



Wysoko wiarygodne metody identyfikacji osób

WIKTOR OLCHOWIK

Wojskowa Akademia Techniczna, Wydział Elektroniki,
00-908 Warszawa, ul. gen. S. Kaliskiego 2, wolchowik@wat.edu.pl

Streszczenie. Spośród znanych technik biometrycznych identyfikacji osób zdecydowanie najwyższy poziom wiarygodności uzyskuje się poprzez porównanie profili DNA. W artykule opisano metodykę identyfikacji opartej o profile STR16. Opracowano oryginalne algorytmy identyfikacji osób i oceny pokrewieństwa bazujące na dwóch uniwersalnych wskaźnikach oraz dokonano statystycznej analizy wiarygodności poszczególnych hipotez. Przetwarzanie danych zostało zintegrowane z innymi biometrycznymi technikami identyfikacji.

Słowa kluczowe: identyfikacja osób, techniki biometryczne, profil DNA

1. Wstęp

Spośród znanych technik biometrycznych identyfikacji osób zdecydowanie najwyższy poziom wiarygodności uzyskuje się poprzez porównanie profili DNA. Wadą tej metody jest długi czas wykonywania badania, które najczęściej obejmuje izolację, pomiar stężenia i amplifikację genomowego DNA oraz rozdział elektroforetyczny, czyli elektroforezę kapilarną. Cały proces badania i identyfikacji trwa więc co najmniej kilka godzin.

2. Wybór metody badania

Istnieje szereg typów markerów stosowanych w badaniach DNA do identyfikacji osób, np. VNTR (*Variable Number Tandem Repeats* — zmienna liczba tandemowych powtórzeń), RFLP (*Restriction Fragment Length Polymorphism* — polimorfizm

długości fragmentów restrykcyjnych), STR (*Short Tandem Repeats* — krótkie powtórzenia tandemowe), SNP (*Single-nucleotide polymorphism* — polimorfizm pojedynczych nukleotydów) i inne. Ze względu na koszt oraz popularność badań oraz efektywność identyfikacji wybrano metodę STR. W jej ramach wybrano Autosomalne STR, ponieważ alternatywne warianty Y STR lub X STR odnoszą się tylko do jednej płci. Oprócz wyboru typu profilu istotne znaczenie ma ilość i zestaw markerów. Kierując się głównie efektywnością i kosztami identyfikacji, przyjęto powszechnie stosowany zestaw piętnastu markerów autosomalnych STR + marker określający płeć (D8S1179, D21S11, D7S820, CSF1PO, D3S1358, THO1, D13S317, D16S539, D2S1338, D19S433, vWA, TPOX, D18S51, AMEL, D5S818, FGA). Ponadto ten format profilu DNA jest zgodny z określonymi przez ENFSI (*European Network of Forensic Science Institutes*) oraz EDNAP (*European DNA Profiling Group*).

Metoda STR polega na badaniu liczby powtórzeń łańcuchów nukleotydowych w pewnych, ściśle określonych lokalizacjach chromosomów. Ze względu na duży polimorfizm markerów (czyli zmienność osobniczą) profil genetyczny STR jest praktycznie niepowtarzalnym ID osoby. Ponadto tzw. mendelowski sposób dziedziczenia powoduje, że można określić prawdopodobieństwo podobieństwa profilu w ramach osób spokrewnionych. Możliwa jest identyfikacja osoby, identyfikacja par rodzic–dziecko, określanie prawdopodobieństwa pokrewieństwa, określanie podobieństwa lub dystansu genetycznego.

3. Identyfikacja osoby

Z wyżej wymienionych cech profilu STR wynika, że dla DNA pochodzącego od tej samej osoby allele dla wszystkich markerów profilu DNA muszą być zgodne.

Algorytm identyfikacji dokonuje więc porównania par alleli dla wszystkich markerów i w przypadku pełnej zgodności określa, że osoba została jednoznacznie zidentyfikowana.

O unikalnej właściwości identyfikacji decyduje bardzo małe prawdopodobieństwo powtórzenia profilu genetycznego w populacji. Prawdopodobieństwo to oblicza się według algorytmu:

Dla każdego markera oblicza się prawdopodobieństwo powtórzenia pary alleli w danej populacji (wzór 1 i 2):

$$\text{Dla } a_1 \neq a_2 \quad p_1 = 2p(a_1)p(a_2), \quad (1)$$

$$\text{Dla } a_1 = a_2 \quad p_1 = p(a_1)p(a_2), \quad (2)$$

gdzie: p_1 — prawdopodobieństwo powtórzenia markera 1 w danej populacji;
 $p(a_1)$ — prawdopodobieństwo wystąpienia allela a_1 ;
 $p(a_2)$ — prawdopodobieństwo wystąpienia allela a_2 .

