

**Tomasz Grzybowski¹, Boris A. Malyarchuk², Jarosław Bednarek¹, Marcin Woźniak¹,
Marta Papuga¹, Katarzyna Stopińska¹, Sylwia Łuczak¹**

Zastosowanie analizy filogeograficznej w interpretacji wyników sekwencjonowania mitochondrialnego DNA dla celów sądowych

Phylogeographic approach in the interpretation of mitochondrial DNA sequencing results in forensics

¹ Z Zakładu Genetyki Molekularnej i Sądowej Katedry Medycyny Sądowej Collegium Medicum im. Ludwika Rydygiera w Bydgoszczy Uniwersytetu Mikołaja Kopernika w Toruniu
Kierownik: prof. dr hab. Karol Śliwka

² Z Laboratorium Genetyki Instytutu Biologicznych Problemów Północy Rosyjskiej Akademii Nauk, Magadan, Rosja

W ostatnim czasie badaniom sekwencji mitochondrialnego DNA dla celów sądowych towarzyszy wyraźna perspektywa ewolucyjna. Szczególnie przydatne w interpretacji wyników sekwencjonowania mtDNA w sprawach identyfikacyjnych okazuje się podejście filogeograficzne, które obejmuje filogenetyczną analizę przestrzennego rozkładu haplotypów i haplogrup w populacjach. W ramach niniejszej pracy wykazano przydatność podejścia filogeograficznego w interpretacji trudniejszych przypadków identyfikacyjnych, w których obserwowano jednoklebotowe, homoplazmatyczne różnice sekwencji mtDNA pomiędzy haplotypami dowodowymi i porównawczymi. Dodatkowo, w celu ustalenia przynależności haplogrupowej wybranych haplotypów przeprowadzono analizę pełnych sekwencji dwóch genomów mitochondrialnych należących do podkladu J1b.

In recent years, forensic mitochondrial DNA analysis has been undertaken from an evolutionary perspective. In particular, the phylogeographic approach based on a phylogenetic analysis of the spatial distribution of mitochondrial haplotypes and haplogroups appears to be a useful tool in the interpretation of identification cases. In this study, the phylogeographic approach has been employed in the analysis of three difficult forensic cases, where single nucleotide, homoplasmic differences were found between the reference and evidentiary haplotypes. mtDNA sequence variation has been examined by the control region (HVS I and HVS II) direct sequencing. Additionally, in order to clarify

the subhaplogroup status of the selected haplotypes, DNA sequences of entire mitochondrial genomes obtained from two samples representing J1b subclade have been analyzed.

WSTĘP

Choć analiza sekwencji mitochondrialnego DNA (mtDNA) stała się jedną z dobrze ugruntowanych technologii badania polimorfizmu DNA dla potrzeb wymiaru sprawiedliwości [1], niektóre zagadnienia dotyczące interpretacji wyników sekwencjonowania mtDNA dla celów sądowych wciąż są przedmiotem ożywionych dyskusji. Do zagadnień tych należą m.in. wielkość, zasięg geograficzny oraz jakość populacyjnych baz danych mtDNA wykorzystywanych dla oceny częstości haplotypów, zróżnicowane tempo mutacji dla różnych pozycji sekwencji regionu kontrolnego i związane z nim różne kryteria statystycznej oceny wyników potwierdzających lub wykluczających pochodzenie badanych próbek z tej samej linii matczynej, czy też ograniczona wartość dowodowa analizy sekwencji regionu kontrolnego, a stąd konieczność włączenia do badań wybranych wariantów regionu kodującego [2, 3, 4]. W ostatnim czasie analizom sekwencji mtDNA dla celów sądowych towarzyszy wyraźna perspektywa ewolucyjna [2]. W szczególności, dzięki podejściu filogeograficznemu, które obejmuje analizę prze-

strzennego, geograficznego rozkładu haplotypów w drzewie filogenetycznym mtDNA, podjęto próby identyfikacji regionu geograficznego, z którego pochodzą badane haplotypy mtDNA oraz *a posteriori* dokonano kontroli jakości danych zdeponowanych w niektórych sądowych populacyjnych bazach profili mtDNA [2 oraz zawarte tam pozycje piśmiennictwa]. Ponieważ rozdzielczość analiz filogenetycznych osiągnęła w ostatnim czasie maksymalny poziom dzięki badaniom populacyjnym wykorzystującym sekwencjonowanie pełnych genomów mitochondrialnych [5 oraz zawarte tam pozycje piśmiennictwa], znacznie łatwiejsze niż dotychczas stały się poszukiwania dodatkowych markerów regionu kodującego, które mogą zwiększyć informatywność analiz mtDNA dla celów sądowych. Ponadto, szczegółowa znajomość filogenezy mtDNA pozwala na bardziej wiarygodną ocenę częstości haplotypów regionu kontrolnego w populacjach zamieszkujących różne regiony geograficzne.

