno do celów badania pokrewieństwa jak i identyfikacji śladów biologicznych. Potwierdzają to również badania przeprowadzone dla dużej próbki populacyjnej z obszaru Polski północnej. W żadnym z dotychczas rutynowo badanych loci dostępnych w komercyjnych zestawach nie obser-

Tabela I. Częstości alleli oraz podstawowe parametry statystyczne dla locus SE33 w trzech populacjach z obszaru Polski.

Table I. Allele frequency and major statistical parameters for the SE33 locus in three Polish populations.

Allele	Częstości alleli w populacji Polski północnej A (N=1007) Allele frequency in the Northern Polish population A (N=1007)		Częstości alleli w populacji Polski północnej B (N=255) Allele frequency in the Northern Polish population B (N=255)		Częstości alleli w populacji Podlasia (N=220) Allele frequency in the Podlasie region population (N=220)	
11	0.0000	0	0.0000	0	0.0023	1
11.2	0.0000	0	0.0020	1	0.0000	0
12	0.0030	6	0.0059	3	0.0136	6
12.2	0.0005	1	0.0000	0	0.0000	0
13	0.0070	14	0.0098	5	0.0023	1
13.2	0.0015	3	0.0020	1	0.0023	1
14	0.0318	64	0.0235	12	0.0273	12
14.2	0.0070	14	0.0000	0	0.0000	0
15	0.0422	85	0.0510	26	0.0500	22
15.2	0.0010	2	0.0000	0	0.0000	0
15.3	0.0005	1	0.0020	1	0.0000	0
16	0.0501	101	0.0431	22	0.0477	21
16.2	0.0010	2	0.0000	0	0.0000	0
17	0.0775	156	0.0529	27	0.0705	31
17.3	0.0000	0	0.0000	0	0.0023	1
18	0.0546	110	0.0667	34	0.0432	19
18.3	0.0000	0	0.0000	0	0.0023	1
19	0.0650	131	0.0804	41	0.0818	36
19.2	0.0050	10	0.0059	3	0.0091	4
20	0.0546	110	0.0588	30	0.0432	19
20.2	0.0119	24	0.0020	1	0.0091	4
21	0.0223	45	0.0294	15	0.0341	15
21.1	0.0000	0	0.0020	1	0.0000	0
21.2	0.0209	42	0.0137	7	0.0182	8
22	0.0094	19	0.0039	2	0.0068	3
22.2	0.0283	57	0.0275	14	0.0341	15
23	0.0015	3	0.0020	1	0.0000	0
23.2	0.0397	80	0.0412	21	0.0227	10
24	0.0030	6	0.0000	0	0.0023	1
24.2	0.0482	97	0.0373	19	0.0341	15
25	0.0005	1	0.0000	0	0.0000	0
25.2	0.0407	82	0.0510	26	0.0432	19
26	0.0015	3	0.0000	0	0.0000	0
26.2	0.0586	118	0.0608	31	0.0432	19
27	0.0005	1	0.0000	0	0.0000	0
27.2	0.0765	154	0.0667	34	0.0000	0
28	0.0015	3	0.0000	0	0.0841	37
28.2	0.0755	152	0.0922	47	0.0841	37
29	0.0015	3	0.0000	0	0.0000	0
29.2	0.0611	123	0.0745	38	0.0523	23
30	0.0005	1	0.0000	0	0.0023	1
30.2	0.0427	86	0.0373	19	0.0636	28
30.3	0.0005	1	0.0000	0	0.0000	0
31	0.0015	3	0.0020	1	0.0000	0
31.2	0.0248	50	0.0196	10	0.0273	12
32	0.0000	0	0.0020	1	0.0000	0
32.2	0.0114	23	0.0118	6	0.0227	10
33	0.0020	4	0.0000	0	0.0091	4
33.2	0.0045	9	0.0118	6	0.0045	2
34	0.0035	7	0.0020	1	0.0023	1
34.2	0.0015	3	0.0020	1	0.0000	0
35.2	0.0020	4	0.0020	1	0.0000	0
36	0.0000	0	0.0020	1	0.0023	1
Het		0.938		0.941		0.949
PIC		0.950		0.940		0.950
PD		0.994		0.991		0.990
PE		0.874		0.880		-
pM		0.006		0.009		0.010

wowano tak wielu alleli o tak korzystnym dla celów genetyki sądowej rozkładzie częstości. Zgodnie z oczekiwaniami dla locus SE33 obserwowano bardzo wysokie wartości parametrów statystycznych jak: PIC (zawartość informacji polimorficznej), Het (obserwowana heterozygotyczność), PD (siła dyskryminacji), PE (Power of exclusion – siła wykluczenia) czy pM (prawdopodobieństwo przypadkowej zgodności) potwierdzających jego ogromną

przydatność do celów identyfikacyjnych (tab. I). Najwyższe wartości siły dyskryminacji (PD), a tym samym najmniejsze wartości przypadkowej zgodności (pM) uzyskano dla populacji z Polski północnej A. Wynosiły one odpowiednio 0,994 i 0,006 i potwierdzają ogromną przydatność tego locus do celów identyfikacji śladów biologicznych, w tym interpretacji złożonych profili DNA w postaci ich mieszanin.