Obliczamy prawdopodobieństwa powtórzenia alleli dla każdego markera oddzielnie. W załączonym przykładzie prawdopodobieństwa dla markerów THO1 i D13S317 obliczamy ze wzoru (1), a dla CSF1PO według wzoru (2).

TABELA 1

Przykładowe trzy allele profilu genetycznego

Marker	THO1	D13S317	CSF1PO
allel a_1	7	12	11
allel a_2	9	13	11

Następnie oblicza się prawdopodobieństwo powtórzenia całego profilu genetycznego w danej populacji jako iloczyn prawdopodobieństw powtórzenia wszystkich piętnastu markerów.

$$p = \prod_{i=1}^{15} p_i \quad (3)$$

gdzie: p — prawdopodobieństwo powtórzenia profilu genetycznego;

p_i — prawdopodobieństwo powtórzenia i -tego markera.

Dla piętnastu markerów STR przykładowego profilu prawdopodobieństwo powtórzenia alleli przedstawiono w tabeli 2.

TABELA 2

Przykładowy profil genetyczny STR 15

Marker	allel a_1	allel a_2	Prawdopodobieństwo p_i
D8S1179	14	14	0,0433
D21S11	31,2	32,2	0,0195
D7S820	8	12	0,0397
CSF1PO	11	11	0,052
D3S1358	16	18	0,0642
THO1	7	9	0,0516
D13S317	12	13	0,041
D16S539	12	14	0,0285
D2S1338	17	17	0,0324
D19S433	13	15	0,0653
vWA	18	18	0,0635
TPOX	8	9	0,1
D18S51	14	14	0,0282
D5S818	11	12	0,2467
FGA	23	23	0,0154

Prawdopodobieństwo powtórzenia przykładowego profilu w populacji polskiej [1] wyniosłoby $P_{(STR-PL)} = 9,67 \times 10^{-21}$. Przedstawiona powyżej metodyka umożliwia więc identyfikację tożsamości na bardzo wysokim poziomie wiarygodności.

4. Identyfikacja par rodzic–dziecko

Powszechnie wykorzystuje się metodę STR do ustalenia ojcostwa. W szerszym ujęciu umożliwia to identyfikację par rodzic–dziecko. W profilu rodzica i dziecka dla każdego markera powinna wystąpić zgodność co najmniej jednego allela.

W tabeli 3 przedstawiono przykład dokładnej identyfikacji pary rodzic–dziecko, ponieważ dla każdego markera istnieje co najmniej jeden identyczny allel (zaznaczony wytłuszczonym drukiem).

Algorytm identyfikacji dokonuje więc porównania par alleli dla wszystkich markerów i w przypadku zgodności co najmniej jednego allela dla wszystkich markerów określa, że para rodzic–dziecko została jednoznacznie zidentyfikowana. Liczba markerów, dla których oba allele są różne, musi więc być zerowa.

TABELA 3
Przykładowy profil genetyczny STR 15 rodzic–dziecko

Marker	Profil 1		Profil 2	
D8S1179	14	14	11	14
D21S11	31,2	32,2	29	32,2
D7S820	8	12	8	8
CSF1PO	11	11	11	12
D3S1358	16	18	15	16
THO1	7	9	7	8
D13S317	12	13	11	13
D16S539	12	14	12	14
D2S1338	17	17	17	25
D19S433	13	15	13	16
vWA	18	18	18	18
TPOX	8	9	8	9
D18S51	14	14	14	14
D5S818	11	12	11	12
FGA	23	23	20	23

Wiarygodność identyfikacji można określić prawdopodobieństwem wystąpienia innej niespokrewnionej osoby, która pasowałaby do pary rodzic–dziecko w danej populacji. Prawdopodobieństwo to można obliczyć jako iloczyn prawdopodobieństw dla wszystkich badanych markerów wystąpienia co najmniej jednego z pary alleli w populacji. Prawdopodobieństwo to jest rzędu 10^{-8} .

5. Identyfikacja pokrewieństwa

Identyfikacja pokrewieństwa (poza ustalaniem rodzicielstwa) na podstawie piętnastu markerów STR jest oparta tylko o wskaźniki statystyczne, teoretycznie nie musi być spełniony żaden warunek zgodności alleli. Jednak dla celów identyfikacji osób podejrzanych o terroryzm wydaje się być wystarczające określenie stopnia zgodności danego profilu genetycznego z profilami znajdującymi się w bazie oraz szacunkowe wskazanie stopnia pokrewieństwa.