Dane przedstawione w niniejszej pracy ilustrują użyteczność podejścia filogeograficznego w interpretacji trudniejszych przypadków identyfikacyjnych pochodzących z praktyki Zakładu Genetyki Molekularnej i Sądowej, w których zachodziła konieczność analizy mtDNA. W przypadkach tych zanotowano jednonukleotydowe, homoplazmatyczne różnice pomiędzy haplotypami uzyskanymi dla próbek dowodowych i porównawczych. Interpretacji wyników dokonano z wykorzystaniem własnych baz danych regionu kontrolnego mtDNA szczegółowo scharakteryzowanych pod względem filogeograficznym. Dodatkowo, dla uaktualnienia filogenezy podhaplogrupy J1b, z której pochodziły haplotypy uzyskane w jednym z analizowanych przypadków oraz dla uprawdopodobnienia interpretacji, przeprowadzono analizę sekwencji pełnych genomów mitochondrialnych.

MATERIAŁ I METODY

Badany materiał biologiczny pochodził z trzech spośród około 225 przypadków, w których Zakład Genetyki Molekularnej i Sądowej w latach 1999-2006 wydał ekspertyzy na podstawie badań sekwencji mtDNA.

Przypadek 1

W sprawie identyfikacyjnej z terenu Małopolski zabezpieczono pięć włosów dowodowych pozbawionych cebulek włosowych (oznaczone numerami 1-5). Materiałem porównawczym były włosy zabezpieczone z odciętej głowy odnalezionej w niedalekiej odległości od miejsca, w którym zabezpieczo-

no włosy dowodowe. W wyniku wstępnej analizy morfologicznej – porównawczej przeprowadzonej z wykorzystaniem komputerowego systemu analizy obrazu stwierdzono, że włosy dowodowe mają swe odpowiedniki morfologiczne wśród włosów porównawczych. Do badań mtDNA zakwalifikowano wszystkie włosy dowodowe oraz trzy włosy porównawcze (oznaczone jako P1-P3).

Przypadek 2

W sprawie zabójstwa mężczyzny dokonanego na terenie Pomorza i Kujaw zabezpieczono cztery włosy dowodowe pozbawione cebulek włosowych, ujawnione na ubraniu ofiary (oznaczone numerami 1-4). Materiałem porównawczym do badań identyfikacyjnych była krew zabezpieczona podczas sekcji zwłok denata (oznaczona jako SER_2).

Przypadek 3

W sprawie identyfikacji zwłok nieznanego mężczyzny, który zginął podczas pożaru na terenie Pomorza i Kujaw, zabezpieczono materiał w postaci krwi denata (oznaczony jako NN). Materiałem porównawczym do badań identyfikacyjnych była krew uzyskana od matki osoby zaginionej (oznaczona jako SER_3).

MOLEKULARNE ANALIZY mtDNA

Wszystkie próbki materiału biologicznego poddano badaniom regionu kontrolnego mtDNA zgodnie z procedurą opisaną we wcześniejszej pracy [6]. Uzyskane sekwencje mtDNA, obejmujące fragment regionu kontrolnego pomiędzy nukleotydem 15999 a 16400 dla HVS I oraz 30-407 dla HVS II porównywano z sekwencją wzorcową CRS [7]. Analizie sekwencji pełnych genomów mitochondrialnych poddano próbki pochodzące z przypadku 3 (NN oraz SER_3). Pełne genomy amplifikowano metodą PCR w jedenastu nakładających się fragmentach, a następnie sekwencjonowano za pomocą 32 opisanych poprzednio wewnętrznych starterów, w warunkach podanych we wcześniejszej pracy [6]. Uzyskane sekwencje mtDNA porównywano z poprawioną sekwencją wzorcową rCRS [8].

FILOGENETYCZNE I STATYSTYCZNE ANALIZY mtDNA

Drzewo filogenetyczne pełnych genomów mitochondrialnych zrekonstruowano manualnie metodą największej oszczędności i ukorzeniono za pomocą sekwencji rCRS [8] jako grupy zewnętrznej,

a następnie sprawdzono z wykorzystaniem oprogramowania Network v. 4.1.1.1 (A. Röhl; © Fluxus Technology; www.fluxus-engineering.com). W rekonstrukcji drzewa wykorzystano sekwencję pełnego genomu mitochondrialnego należącego do podhaplogrupy J1b, opublikowanego przez Palanichamy i wsp. [9].

