OCENA

rozprawy doktorskiej mgr Michała Boronia zatytułowanej

Ocena przydatności polimorfizmu sekwencji flankujących markerów mikrosatelitarnych w identyfikacji genetycznej człowieka.

Mgr Michał Boroń przedłożył ocenianą pracę Uniwersytetowi Medycznemu w Białymstoku jako rozprawę na stopień doktora nauk medycznych.

Odkrycie polimorfizmów DNA i nowoczesne technologie ich oznaczania wywarły olbrzymi wpływ na rozwój wielu dyscyplin biologii. Badania polimorfizmu DNA typu STR i SNP są obecnie podstawą profilowania genetycznego osób, co pozwala między innymi na ustalanie pokrewieństwa, pochodzenia jak również indywidualizację materiału biologicznego. Unikalność profili DNA zależy w największym stopniu od ilości oznaczanych markerów i ich heterozygotyczności ale niezwykle ważna jest technologia badania, która pozwala na uzyskanie niepodważalnych wyników fenotypowania lub genotypowania. Technika elektroforezy kapilarnej pozwala na zróżnicowanie długości alleli STR z dokładnością do nukleotydu, ale nie pozwala na wykrywanie wariantów w sekwencjach powtarzalnych, jak również polimorfizmu wynikającego z mutacji w regionach flankujących.

Podjęcie przez Doktoranta badań zmienności genetycznej regionów flankujących 27 układów STR używanych w genetyce sądowej w oparciu o dane sekwencji wygenerowanych w technologii NGS należy uznać za bardzo celowe.

Przedłożona do oceny rozprawa doktorska jest przedstawiona na 98 stronach, zawiera 22 tabele, 10 wykresów oraz 6 rycin. Na początku pracy jest podany *Spis treści* a następnie *Streszczenie* w języku polskim i angielskim. Wyodrębnione zostały następujące rozdziały: *Wstęp, Materiały i metody, Wyniki, Dyskusja, Wnioski, Wykaz rycin, tabel i wykresów, Bibliografia.* Piśmiennictwo obejmuje 106 pozycji literatury z renomowanych czasopism przedmiotu, prawidłowo dobranych do tematyki rozprawy.

Rozprawa ma bardzo zbitą formę edytorską (do 31 wierszy na stronie), co utrudnia lekturę. Część tabel jest łamana i wydrukowana na następnej stronie, a nawet, jak w przypadku tabeli XXII, bez zachowania ciągłości. Te same markery zmieniają pozycje w kolejnych tabelach, co bardzo utrudnia ich analizowanie. Polszczyzna doktoranta pozostawia wiele do życzenia. Zob. np. zdania zaczynające się od: „Stąd wzięła się..." (str. 4), „Inaczej sprawa ma się" (str. 8), „Aby podsumować.." (str. 52), „Jak widać" (str. 56). Autor wprowadza swoje tłumaczenia angielskich terminów specjalistycznych, które mają inne odpowiedniki w języku polskim. Np. „przyczep primerów„ zamiast „hybrydyzacja starterów", „allel zerowy" zamiast „allel niemy" (str. 8). Przedstawiając technologię NGS, poszczególne etapy i narzędzia informatyczne do analizy sekwencji flankujących Doktorant podaje czasem terminy w wersji angielskiej, czasem dołącza swoje tłumaczenie na język polski, a czasem wyjaśnienie można znaleźć jedynie czytając tekst pracy. Rozprawie brakuje zestawienia używanych skrótów i tłumaczeń, które ułatwiłoby czytanie pracy. Zob. zwłaszcza ryc. 3, która jest bardzo przydatna do zrozumienia technologii NGS, ale trudna w percepcji bez szczegółowej wiedzy o tej metodzie sekwencjonowania.

Wstęp pracy podzielony jest na 4 podrozdziały: *Metody genetycznej identyfikacji człowieka, Zastosowanie technologii NGS w Kryminalistyce, Zmienność genetyczna regionów flankujących markerów STR* i *Cel pracy*. Informacje zawarte we *Wstępie* wyczerpująco przedstawiają

współczesną wiedzę na w/w tematy i w klarowny sposób uzasadniają cel podjętych badań oraz ich aktualność.

Materiał do badań regionów flankujących stanowią wyniki sekwencjonowania 308 niespokrewnionych osób z populacji Polski uzyskane technologią NGS (Illumina) przy użyciu zestawu ForenSeq TM DNA Signature Prep Kit w ramach projektu NEXT(DOB-BIO/17/01/2015). Jakoże nie jestem w stanie sprawdzić czy w/w projekt uwzględniał badanie regionów flankujących, zakładam, że wyniki sekwencjonowania zostały udostępnione Doktorantowi za zgodą autorów projektu.

W podrozdziale 2.5 *Analiza danych* autor przedstawia krok po kroku, w jaki sposób ustalił sekwencje referencyjne badanych 27 markerów STR, które przedstawił w tabeli VI, następnie przygotował do pracy z programem FDStools i tak uzyskane pliki Multi-FASTA przeanalizował używając program BioEdit Sequence Alignment Editor.

Wyniki badań przedstawione są w trzech podrozdziałach: 3.1 *Metoda prezentacji wyników*, 3.2 *Zaobserwowane polimorfizmy sekwencji* i 3.3 *Parametry biostatystyczne*.

Zaproponowana przez Doktoranta nomenklatura izoalleli uwzględniająca zmiany w sekwencjach powtarzalnych jak i flankujących jest logiczna i prosta. Wprowadzenie takiej nomenklatury w prezentowanej dysertacji pozwala na przejrzyste prezentowanie wyników. Powodzenie takiego nazewnictwa jest jednak uwarunkowane od powszechnej akceptacji oraz centralnego katalogowania wariantów allelicznych.