Mechanizm dziedziczenia powoduje, że w pierwszym pokoleniu potomek posiada dokładnie po jednym chromosomie rodzica. W przypadku pary rodzic–dziecko dla wszystkich piętnastu markerów STR musi być więc zgodność co najmniej jednego allela. W przypadku różnicy dwóch pokoleń lub rodzeństwa taka zgodność nie musi wystąpić, jednak prawdopodobieństwo zgodności jednego allela jest bardzo duże. Im dalsze pokrewieństwo, tym prawdopodobieństwo wystąpienia identycznych alleli dla danego markera jako efektu dziedziczenia jest mniejsze. W przypadku gdy jest mniejsze niż prawdopodobieństwo powtórzenia alleli wynikające ze statystycznych rozkładów częstości wystąpienia poszczególnych alleli, hipoteza o pokrewieństwie staje się mało wiarygodna.

Liczba markerów zawierających pary różnych alleli jest więc pewnym wskaźnikiem pokrewieństwa. W przypadku zgodności jednego allela a_1 dodatkowo przyjęto statystyczną miarę bliskości genetycznej w postaci (4).

$$b_i = -\ln(p(a_i)), \quad (4)$$

gdzie: b_i — miara bliskości genetycznej dla i -tego markera;

$p(a_1)$ — prawdopodobieństwo wystąpienia allela a_1 w populacji.

W przypadku zgodności pary alleli a_1 i a_2 przyjęto miarę bliskości genetycznej w postaci (4).

$$b_i = \sqrt{\ln^2(p(a_1)) + \ln^2(p(a_2))}. \quad (5)$$

W przypadku braku zgodności alleli przyjmuje się $b_i = 0$.

Dla piętnastu markerów miara bliskości będzie obliczana ze wzoru (6) jako średnia arytmetyczna miar dla poszczególnych markerów.

$$b = \frac{1}{15} \sum^{15} b \quad (6)$$

Powyższa metoda umożliwia określenie miar bliskości genetycznej b pomiędzy dowolną parą osób. Przy czym bardziej komunikatywny jest względny wskaźnik procentowy B odniesiony do wartości współczynnika b_w jako bezwzględnej bliskości genetycznej do samego siebie, czyli bliskości genetycznej własnej. Przedstawiona metoda może pozwolić na wskazanie pokrewieństwa pomiędzy badaną osobą a profilami genetycznymi znajdującymi się w bazie danych.

6. Określenie indywidualnego dystansu genetycznego do populacji

Dystans genetyczny jest miarą różnic genetycznych między populacjami. Wynika on z matrycy podobieństwa genetycznego między dwiema lub większą liczbą populacji. Wartość liczbową dystansu dla pojedynczego markera najczęściej określa się ze wzoru (4).

$$D = -\ln \left(\frac{\sum p_{x(i)} p_{y(i)}}{\sqrt{\sum p_{x(i)}^2 \sum p_{y(i)}^2}} \right), \quad (7)$$

gdzie: $p_x(i)$ — prawdopodobieństwo wystąpienia i -tego allela w populacji X;
 $p_y(i)$ — prawdopodobieństwo wystąpienia i -tego allela w populacji Y;
 i — numer allela.

Dla wielu markerów dystans genetyczny określa się jako średnią arytmetyczną.

Natomiast bazując na wzorze (7) określono metodę obliczania indywidualnego dystansu genetycznego pomiędzy profilem pojedynczej osoby a populacją:

- w przypadku gdy dla danego markera występują dwa różne allele w profilu osoby, podstawiamy $p_x(i) = 0,5$ dla alleli występujących w profilu oraz $p_x(i) = 0$ dla pozostałych alleli;
- w przypadku gdy dla danego markera występuje tylko jeden allel w profilu osoby, podstawiam $p_x(i) = 1$ dla danego allela oraz $p_x(i) = 0$ dla pozostałych alleli;
- dla wielu markerów indywidualny dystans genetyczny określa się jako średnią arytmetyczną.

Przy czym bezwzględne wartości tak obliczonego dystansu genetycznego bardziej zależą od globalnej częstości występowania alleli zawartych w profilu danej osoby niż od rozkładu prawdopodobieństw dla populacji. Jako wskaźnik przynależności do danej populacji można zastosować ewentualnie względny dystans genetyczny

odniesiony do populacji bazowej. Przy określaniu dystansu genetycznego bardzo istotną kwestią jest właściwy rozkład częstości występowania alleli dla poszczególnych markerów określonych w trakcie badań na skalę masową.

7. Weryfikacja metody i opracowanie tabeli komunikatów

W wyniku powyższych analiz określono dwie wielkości charakteryzujące dany profil w korelacji z dowolnym innym profilem z bazy danych oraz wielkość wskazującą najbliższą populację.