Interpretację statystyczną wyników analizy mtDNA dla celów sądowych przedstawiano za pomocą ilorazu prawdopodobieństw warunkowych $LR = P(E/H1) / P(E/H2)$, gdzie $P(E/H1)$ oznacza prawdopodobieństwo uzyskania wyniku przy założeniu, że materiał dowodowy pochodzi z linii matczynej reprezentowanej przez próbkę referencyjną, zaś $P(E/H2)$ jest prawdopodobieństwem uzyskania wyniku przy założeniu, że próbka dowodowa pochodzi z innej linii matczynej. W przypadku pełnej zgodności sekwencji pomiędzy próbką dowodową i porównawczą, $P(E/H1) = 1$. W przypadku zaistnienia jednonukleotydowej różnicy pomiędzy sekwencją próbki dowodowej i porównawczej, $P(E/H1)$ ma postać iloczynu częstości mutacji w pozycji sekwencji różniącej oba haplotypy oraz tempa mutacji w regionie kontrolnym w czasie jednego pokolenia. Z uwagi na znaczną wewnątrzsobniczą niestabilność szlaków polipirymidynowych regionów HVS I i HVS II oraz związaną z tym ograniczoną wartość identyfikacyjną mutacji w tych regionach, w analizach statystycznych pomijano polimorfizmy długości notowane w pozycjach 16180-16193 oraz 303-315. Częstość substytucji w pozycjach 16357, 16129 oraz 16319 obliczano na podstawie ilorazu liczby mutacji obserwowanych w tych pozycjach w bazie danych i całkowitej liczby mutacji w tej samej bazie (7800). W obliczeniach wykorzystano własną bazę danych haplotypów HVS I i HVS II ($n=1465$) z obszaru zachodniej Eurazji (z populacji polskiej, rosyjskiej, bośniackiej i słoweńskiej) [6, 10, 11]. Wszystkie haplotypy umieszczone w bazie zaliczono do odpowiednich haplogrup i podhaplogrup mtDNA na podstawie syntetycznego systemu klasyfikacji, uwzględniającego dotychczasowe dane uzyskane metodą PCR-RFLP oraz informacje pochodzące z sekwencjonowania regionów HVS I i HVS II. Uśrednione tempo mutacji w regionie kontrolnym rzędu 0.0043 w czasie jednego pokolenia (lub 0.32 na pozycję sekwencji w czasie miliona lat) przyjęto za Sigurðardóttir i wsp. [12]. $P(E/H2)$ obliczano na podstawie liczby obserwacji haplotypu w bazie danych (w przypadku jednonukleotydowej różnicy pomiędzy próbką dowodową i porównawczą była to suma liczby obserwacji haplotypu dowodowego oraz haplotypu różniącego się od dowodowego jedną mutacją). W sytuacji braku haplotypu w bazie danych, w obliczeniach wykorzystywano

częstość „najmniejszej” (najmłodszej ewolucyjnie, czyli znajdującej się na obrzeżach drzewa filogenetycznego) podhaplogrupy, do której można było zaliczyć uzyskany haplotyp, zgodnie z sugestią Forstera i wsp. [13]. W obliczeniach częstości rzadkich haplotypów na podstawie przeszukiwania bazy danych stosowano przedział ufności 95 %, z wykorzystaniem logarytmu naturalnego częstości, przybliżenia normalnego do rozkładu dwumianowego oraz antylogarytmu [1, 14]. Dla zachowania ostrożności, w obliczeniach LR wykorzystywano maksymalne częstości haplotypów (z górnej granicy przedziału ufności).

WYNIKI I OMÓWIENIE

Haplotypy regionu kontrolnego mtDNA uzyskane dla próbek materiału dowodowego i porównawczego oraz ich przynależność haplogrupową przedstawiono w tabeli I.

Z przedstawionych danych wynika, że pomiędzy sekwencjami mtDNA uzyskanymi z włosa dowodowego oznaczonego numerem 4, a sekwencją otrzymaną dla włosów porównawczych analizowanych w Przypadku 1 wystąpiły różnice w postaci dodatkowej tranzycji T16357C oraz insercji 309.1C. Pozostałe haplotypy uzyskane dla włosów dowodowych wykazywały pełną zgodność wobec haplotypu porównawczego (tabela I). Uzyskane haplotypy należą do podhaplogrupy J1c zdefiniowanej poprzez tranzycje w pozycjach 228 i 14798 [9], najczęstszego podkladu w obrębie haplogrupy J w populacjach słowiańskich oraz w innych populacjach europejskich [6]. Haplotypów należących do J1c z tranzycjami w pozycjach 16086 i 16357 nie zanotowano jednak do tej pory w bazie danych [6, 10, 11]. Na podstawie znajomości topologii drzewa filogenetycznego haplogrupy J1 [6] można stwierdzić, że najmniejsza podgrupa w obrębie J1c, do której można zaliczyć haplotypy uzyskane w omawianym przypadku obejmuje sekwencje z motywem regionu kontrolnego 16069-16126-16261-185-228-295, a jej częstość w bazie danych wynosi $3/1465$, czyli 2.055×10^{-3} (wartość będąca górną granicą przedziału ufności). Wartość LR dla obserwacji pełnej zgodności pomiędzy haplotypami dowodowymi i porównawczymi w przypadku pierwszym wynosi 486. Dla obserwacji jednonukleotydowej różnicy pomiędzy haplotypem uzyskanym z włosa numer 4 a haplotypem porównawczym (tabela I), dokonano oceny częstości substytucji T16357C z wykorzystaniem spektrum mutacji regionu kontrolnego obserwowanego w bazie danych. Zanotowano tutaj dwa przypadki substytucji w pozycji 16357 – na