Zaobserwowany polimorfizm sekwencji flankujących został przedstawiony za pomocą dwóch wykresów (ad 1 i 2), jednej ryciny (ad 6) oraz pięciu tabel (ad VII, VIII, IX, X, XI), natomiast parametry biostatystyczne zestawione zostały w czterech tabelach (ad XII, XIII, XIV i XV). Wyniki badań są wszechstronnie opisane i zrozumiale zaprezentowane na rycinach i w tabelach. Przed złożeniem pracy do druku należy skorygować drobne nieścisłości. Np. Tabela VIII zgodnie z opisem ma przedstawiać 40 nowych alleli, podczas gdy zawiera dane dla 39 alleli (str. 27). W tabeli X brakuje danych dla markera vWA w pozycji RF. Najwięcej informacji zawartych jest w tabeli VII, która jest dobrze zredagowana. Moim zdaniem kolejność markerów w dalszych tabelach powinna być taka sama.

Parametry biostatystyczne: MP (pawdopodobieństwo identyczności), PD (siła dyskryminacji), PIC (współczynnik informacji o polimorfizmie) i Ht (heterozygotyczność) zostały prawidłowo dobrane. Wyniki obliczeń zestawione w czterech tabelach pozwalają na ocenę przydatności badanego polimorfizmu sekwencji flankujących przy ustalaniu profili genetycznych.

W podrozdziale 3.1 Doktorant odsyła czytającego do tabeli XXII, która *„przedstawia sekwencje zaobserwowanych alleli a także ich częstości„* dla alleli oznaczonych CE i RR. Dołączenie tej tabeli wymaga wyjaśnienia, ponieważ zamieszczone dane częstości alleli obliczone są na podstawie badań populacji wielokrotnie większej niż badana grupa 308 osób. Częstości te są niezbędne do obliczeń paramentrów biostatystycznych w tej pracy, stanowią podstawę do analiz i dyskusji, a nawet wniosków, należy zatem podać ich pochodzenie.

Dyskusja przeprowadzona przez Doktoranta na 14 stronach pracy wsparta 8 wykresami oraz 6 tabelami jest, moim zdaniem, najlepszą częścią dysertacji. Autor wszechstronne omawia parametry biostatystyczne oceniające wartość polimorfizmu regionów flankujących (RF) analizując zależności między ilością alleli danego markera, ich częstością z uwzględnieniem długości alleli (CE) oraz sekwencji repetytywnych (RR).

Dywagacja Doktoranta dotycząca heterozygotyczności dwu allicznego markera jest błędna, ponieważ w takim układzie Ht nie przekracza wartości 0.5, a najwyższą wartość osiąga przy częstości 0,5 dla obu alleli (str. 47).

Interesujące są natomiast porównania uzyskanych wyników z danymi opublikowanymi dla innych populacji w zakresie częstości SNP w regionach flankujących, średniej heterozygotyczności, wartości MP, PD oraz PIC. W porównaniu zostały uwzględnione również te częstości alleli SNP, których znane są rs i znajdują się w bazie dbSNP.

Na wyróżnienie zasługuje próba obiektywnej oceny parametrów statystycznych oceniających wartość polimorfizmu regionów flankujących, sformułowanie alternatywnych hipotez i szacowanie ich wiarygodności.

Wyciągnięte przez autora wnioski wynikają z przeprowadzonych w pracy badań i obliczeń biostatystyczych.

Reasumując należy uznać przedłożoną do oceny pracę za wartościową, ponieważ wykazano w niej, że:

- zastosowana technologia NGS jest bardzo użyteczna do oznaczania polimorfizmu sekwencyjnego w 27 markerach STR rutynowo stosowanych przy profilowaniu DNA u ludzi,
- analiza sekwencji flankujących pozwoliła na wyodrębnienie 11 markerów STR wykazujących zmienność genetyczną, a w 9 z nich obliczono częstości alleliczne,
- zaproponowane przez Doktoranta nazewnictwo wariantów allelicznych jest logiczne i uzasadnione,
- wykryte delecje w regionach flankujących markerów D13S317, PentaD, D9S1122 zmieniają fenotypowanie STR,
-częstości alleli uzyskane dla badanej grupy mieszkańców Polski nie wykazują znaczących różnic w porównaniu do innych białych populacji Europy ,
-współczynniki biostatystyczne wartościujące polimorfizm DNA dla celów profilowania człowieka (Ht,MP,PD oraz PIC) potwierdzają użyteczność SNP regionów flankujących następujących markerów STR: D13S317, D5S818, D20S482, D7S820 i D2S441.

Warto dodatkowo zaznaczyć, że wykonanie badań polimorfizmu sekwencji flankujących odbywa się na bazie wyników uzyskanych w trakcie sekwencjonowania technologią NGS autosomalnych markerów STR.

Większość moich uwag dotyczy sposobu edytowania rozprawy. Jedynie te uwagi, które dotyczą meritum pracy, domagają się odpowiedzi doktoranta w trakcie obrony.

Przedłożoną do oceny rozprawę doktorską mgr Michała Boronia p. t. *"Ocena przydatności polimorfizmu sekwencji flankujących markerów mikrosatelitarnych w identyfikacji genetycznej człowieka"* oceniam pozytywnie.

Zwracam się zatem do Senatu Uniwersytetu Medycznego w Białymstoku o dopuszczenie Mgr Michała Boronia do dalszych etapów przewodu doktorskiego.

dr hab. n. med. Piotr Kozioł