

Rafał Wojtkiewicz

**Analiza polimorfizmu 12 regionów
autosomalnego DNA systemu HDplex
w badaniach genetyczno-sądowych**

ROZPRAWA DOKTORSKA

PROMOTOR

Dr hab. n.med. Renata Jacewicz

Pracownia Genetyki Medycznej i Sądowej
Katedry i Zakładu Medycyny Sądowej



Łódź 2017

VII. Streszczenie

Opracowanie przedstawia analizę 530 osobowej próby populacyjnej niespokrewnionych osób z regionu łódzkiego Polski Centralnej stanowiącej postawę do oceny częstości występowania określonych cech w zakresie badanych regionów STR autosomalnego DNA wchodzących w skład systemu Investigator HDplex. Jest to pierwsza baza opracowana dla kompletnego zestawu dwunastu markerów tego multipleksu w populacji Polski. Uzyskany rozkład genotypów i alleli wskazuje na niezależność dziedziczenia markerów zarówno w obrębie zestawu HDplex, jak i w odniesieniu do syntenicznych markerów z zestawów Identifiler i NGM. Wykazana zgodność rozkładów cech z równowagą Hardy'ego i Weinberga pozwala na wykorzystanie zebranych danych populacyjnych jako bazy umożliwiającej przeprowadzenie obliczeń wartości dowodowej w genetyce sądowej, jak i ocenę chimeryzmu po allo-HSCT. Z uwagi na to, że badana populacja Polski zalicza się do jednej z najbardziej endogennych populacji Słowian kaukaskich w środkowo-wschodniej części Europy, uzyskana baza danych może stanowić referencyjną do obliczeń częstości wystąpienia profilu DNA.

Zestawienia rozkładów cech uzyskanych globalnie dla populacji całego świata, jak również w obrębie poszczególnych populacji pozwoliły na ujawnienie charakterystycznych rozkładów dla każdego z badanych markerów przy jednoczesnych znacznych różnicach w częstościach alleli w oddalonych terytorialnie i kulturowo populacjach. Wykorzystanie techniki radialnych drzew oraz skalowania wielowymiarowego pozwoliło na zobrazowanie dystansu genetycznego dzielącego badaną populację od innych populacji Europy i świata, który okazał się być tym większy, im większy dystans geograficzny dzielił populacje. Przeprowadzona analiza filogenetyczna stanowi w pewnym stopniu odzwierciedlenie liczby wspólnych przodków i kolejnych rozejść populacji, co wskazuje na przydatność badanych markerów przy ustalaniu przybliżonego pochodzenia biogeograficznego.

Najwyższe wskaźniki zmienności w zakresie heterozygotyczności, siły dyskryminacji, zmienności genowej, siły wyłączenia, indeksu ojcostwa, wskaźnika polimorfizmu oraz najniższe wartości prawdopodobieństwa przypadkowej zgodności odnotowano kolejno dla loci: SE33, D21S2055, D12S391, D18S51, D10S1132, D8S1132, D7S1517. Czyni to te markery najbardziej przydatnymi dowodowo w aspekcie badań genetyczno-sądowych.

Najmniej polimorficznymi, a tym samym dającymi najmniejszą wartość dowodową okazały się kolejno loci: D6S474, D4S2366 oraz D2S1360. Tym niemniej średnie wartości PD i PE dla układu w systemie Investigator HDplex, wynoszące odpowiednio 95,73% i 70,06% są znacznie wyższe w porównaniu z analogicznymi wartościami tych parametrów w systemach Identifier i NGM.

Wysokie wartości parametrów przydatności, m.in. siły dyskryminacji i siły wyłączenia, wynoszące odpowiednio 0,999999999999999996 i 0,99999974 oraz niska średnia wartość współczynnika wsobności 0,0018, potwierdzają bardzo wysoki stopień przydatności badanego multipleksu w badaniach z zakresu genetyki sądowej, zarówno w aspekcie identyfikacji osób, jak i analiz pokrewieństwa.

Analiza spraw ojcostwa wykonanych w latach 2012 – 2016 w Pracowni Genetyki Medycznej i Sądowej ZMS UM Łódź dokumentuje 100% skuteczność wyłączania ojcostwa za pomocą zestawu markerów Investigator HDplex w zakresie od sześciu do dziesięciu markerów (najczęściej ośmiu). Potwierdzenie ojcostwa z wymaganym poziomem prawdopodobieństwa graniczącego z pewnością, tj. co najmniej 99,9999% uzyskano w praktyce badawczej w 98,04% spraw, w których uczestniczył biologiczny ojciec dziecka.