Tabela I. Haplotypy regionu kontrolnego mtDNA (HVS I i HVS II) oraz ich przynależność haplogrupowa w materiale dowodowym i porównawczym. Mutacje identyfikowano względem sekwencji wzorcowej CRS [7]. Brak zasady przy odpowiedniej pozycji sekwencji oznacza transzycję. Insercje reszt cytozyny oznaczono przyrostkami „.1C”.

Table I. mtDNA control region haplotypes and their subhaplogroup assignment in the evidentiary and reference samples. Mutations are shown indicating positions relative to the mtDNA Cambridge reference sequence (CRS) [7]. The nucleotide positions correspond to transitions; insertions of cytosines are referred to by “.1C” following the nucleotide position.

Przypadek	Nazwa próbki	Podhaplogrupa	Sekwencja regionu kontrolnego mtDNA	
			HVS I	HVS II
1	1	J1c	16069 16086 16126 16261	73 185 228 263 295 315.1C
	2		16069 16086 16126 16261	73 185 228 263 295 315.1C
	3		16069 16086 16126 16261	73 185 228 263 295 315.1C
	4		16069 16086 16126 16261 16357	73 185 228 263 295 309.1C 315.1C
	5		16069 16086 16126 16261	73 185 228 263 295 315.1C
	P1		16069 16086 16126 16261	73 185 228 263 295 315.1C
	P2		16069 16086 16126 16261	73 185 228 263 295 315.1C
	P3		16069 16086 16126 16261	73 185 228 263 295 315.1C
2	1	H5	16129 16304	263 309.1C 315.1C
	2		16129 16304	263 309.1C 315.1C
	3		16304	263 309.1C 315.1C
	4		16304	263 309.1C 315.1C
	SER_2		16129 16304	263 309.1C 315.1C
3	NN	J1b2	16069 16126 16145 16222 16234 16261 16319	64 73 152 263 271 295 315.1C
	SER_3		16069 16126 16145 16222 16234 16261	64 73 152 263 271 295 315.1C

tle haplogrup R* oraz U5a [10, 11]. Warto jednak zwrócić uwagę, że pozycja 16357 uważana jest za „gorące miejsce” mutacji w regionie kontrolnym mtDNA, co uzyskało potwierdzenie w co najmniej dwóch niezależnych analizach filogenetycznych [15, 16]. Oznacza to, że substytucja T16357C może być obserwowana na tle różnych haplogrup [16]. Prawdopodobieństwo pojawienia się mutacji w pozycji 16357 wynosi zatem $2/7800$, czyli 2.5×10^{-4} . Przyjmując uśrednione tempo mutacji w czasie jednego pokolenia rzędu 0.0043, P (E/H1) dla obserwacji różnicy jednonukleotydowej pomiędzy próbką dowodową i porównawczą wynosi 1×10^{-6} , a wartość LR w omawianym przypadku równa jest tylko 4.9×10^{-4} . Innymi słowy, przy zaistnieniu różnicy jednonukleotydowej w pozycji 16357, pochodzenie próbki dowodowej z innej linii matczynej niż próbka porównawcza jest ok. 2040 razy bardziej prawdopodobne niż pochodzenie tych próbek z tej samej linii.