PIŚMIENICTWO

1. Ghosh A., Seshadri M.: Forensic assessment of ACTBP2 (SE33) microsatellite. *J Forensic Sci* 2003, 48: 1195-1196.

2. Dauber E. M., Schwartz-Jungl E. M., Wenda S., Dorner G., Glock B., Mayr W. R.: Further allelic variation at the STR-loci ACTBP2 (SE33), D3S1358, D8S1132, D18S51 and D21S11 *Forensic Sci Int: Genetics Supplement Series*. 2009, 1: 41-42.

3. *Advanced Topics In Forensic DNA Typing: Methodology* John M. Butler, Elsevier 2012.

4. Müller K., Klein R., Miltner E., Wiegand P.: Validation of the Short Amplicon Multiplex Q8 Including the German DNA Database Systems. *J Forensic Sci*. 2009, 54: 862-865.

5. Martin P. D., Schmitter H., Schneider P. M.: A brief history of the formation of DNA databases in forensic science within Europe. *Forensic Sci Int*. 2001, 19: 225-231.

6. Hellmann A., Zaucha M., Pawłowski R., Welz A., Hałaburda K.: Amplification of the VNTR systems HUMACTBP2 and HUMTHO1 by the polymerase chain reaction for documentation of residual recipient cells after allogeneic bone marrow transplantation. *Appl. Biol. Commun.* 1996, 6, suppl. 1: 127-134.

7. Zaucha J. M., Pawłowski R., Welz A., Prejzner W., Hauser R., Hellmann A.: Ocena przyjęcia przeszczepu szpiku oraz chimeryzmu hemopoetycznego przy użyciu amplifikacji metodą PCR hiperzmiennych sekwencji typu VNTR. *Pol. Tyg. Lekarski*. 1995, 50: 73-74.

8. Ott J.: Utility programs for analysis of Genetic

Linkage, program HWE version 1.10, Rockefeller University, NY, USA.

9. Pawłowski R.: HUMFIBRA allele distribution in Northern Poland using capillary electrophoresis. *Int. J. Legal Med.* 1999, 112: 139-141.

10. Pawłowski R., Maciejewska A.: The forensic validation of multiplex containing nine STRs – population genetics in Northern Poland. *Int. J. Legal Med.* 2000, 114: 45-49.

11. Pawłowski R., Maciejewska A., Paszkowska R., Welz A.: Frequencies of the five short tandem repeat systems (STR) in a population from North Poland. *Int. J. Legal Med.* 1997, 110: 10-13.

12. Ott J.: Utility programs for analysis of Genetic Linkage, program CONTING v. 2.36, Rockefeller University, NY, USA.

13. Excoffier L., Laval G., Schneider S.: Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evol. Bioinf. Online* 2005, 1: 47-50.

14. Janica J., Pepiński W., Skawrońska M., Niemcunowicz-Janica A., Koc-Żórawska E.: Polimorfizm 10 autosomalnych loci STR w populacji Podlasia – rozszerzenie typowego panelu badawczego. *Arch. Med. Sąd. Kryminol.* 2007, 57: 248-251.

15. Wysocka J., Kapińska E., Cybulska L., Rębała K., Szczerkowska Z.: Identyfikacja osobnicza w oparciu o analizę 11 polimorficznych loci DNA typu STR. *Badania populacji Polski północnej z wykorzystaniem zestawu AmpFISTR SEfiler TM Kit*, *Ann. Acad. Med. Gedan.* 2008, 38: 91-96.

16. Pawłowski R., Maciejewska A., Paszkowska R. and Welz A.: The current status of forensic DNA analysis in Poland. *Proceedings from the First*

European Symposium on Human Identification. 1996: 116-125.

17. Davis C., Ge J., King J., Malik N., Weirich V. Eisenberg A. J., Budowle B.: Variants observed for STR locus SE33: A concordance study. *Forensic Sci. Int: Genetics* 2012, 6: 494-497.

18. Butler J. M., Hill C. R., Kline M. C., Bastisch I., Weirich V., McLaren R. S., Storts D. R: SE33 variant alleles: sequences and implications. *Forensic Sci. Int. Genet.*: 2011, Suppl. Ser. 3: e502-e503.

Adres do korespondencji:

Zakład Medycyny Sądowej GUMed,

ul. Dębowa 23

80-204 Gdańsk

e-mail: Richard@gumed.edu.pl (R. Pawłowski)