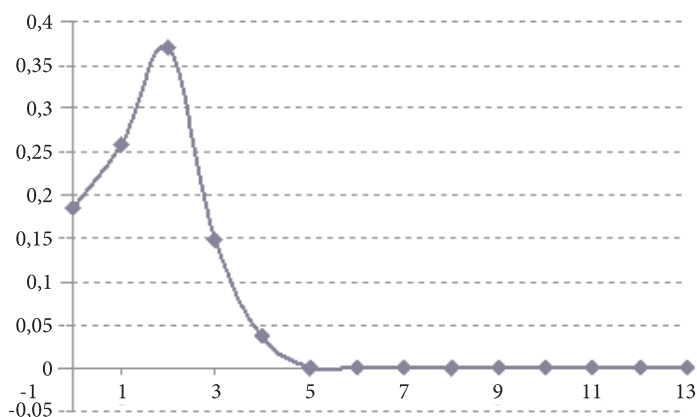
Wielkości te zdefiniowano jako:

- indywidualna względna bliskość genetyczna B , czyli miara zgodności dwóch profili,
- liczba niezgodnych markerów w analizowanych dwóch profilach L ,
- dystans genetyczny do populacji określony względem wybranej populacji bazowej E .

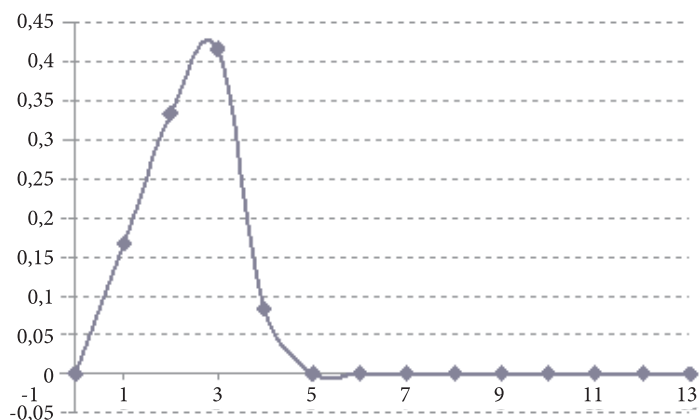
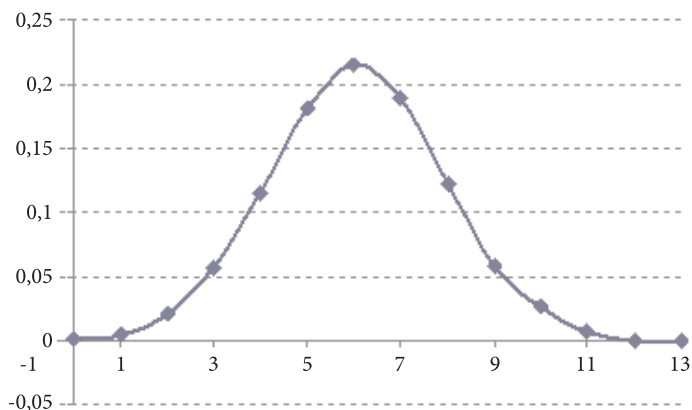
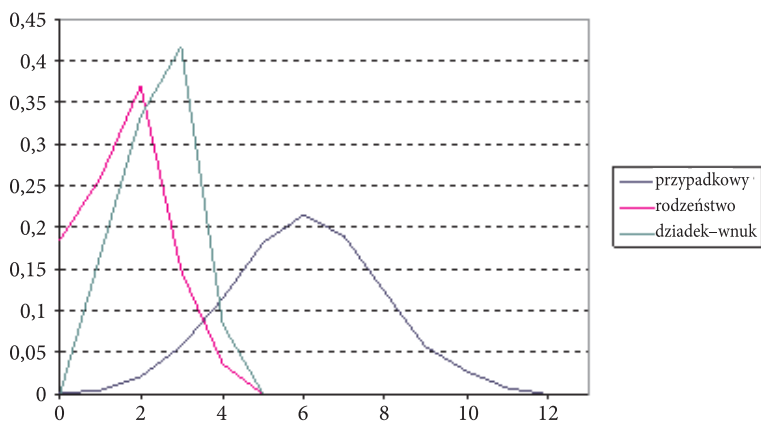
Istotna jest kwestia ustalenia informacji, które można uzyskać na bazie obliczonych wartości (czyli ewentualne pokrewieństwo) oraz oszacować ich wiarygodność.