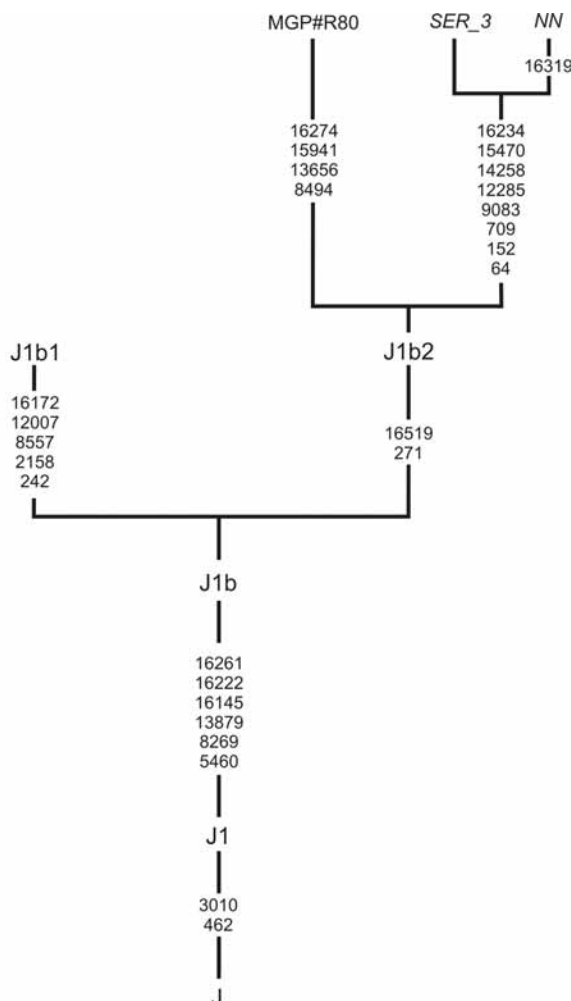
Zakładając, że mutacja T16304C ma charakter monofiletyczny w obrębie haplogrupy H [6] można uznać, że haplotypy obserwowane w Przypadku 2 należą do podkladu H5, zdefiniowanego poprzez transzycje w pozycjach 456 i 16304 [17]. H-16304 jest drugim co do częstości motywem wśród haplotypów HVS I należących do haplogrupy H w Euro-

pie (ok. 4 %), choć występuje również na Bliskim Wschodzie, z częstościami rzędu 1-3 % [18]. W bazie danych odnaleziono jednak tylko jeden haplotyp z motywem H-16129-16304 (częstość profilu w bazie 6.8×10^{-4}). Mimo to pozycja 16129 uważana jest za jedno z „najszybszych” miejsc w regionie kontrolnym [15, 16, 19]. Przykładowo, w wyniku populacyjnych badań mtDNA Polaków i Rosjan stwierdzono równoległe występowanie transzycji G16129A w haplotypach należących do haplogrup H, U5, T*, T1, I, W oraz M* [10]. Na podstawie liczby obserwacji tej mutacji w bazie danych ($22/7800$) określono częstość substytucji w pozycji 16129 na 2.82×10^{-3} . Wartości LR dla obserwacji pełnej zgodności (próbki oznaczone jako 1 i 2, tabela I) oraz jednonukleotydowej różnicy (próbki oznaczone numerami 3 i 4, tabela I) pomiędzy haplotypami dowodowym i porównawczym wynoszą odpowiednio 1470 oraz 1.047×10^{-4} , a zatem w tym drugim przypadku dowód genetyczny wskazuje aż na ok. 9551-krotnie większe prawdopodobieństwo pochodzenia haplotypów z różnych linii matczynej.

W wyniku sekwencjonowania regionów HVS I i HVS II dwóch próbek badanych w ramach Przypadku 3 stwierdzono, że uzyskane haplotypy, z motywem 16069-16126-16145-16222-16261-271 należą najprawdopodobniej do niezwykle rzadkiego

Ryc. 1. Drzewo filogenetyczne podhaplogrupy J1b utworzone metodą największej oszczędności, zrekonstruowane z wykorzystaniem sekwencji pełnych genomów mitochondrialnych i ukorzenione za pomocą poprawionej sekwencji referencyjnej (rCRS) [8] jako grupy zewnętrznej. Sekwencje uzyskane w ramach niniejszej pracy opisano jako „NN” i „SER_3”. Sekwencję mtDNA uzyskaną przez Palanichamy i wsp. [9] z populacji Indii oznaczono jako „MGP#R80”. Gałęzie drzewa oznaczono za pomocą pozycji nukleotydowych sekwencji mtDNA, według numeracji zastosowanej dla rCRS [8]. Brak zasady obok pozycji sekwencji oznacza transycję. Z analizy wyłączono polimorfizmy długości towarzyszące pozycjom 309 i 315 regionu HVS II.

Fig. 1. The most parsimonious tree of the complete subhaplogroup J1b mtDNA sequences, rooted with the revised Cambridge reference sequence (rCRS) [8] as an outgroup. The tree includes 3 mtDNAs, out of which two are novel (“NN” and “SER_3”) and one (from the Indian population, designated as “MGP#R80”) was reported by Palanichamy et al. [9]. Mutations are shown on the branches and are transitions. The hypervariable nucleotides (the indels) at nucleotides 309 and 315 in HVS II were excluded.



w populacjach europejskich podkladu w obrębie J1b. Przepuszczenie takie oparte jest na obserwacji mutacji C271T w HVS II, przy jednoczesnym braku transycji w pozycji 242, diagnostycznej dla podhaplogrupy J1b1 [9]. Do chwili obecnej opublikowano tylko jedną pełną sekwencję genomu mitochondrialnego należącego do J1b z transycją w pozycji 271 (z populacji Indii) [9]. W celu dokładnego określenia statusu filogenetycznego haplotypów uzyskanych w Przypadku 3, przeprowadzono analizę sekwencji pełnych genomów mitochondrialnych próbek oznaczonych jako „NN” oraz „SER_3”, po czym dokonano rekonstrukcji filogenezy podhaplogrupy J1b w wykorzystaniu nowych danych (rycina 1). Analiza sekwencji wykazała, że pełne genomy mitochondrialne próbki dowodowej i porównawczej różnią się tylko w pozycji 16319.