Analiza dziedziczenia w parach rodzic dziecko, przeprowadzona na podstawie analizy ogółem 2868 mejoz, wykazała wyłącznie odojcowskie pochodzenie mutacji w 4 badanych układach, tj. D21S2055, D6S474, D8S1132, SE. Określona na tej podstawie wartość współczynnika mutacji dla badanego zestawu markerów Investigator HDplex osiągnęła stosunkowo niską wartość $1,45 \times 10^{-4}$. Średnia wartość mutacyjności dla pojedynczego markera układu wyniosła $1,74 \times 10^{-3}$ i nie przekroczyła powszechnie akceptowanego poziomu mutacyjności dla przydatnych w badaniach sądowych układów STR.

Z uwagi na wysoką siłę dyskryminacji badanych markerów, z których dziewięć nie jest dostępnych w żadnym innym komercyjnym multipleksie, zestaw markerów Investigator HDplex stanowi doskonałe wsparcie systemów: Combined DNA Index System (CODIS) i European Standard Set (ESS) przy rozwiązywaniu najtrudniejszych spraw z zakresu identyfikacji osób i analizy pokrewieństwa, jak również w ocenie chimeryzmu po alloprzeszczepie.

VII. Summary

The study presents an analysis of population sample of 530 unrelated people from Lodz region of Central Poland, which is the basis for the assessment of allele frequency occurrence in the range of investigated STR regions of autosomal DNA included in the Investigator HDplex system. This is the first database developed for the complete set of twelve markers of this multiplex in Polish population. The distribution of genotypes and alleles shows the independence of inheritance both within the HDplex markers as well as in their reference to the syntenic markers from Identifiler and NGM sets. The received agreement of the distributions with Hardy and Weinberg equilibrium allows the use of the collected population data as a platform that allows to calculate the evidence value in forensic genetics as well as the chimerism state after allo-HSCT. Due to the fact that studied Polish population belongs to one of the most endogenous populations of Caucasian Slavs in the central-eastern part of Europe, the established database may constitute a referential database for the evaluation of the DNA profile frequency.

The juxtaposition of the allele distribution obtained globally in the world's population as well as within particular populations allowed to reveal the characteristic distribution for each of the markers with simultaneous significant differences in allele frequencies in territorially and culturally distant populations. The use of the technique of radial trees and multidimensional scaling allowed to illustrate the genetic distance separating the investigated population from other populations of Europe and the world, which turned out to be the greater, the greater was the geographical distance between populations. The performed phylogenetic analysis reflects to a certain extent the number of common ancestors and subsequent divergences of the populations, which indicates the usefulness of the investigated markers in determining of approximate biogeographical origin.

The highest indicators of variability in the range of heterozygosity, power of discrimination, gene variability, power of exclusion, paternity index, polymorphism information content as well as the lowest values of matching probability were noted consecutively for the loci: SE33, D21S2055, D12S391, D18S51, D10S2325, D8S1132, D7S1517. That makes these markers the most useful evidentially in the aspect of forensic genetics research.

The least polymorphic, thus giving the smallest value of the evidence proved to be successively loci: D6S474, D4S2366 and D2S1360. However, the average values of PD and PE for the Investigator HDplex system, amounting to 95.73% and 70.06%, are significantly higher in comparison with the analogical values of these parameters in the in the Identifiler and NGM systems.

The high values of usefulness parameters, among others the power of discrimination and the power of exclusion amounting to 0.999999999999999996 and 0.99999974 respectively, as well as low average value of inbreeding coefficient of 0.0018, confirm the very high degree of usefulness of the investigated multiplex in the research of forensic genetics, both in the term of personal identification and family relationship.

Evaluation of paternity cases conducted in the years 2012-2016 in Medical and Forensic Genetics Laboratory, Department of Forensic Medicine, Medical University of Lodz documents 100% effectiveness of paternity exclusion with the use of Investigator HDplex in the range from six to ten markers (most often eight). The confirmation of paternity with probability close to certainty, i.e. at least 99.9999%, was obtained in the research practice in 98.04% of cases with the biological father of the child.

Analysis of inheritance in parent-child pairs, carried out on the basis of the analysis of a total 2868 meioses, showed only paternal origin of mutation in 4 investigated markers, i.e. D21S2055, D6S474, D8S1132, SE33. In this way determined value of the mutation coefficient for the tested set of markers of Investigator HDplex reached relatively low value of $1,45 \times 10^{-4}$. The average value of mutability for the single marker of the set equaled $1,74 \times 10^{-3}$ and did not exceeded commonly accepted level of mutability for STR markers, that are useful in forensic research.

Due to the high discriminatory power of the investigated markers, of which nine are not available in any other commercially available multiplex, the Investigator HDplex constitutes an excellent support of the systems: Combined DNA Index System (CODIS) and European Standard Set (ESS) in resolving the most difficult cases in the field of person identification and kinship analysis, as well as in the assessment of chimerism state after allotransplantation.