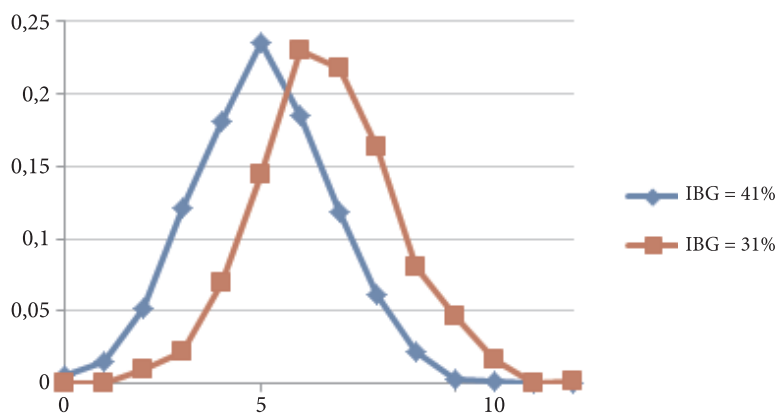
Otrzymanie miarodajnych wyników wymagało dokonania analizy współzależności wzajemnej profili z uwzględnieniem pokrewieństwa. Przetworzone dane pogrupowano w zależności od pokrewieństwa i otrzymano następujące unormowane rozkłady:

- rozkład L dla rodzeństwa (rys. 1),
- rozkład L dla relacji dziadek–wnuk (rys. 2),
- rozkład L dla przypadkowej relacji (rys. 3),
- porównawczy zestaw rozkładów L dla powyższych relacji (rys. 4),
- rozkłady L dla przypadkowych relacji przy zróżnicowanych średnich B (rys. 5).



Rys. 1. Rozkład L dla rodzeństwa

Rys. 2. Rozkład L dla relacji dziadek-wnukRys. 3. Rozkład L dla przypadkowej relacjiRys. 4. Porównawczy zestaw rozkładów L

Rys. 5. Rozkłady L dla przypadkowych relacji, przy różnych średnich B

Dla wymienionych relacji statystycznie najmniejsza liczba niezgodnych markerów występuje w przypadku rodzeństwa, najczęściej 0-2, w relacji dziadek-wnuk najczęściej 2-3, a dla przypadkowych relacji najczęściej 5-8, przy czym występują także wartości L od 1 do 3. Oprócz powyżej wymienionych występuje także relacja rodzic-dziecko oraz identyfikacja osoby, ale wówczas $L = 0$ i przedstawianie na wykresie mija się z celem. Przebadano również powyższe relacje pod względem wartości współczynnika indywidualnej bliskości genetycznej. Statystyczne rozkłady wartości B opracowano oddzielnie dla każdej relacji oraz wartości L . Opracowane dane przedstawiono w tabeli 4. Jako przedział przyjęto wartości z zakresu: (średnia B) \pm (odchylenie standardowe B). Wartości B zostały przedstawione w %. Ograniczono się do analizy dla $L = \{1, 2, 3\}$, ponieważ dla wyższych wartości L częstość relacji losowych jest dominująca.

TABELA 4

Relacje pokrewieństwa i bliskości genetycznej

L	Identyfikacja osoby		Rodzic-dziecko		Rodzeństwo		Dziadek-wnuk		Relacja losowa	
	B -min	B -max	B -min	B -max	B -min	B -max	B -min	B -max	B -min	B -max
0	100%	100%	62,9%	84,2%	68,5%	93,1%				
1					63,5%	79,9%	59,6%	67,5%	53,3%	78,0%
2					55,1%	78,3%	52,4%	71,9%	44,4%	73,4%
3					55,1%	72,5%	45,6%	59,9%	37,5%	67,6%

Tabela 4 stanowiła podstawę do opracowania tabeli komunikatów, w której uwzględniono fakt, że powyższe dane powstały na bazie stosunkowo niewielkiej liczby relacji. Tylko liczba relacji losowych i rodzic-dziecko jest relatywnie duża.

Ponadto wartość B zależna jest w istotny sposób od rozkładu prawdopodobieństwa alleli, czyli od przyjętej populacji bazowej. Istnieje też duża zależność rozkładu relacji losowych w zależności od wartości średniej B_{sr} (patrz wykres 5). Przy wysokiej średniej B_{sr} dla analizowanego profilu prawdopodobieństwo wystąpienia relacji losowej wzrasta. Dodatkowym elementem, który powinien być uwzględniony, jest aspekt ergonomiczny, czyli potrzeba przekazania komunikatywnej informacji w jak najprostszej formie. Należy też brać pod uwagę fakt częstości występowania dla danego L_n profili spokrewnionych i losowych.

Uwzględniając powyższe uwarunkowania, opracowano tabelę komunikatów, które mogą być pomocne dla operatora. Treść komunikatu jest zależna od wartości L_n oraz B_n i zawiera dwa elementy: hipotezę oraz jej wiarygodność.