Z drzewa filogenetycznego przedstawionego na rycinie 1 wynika, że genomy analizowane w Przypadku 3 należą do podhaplogrupy J1b2, zdefiniowanej poprzez transycje w pozycjach 271 i 16519. Należy zwrócić uwagę, że taka definicja może mieć charakter tymczasowy, gdyż włączenie do analizy kolejnych danych o sekwencjach pełnych genomów może zmienić topologię drzewa filogenetycznego. W szczególności warto zauważyć, że pozycja 16519 uważana jest za kolejne „gorące miejsce” mutacji. Częstość podhaplogrupy J1b2 jest w chwili obecnej trudna do ustalenia, gdyż wiele zespołów badawczych nie uwzględnia w swych badaniach populacyjnych sekwencji regionu HVS II. We wcześniejszych badaniach własnych zidentyfikowano haplotypy pozbawione transycji w pozycji 242, a jednocześnie posiadające mutację C271T (lub jej pozbawione) w populacjach Bliskiego Wschodu (Persowie z Iranu) oraz południowo-zachodniej i południowo-wschodniej Syberii (Altajczycy, Buriaci) [6]. Najwyższe częstości J1b2 zanotowano u Persów z Iranu (ok. 3.7 %) [6], jednak w populacji tej wszystkie haplotypy należące do J1b2 były pozbawione transycji C1622T, zanotowanej w próbkach badanych w ramach Przypadku 3 (tabela I). Biorąc pod uwagę całokształt przytoczonych danych filogeograficznych należy się spodziewać, że podhaplogrupa J1b2 osiąga najwyższe częstości i wewnętrzne zróżnicowanie w populacjach północno-zachodniego Iranu. Zważywszy jednak na niewielką liczebność baz danych sekwencji HVS I i HVS II z tego regionu, ich zastosowanie do obliczeń częstości haplotypów w sprawach sądowych nie wydaje się stosowne. Najmniejszym podkładem, do którego można zaliczyć haplotypy uzyskane w ramach Przypadku 3 jest zatem podhaplogrupa J1b. Częstość tego podkladu w bazie danych wynosi 17/1465, czyli 1.2×10^{-2} . Mutacja G16319A należy do

„gorących miejsc” w regionie kontrolnym, a jej częstość oszacowano na 22/7800, czyli 2.8×10^{-3} . LR dla obserwacji poczynionych w ramach Przypadku 3 wynosi 0.001, czyli 1000-krotnie faworyzuje hipotezę wykluczającą pochodzenie próbek NN i SER_3 z tej samej linii matczynej. W przypadku zaistnienia pełnej zgodności pomiędzy haplotypami dowodowym i porównawczym, LR wynosiłby 83. W omawianym przypadku przeprowadzono dodatkowo badania piętnastu jądrowych mikrosatelitów. Wyniki tych badań nie pozwoliły na wykluczenie macierzyństwa (MI = 861 422, dane nie prezentowane).

PODSUMOWANIE

Przedstawione dane ilustrują przydatność podejścia filogeograficznego w interpretacji wyników sekwencjonowania mtDNA dla celów sądowych. Podejście takie jest szczególnie użyteczne w ocenach częstości bardzo rzadkich haplotypów, których nie obserwowano do tej pory w bazach danych. W takich przypadkach ustalanie częstości najmniejszych podkladów obejmujących rzadkie haplotypy może stanowić alternatywę dla oceny maksymalnego, teoretycznego progu częstości danego haplotypu w populacji z wykorzystaniem tzw. przedziału ufności z proporcji zero [1]. Jak wykazano w niniejszej pracy, w niektórych przypadkach dla dokładnego ustalenia przynależności haplogrupowej uzyskanego haplotypu można się uciec do analizy sekwencji pełnych genomów, pełnych sekwencji regionu kodującego lub chociażby wybranych polimorfizmów regionu kodującego.

We wszystkich analizowanych przypadkach niezgodności jednonukleotydowych pomiędzy haplotypami dowodowymi i porównawczymi uzyskiwano bardzo niskie wartości LR, faworyzujące hipotezę wykluczającą pochodzenie próbek z tej samej linii matczynej. Nie upoważnia to jednak do sformułowania ogólnej reguły interpretacji, według której obserwacja różnicy jednonukleotydowej prowadzi do jednoznacznie wykluczającej sentencji opinii. Warto zwrócić uwagę, że wszystkie pozycje sekwencji, w których zaobserwowano różnice należą do „szybkich” (16357) lub „bardzo szybkich” (16319, 16129) miejsc w regionie kontrolnym. Substytucje w pozycjach 16357 i 16319 powstają najprawdopodobniej na nici H na skutek przejściowego, nieprawidłowego połączenia startera i matrycy podczas replikacji mtDNA przez mitochondrialną polimerazę gamma [16, 19]. Warto zwrócić uwagę, że ten sam mechanizm molekularny wyjaśnia powstawanie polimorfizmu długości w superzmiennych regionach polipirymidynowych HVS I i HVS II,

a zmienność tego rodzaju zdaje się nie posiadać większego znaczenia identyfikacyjnego. Ponadto nie można wykluczyć, że przedstawione przypadki homoplazmatycznych różnic są w istocie przypadkami heteroplazmii, niemożliwych do wykrycia z uwagi na niską czułość metody bezpośredniego sekwencjonowania produktów PCR.

Dokonane w ramach pracy obliczenia częstości mutacji właściwych dla poszczególnych pozycji regionu kontrolnego wymagają przyjęcia pewnych wstępnych założeń, które mogą być poddane dyskusji. Przykładowo, do tej pory nie ustalono w sposób definitywny przyczyn rozbieżności między częstościami mutacji w regionie kontrolnym obliczonymi na podstawie rekonstrukcji filogenetycznych oraz w wyniku kalkulacji bezpośrednich, poprzez badania mtDNA w rodzinach [20]. Szczegółowe omówienie tego zagadnienia przekraczałoby ramy niniejszej pracy. Należy jednak zaznaczyć, że dla potrzeb interpretacji wyników w sprawach sądowych bardziej stosowne wydaje się przyjęcie założeń co do średniego tempa mutacji wyznaczonego na podstawie badań bezpośrednich. W badaniach takich preferencyjnie wykrywane są mutacje w „gorących miejscach”, które w rekonstrukcjach filogenetycznych mogą pozostać niezauważone bądź niedoszacowane z uwagi na możliwe przypadki rewersji, tj. powrotu danej pozycji sekwencji do stanu wyjściowego [20]. Dla potrzeb niniejszej pracy przyjęto wartość średniego tempa mutacji wyznaczonego w wyniku jednego z szerszych badań rodzinnych [12]. Wartość ta jest porównywalna z wynikami uzyskanymi w innych badaniach bezpośrednich [12]. Obliczenia częstości mutacji w konkretnych pozycjach regionu kontrolnego wymagają również wykorzystania baz danych o odpowiedniej liczebności oraz rozpoznania, czy dana mutacja nie występuje preferencyjnie w obrębie określonej haplogrupy. Substytucje T16357C, G16319A oraz G16129A zanotowano w obrębie wielu różnych haplogrup, toteż w obliczeniach wykorzystano liczbę mutacji obserwowanych w tych pozycjach w stosunku do całkowitej liczby mutacji w bazie danych. W przypadkach różnic jednonukleotydowych podejście takie ma charakter bardzo konserwatywny i prowadzi do znaczącego obniżenia wartości LR, przez co silnie faworyzuje hipotezę wykluczającą pochodzenie próbek z jednej linii matczynej.

Należy się spodziewać, że wzrost liczebności baz danych mtDNA oraz ich bardziej szczegółowa charakterystyka filogenetyczna pozwolą na dokonanie bardziej precyzyjnych obliczeń częstości mutacji w poszczególnych pozycjach regionu kontrolnego, a także bardziej wiarygodnych ocen częstości haplotypów. W szczególności wzrastająca ilość da-

nych populacyjnych dotyczących sekwencji pełnych genomów przyczyni się do jeszcze szerszego wykorzystania podejścia filogeograficznego w genetyce sądowej, przede wszystkim w trudniejszych przypadkach identyfikacyjnych.