TABELA 5

Komunikaty

Parametr			Komunikat	
L_n	B_n	B_{sr}	Hipoteza	Wiarygodność
0	100%		Identyfikacja osoby	100%
0	75-99%	< 36%	Rodzic-dziecko	30%
			Rodzeństwo	70%
0	< 75%	< 36%	Rodzic-dziecko	70%
			Rodzeństwo	25%
			Dziadek-wnuk	5%
0	< 75%	> 36%	Rodzic-dziecko	70%
			Rodzeństwo	25
			Dziadek-wnuk	<3%
			Relacja losowa	<1%
1	> 65%	< 36%	Rodzeństwo	70%
			Dziadek-wnuk	20%
			Relacja losowa	10%
1	> 65%	> 36%	Rodzeństwo	60%
			Dziadek-wnuk	20%
			Relacja losowa	20%
1	< 65%	< 36%	Rodzeństwo	30%
			Dziadek-wnuk	60%
			Relacja losowa	10%
1	< 65%	> 36%	Rodzeństwo	20%
			Dziadek-wnuk	60%
			Relacja losowa	20%

cd. tabeli 5

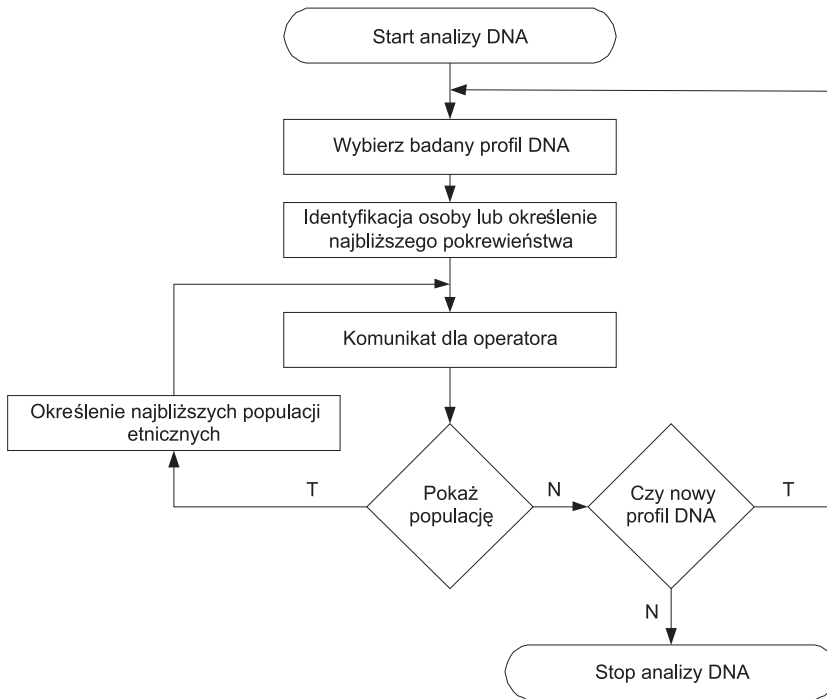
2	> 60%	< 36%	Rodzeństwo	40%
			Dziadek-wnuk	40%
			Relacja losowa	20%
2	> 60%	> 36%	Rodzeństwo	30%
			Dziadek-wnuk	40%
			Relacja losowa	30%
2	< 60%	< 36%	Rodzeństwo	20%
			Dziadek-wnuk	50%
			Relacja losowa	30%
2	< 60%	> 36%	Rodzeństwo	10%
			Dziadek-wnuk	40%
			Relacja losowa	40%
3	> 55%	< 36%	Rodzeństwo	30%
			Dziadek-wnuk	30%
			Relacja losowa	40%
3	> 55%	> 36%	Rodzeństwo	20%
			Dziadek-wnuk	30%
			Relacja losowa	50%
3	< 55%	< 36%	Rodzeństwo	15%
			Dziadek-wnuk	45%
			Relacja losowa	40%
3	< 55%	> 36%	Rodzeństwo	10%
			Dziadek-wnuk	30%
			Relacja losowa	60%

B_{sr} średnie jest obliczane jako wartość średnia z B_n określonych dla wszystkich profili znajdujących się w bazie danych. Wiarygodność należy interpretować jako prawdopodobieństwo wystąpienia hipotezy oszacowane z dokładnością $\pm 5\%$.

Efektom analizy informacyjnej profili DNA STR 16 są metody przetwarzania danych, wzory obliczeniowe i komunikaty dla operatorów. Umożliwia to opracowanie odpowiednich algorytmów.

8. Algorytmy ekstrakcji informacji z profili DNA STR 16

Opierając się na zweryfikowanych metodach określono algorytmy analizy profilu DNA i przetwarzania danych. Algorytm główny przedstawiony jest na rysunku 6.