PIŚMIENNICTWO

- Holland M. M., Parsons T. J.: Mitochondrial DNA sequence analysis – validation and use for forensic casework. *Forensic Sci. Rev.* 1999, 11, 21.
- Salas A., Bandelt H. J., Macaulay V., Richards M.: Phylogeographic investigations: the role of trees in forensic genetics. *Forensic Sci. Int.* 2006, w druku.
- Coble M. D., Vallone P. M., Just R. S., Diegoli T. M., Smith B. C., Parsons T. J.: Effective strategies for forensic analysis in the mitochondrial DNA coding region. *Int. J. Legal Med.* 2006, 120, 27-32.
- Brandstätter A., Salas A., Niederstätter H., Gassner C., Carracedo A., Parson W.: Dissection of mitochondrial superhaplogroup H using coding region SNPs. *Electrophoresis* 2006, 27, w druku.
- Torroni A., Achilli A., Macaulay V., Richards M., Bandelt H. J.: Harvesting the fruit of the human mtDNA tree. *Trends Genet.* 2006, 22, 339-345.
- Grzybowski T.: Zróżnicowanie sekwencji mitochondrialnego DNA w populacjach ludzkich północnej Eurazji. Wybrane aspekty filogeograficzne. Uniwersytet Mikołaja Kopernika w Toruniu, Collegium Medicum im. Ludwika Rydygiera, Bydgoszcz, 2006.
- Anderson S., Bankier A. T., Borel B. G., de Bruijn M. H., Coulson A. R., Drouin J., Eperon I. C., Nierlich D. P., Roe B. A., Sanger F. i wsp.: Sequence and organization of the human mitochondrial genome. *Nature* 1981, 290, 457-465.
- Andrews R. M., Kubacka I., Chinnery P. F., Lightowlers R. N., Turnbull D. M., Howell N.: Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat. Genet.* 1999, 23, 147.
- Palanichamy M. G., Sun C., Agrawal S., Bandelt H. J., Kong Q. P., Khan F., Wang C. Y., Chaudhuri T. K., Palla V., Zhang Y. P.: Phylogeny of mitochondrial DNA macrohaplogroup N in India, based on complete sequencing: implications for the peopling of South Asia. *Am. J. Hum. Genet.* 2004, 75, 966-978.
- Malyarchuk B. A., Grzybowski T., Derenko M. V., Czarny J., Woźniak M., Miścicka-Śliwka D.: Mitochondrial DNA variability in Poles and Russians. *Ann. Hum. Genet.* 2002, 66, 261-283.
- Malyarchuk B. A., Grzybowski T., Derenko M. V., Czarny J., Drobnic K., Miścicka-Śliwka D.: Mitochondrial DNA variability in Bosnians and Slovenians. *Ann. Hum. Genet.* 2003, 67, 412-425.
- Sigurðardóttir S., Helgason A., Gulcher J. R., Stefansson K., Donnelly P.: The mutation rate in the human mtDNA control region. *Am. J. Hum. Genet.* 2000, 66, 1599-1609.
- Forster P., Cali F., Rohl A., Metspalu E., D'Anna R., Mirisola M., De Leo G., Flugy A., Salerno A., Ayala G. i wsp.: Continental and subcontinental distributions of mtDNA control region types. *Int. J. Legal Med.* 2002, 116, 99-108.
- National Research Council: The Evaluation of Forensic DNA Evidence. National Academy Press, Washington D.C. 1996.
- Bandelt H. J., Quintana-Murci L., Salas A., Macaulay V.: The fingerprint of phantom mutations in mitochondrial DNA data. *Am. J. Hum. Genet.* 2002, 71, 1150-1160.
- Malyarchuk B. A., Rogozin I. B.: Mutagenesis by transient misalignment in the human mitochondrial DNA control region. *Ann. Hum. Genet.* 2004, 68, 324-339.
- Loogväli E. L., Roostalu U., Malyarchuk B. A., Derenko M. V., Kivisild T., Metspalu E., Tambets K., Reidla M., Tolk H. V., Parik J. i wsp.: Disuniting uniformity: a pied cladistic canvas of mtDNA haplogroup H in Eurasia. *Mol. Biol. Evol.* 2004, 21, 2012-2021.
- Richards M., Macaulay V., Hickey E., Vega E., Sykes B., Guida V., Rengo C., Sellito D., Cruciani F., Kivisild T. i wsp.: Tracing European founder lineages in the Near Eastern mtDNA pool. *Am. J. Hum. Genet.* 2000, 67, 1251-1276.
- Malyarchuk B. A., Rogozin I. B., Berikov V. B., Derenko M. V.: Analysis of phylogenetically reconstructed mutational spectra in human mitochondrial DNA control region. *Hum. Genet.* 2002, 111, 46-53.
- Howell N., Smejkal C. B., Mackey D. A., Chinnery P. F., Turnbull D. M., Herrnstadt C.: The pedigree rate of sequence divergence in the human mitochondrial genome: there is a difference between phylogenetic and pedigree rates. *Am. J. Hum. Genet.* 2003, 72, 659-670.

Adres do korespondencji:

Dr Tomasz Grzybowski
Katedra Medycyny Sądowej
Zakład Genetyki Molekularnej i Sądowej
Collegium Medicum UMK
ul. M. Curie-Skłodowskiej 9
85-094 Bydgoszcz
Tel: +48 52 585 3556
e-mail: tgrzyb@cm.umk.pl