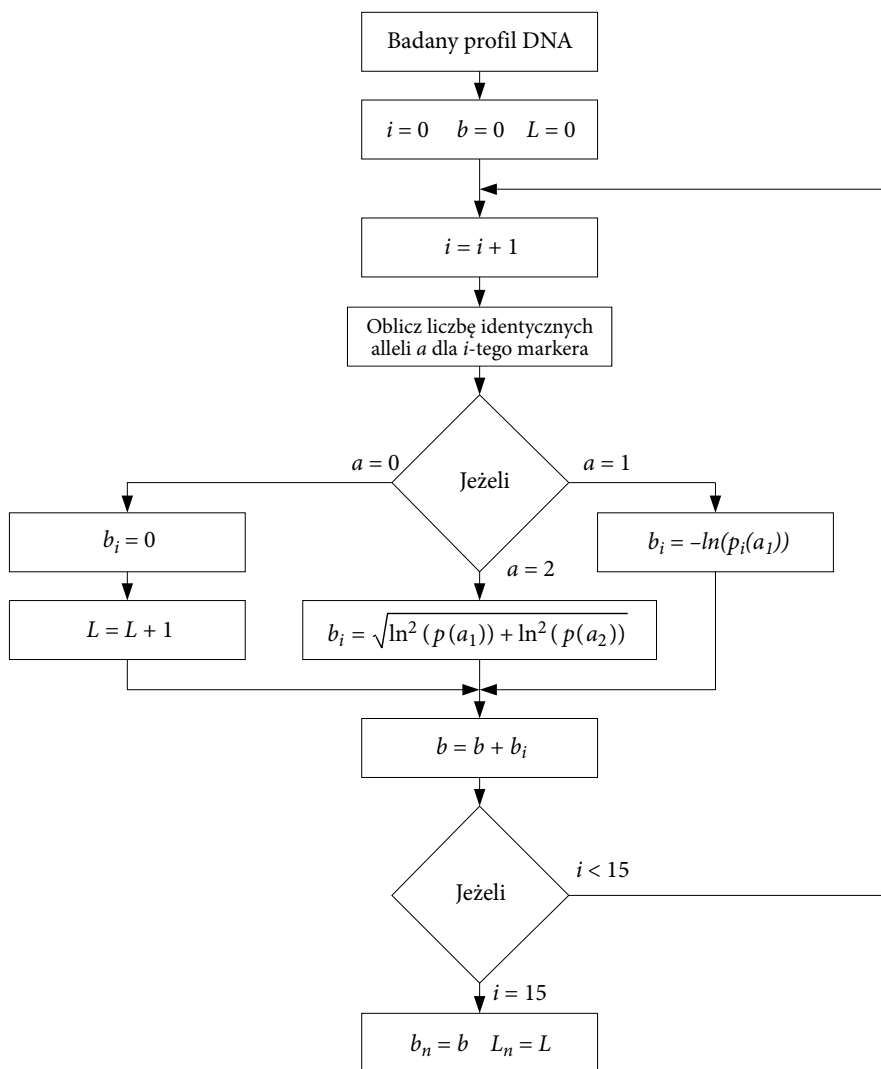


Rys. 6. Główny algorytm analizy badanego profilu DNA

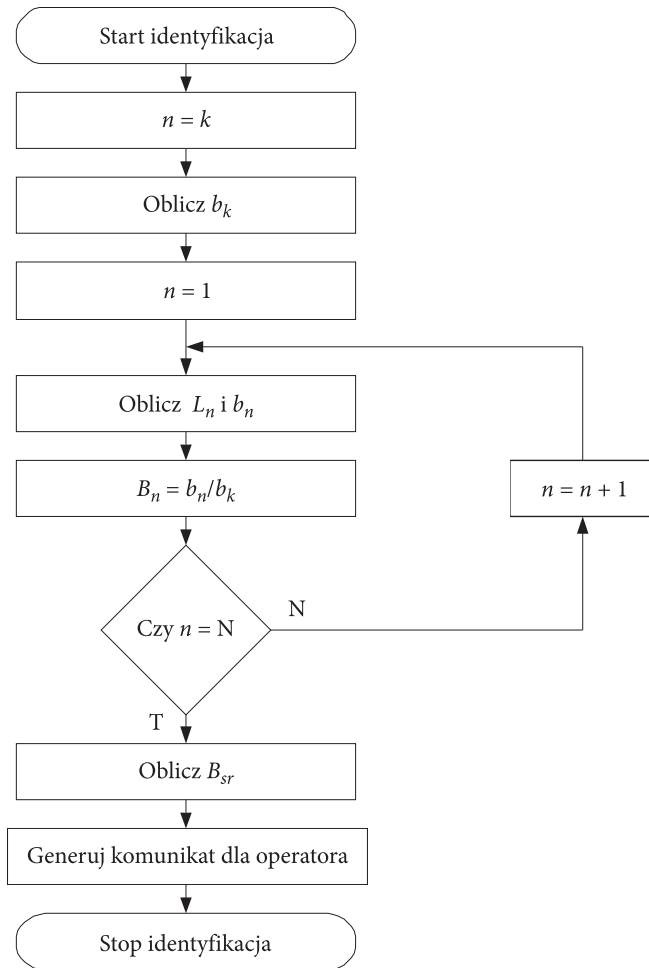
Po uruchomieniu procedury analizy profili DNA powinien pojawić się folder umożliwiający wczytanie lub wybór profilu do analizy. Akceptacja wybranego profilu automatycznie wywołuje procedurę identyfikacji osoby i oceny stopnia pokrewieństwa oraz wyświetlenie komunikatu dla operatora. Dodatkowo może być dokonany wybór opcji określenia najbardziej prawdopodobnej przynależności do grup etnicznych. Następnie operator może wybrać kolejny profil lub zakończyć analizę DNA. W przedstawionym algorytmie moduły: *Identyfikacja osoby lub określenie najbliższego pokrewieństwa* oraz *Określenie najbliższych populacji* są oddzielnymi, złożonymi algorytmami.

Optymalizując przetwarzanie danych, przyjęto, że identyfikację osoby, identyfikację rodzicielstwa oraz określenie pokrewieństwa można ustalić na podstawie tych samych dwóch parametrów: względnej bliskości genetycznej B oraz liczby niezgodnych markerów L , przy czym dla potrzeb generowania automatycznej decyzji niezbędne jest także obliczenie wartości B średnie dla profili znajdujących się w bazie. Obliczenia realizują algorytmy przedstawione na rysunkach 7 i 8.

Identyfikacja osoby i pokrewieństwa polega na obliczeniu tablicy wartości bezwzględnych bliskości genetycznej B_n , wartości średniej B_{sr} oraz liczby niezgodnych markerów L_n . Wielkości te umożliwiają wygenerowanie komunikatu dla operatora zawierającego hipotezę wraz z oceną jej wiarygodności. Listę komunikatów oraz wartości determinujące ich generowanie zawiera tabela 5.

**Oznaczenia:** i — numer markera; a — liczba zgodnych alleli; b — wartość bezwzględnej bliskości genetycznej; L — liczba niezgodnych markerów; b_i — wartość bliskości genetycznej dla i -tego markera; $p_i(a_1)$ — prawdopodobieństwo występowania allela dla i -tego markera w populacji; n — indeks profilu w bazie danych; b_n — wartość bezwzględnej bliskości genetycznej dla n -tego profilu; L_n — liczba niezgodnych markerów dla n -tego profilu; B_n — wartość względnej bliskości genetycznej dla n -tego profilu; B_{sr} — wartość średnia bliskości genetycznej.

Rys. 7. Algorytm obliczania bliskości genetycznej bezwzględnej b oraz liczby niezgodnych markerów L dla profilu DNA o indeksie n



Rys. 8. Algorytm identyfikacji osoby i określania pokrewieństwa

9. Podsumowanie

Badanie DNA oprócz identyfikacji tożsamości na bardzo wysokim poziomie wiarygodności umożliwia także ocenę pokrewieństwa. Opracowane algorytmy i metody umożliwiają identyfikację i ocenę pokrewieństwa w oparciu o dwa uniwersalne wskaźniki. Jest to podejście różniące się od dotychczas stosowanych, ale jego skuteczność została potwierdzona statystycznie.

Opracowano również metodę oceny dystansu genetycznego pojedynczej osoby do populacji. Ten typ oceny w zasadzie nie jest stosowany na bazie badania STR. Jednak analizy przeprowadzone na bardzo licznej próbie wskazują na możliwość określenia odległości genetycznej względnej w relacji do populacji bazowej.

Publikacja powstała w ramach pracy badawczej: OR 00002912.

LITERATURA

- [1] J. CZARNY, T. GRZYBOWSKI, M. DERENKO, B. MALYARCHUK, D. MIŚCICKA-ŚLIWKA, *Genetic variation of 15 STR loci in populations of north and central Poland*, Forensic Science International, 147, January 2005, 97-100.
- [2] P. KUZNIAR, R. PŁOSKI, *STR data for the power plex-16 loci in a population, from Central Poland*, Forensic Science International, 139, 2004, 261-263.
- [3] D. PAETKAU, L. WAITS, P. CLARKSON, L. CRAIGHEADG, C. STROBE, *An Empirical Evaluation of Genetic Distance Statistics Using Microsatellite Data From Bear (Ursidae) Populations*, Genetics Society of America; Genetics 147, December 1997, 1943-1957.

W. OLCHOWIK

Methods of personal identification with high credibility

Abstract. Amongst the known biometric authentication techniques, the highest rate of credibility is definitely attained by usage of DNA profiling. This paper describes the methodology of identification based on STR profiles. The created algorithms and methods will allow the identification and evaluation of kinship on the basis of two fundamental indicators. It is a different approach from the ones, which are currently being used. However, its credibility has been proven statistically. Data processing has been integrated with other biometric identification techniques.

Keywords: personal identification, biometric technology, DNA profile

