

Recenzja rozprawa na stopień doktora nauk medycznych pod tytułem:

**Sekwencjonowanie regionów kodujących genomu w
poszukiwaniu podłoża genetycznego wybranych chorób
wieloczynnikowych**

autorstwa **mgr inż. Marii Kuleckiej**

Sekwencjonowanie następnej generacji (NGS) zrewolucjonizowało genetykę, w tym także genetykę medyczną. Obecnie nie ma wątpliwości, że NGS stosowany do badań paneli genów lub całego eksomu (ang. whole-exome sequencing, WES) jest metodą z wyboru w diagnostyce większości chorób uwarunkowanych jednogenowo. Ponieważ liczne choroby mają uwarunkowanie nie jednogenowe a wieloczynnikowe istotnym zagadnieniem jest wykorzystanie technologii NGS, w tym WES, także na tym obszarze. Ze względu na wysoką częstość wielu chorób wieloczynnikowych możliwości wykorzystania technologii NGS w ich diagnostyce i badaniu należy uznać za szczególnie istotny kierunek badawczy.

Jednocześnie należy podkreślić, że nie można automatycznie przenosić metod analizy wyników WES wypracowanych do analiz chorób rzadkich i ultraradkich na pole chorób wieloczynnikowych. Jest to spowodowane stosunkowo dużą liczbą, słabą penetracją i stosunkowo wysokimi częstościami w populacji zdrowej wariantów, które odpowiadają za predyspozycję do chorób wieloczynnikowych.

Kolejnym istotnym aspektem badań w technologii NGS jest wykorzystywana platforma analityczna. Podczas gdy zdecydowaną przewagę na rynku mają platformy firmy Illumina, w niektórych klinicznych zastosowaniach pewną przewagę mają inne aparaty, np. te oparte na technologii półprzewodnikowej. Dla pełnego wykorzystania ich zalet niezbędne jest jednak opracowywanie dedykowanych ścieżek analizy wyników.

Praca Kandydatki dobrze wpisuje się w opisane powyżej nurty badań. Pierwszym celem pracy Kandydatki było opracowanie bioinformatycznych metod analizy danych pochodzących z półprzewodnikowego sekwencjonowania egzomów. Drugą część pracy dotyczyła zastosowań opracowanych metod do

analizy trzech grup pacjentów: z kryptogenną cholestazą wewnątrzwątrobową, z nieswoistymi chorobami zapalnymi jelit oraz z macroAST. Dodatkowo, w drugiej części pracy Autorka zajmowała się opracowaniem metod analizy biostatystycznej odpowiednich do badanych zagadnień.

Jestem głęboko przekonany, że najlepszym kryterium oceny każdego opracowywanego sposobu analizy danych są wyniki jego zastosowania w praktyce. W przypadku Pracy Kandydatki wyniki te są spektakularne – w jednej z badanych grup, tj. wśród pacjentów z macroAST, analiza Kandydatki pozwoliła na dokonanie w pełni oryginalnego odkrycia efektu wariantu *GOT1* p.Gln208Glu. Kandydatka nie tylko wykryła zwiększoną częstość tego wariantu w grupie pacjentów ale pokazała przekonująco statystyczną znamienność (w skali całogenomowej) stwierdzonej asocjacji. Ostatecznym potwierdzeniem rangi odkrycia jest jego publikacja w znakomitym piśmie (A heterozygous mutation in *GOT1* is associated with familial macro-aspartate aminotransferase. *Journal of Hepatology*. 2017; 67: 1026–30). Należy podkreślić, że Kandydatka jest pierwszym autorem ww. publikacji.

Przedstawiona rozprawa liczy 74 strony. Cytuje ponad 175 pozycji aktualnego piśmiennictwa. Główne części pracy to wstęp, opis celu pracy, sekcja materiały i metody, omówienie wyników i dyskusja i wnioski. Dodatkowo praca zawiera spis treści, Spis rysunków i tabel, Wykaz skrótów i pojęć, streszczenie w języku polskim i angielskim oraz informacje bibliograficzne o 3 artykułach opublikowanych w piśmiennictwie specjalistycznym zawierających wyniki opisane w rozprawie. Układ pracy jest typowy i przejrzysty. Wstęp stanowi dobre wprowadzenie do tematu omawiając zagadnienie NGS z uwzględnieniem specyfiki analiz na platformie półprzewodnikowej.

Podczas lektury nasunęły mi się drobne uwagi, które przytaczam z obowiązku recenzenta:

Skrót macroAST (Streszczenie) nie jest wyjaśniony

Legenda Rysunek 3. (Drzewa rodzinne pacjentów z cholestatycznymi chorobami wątroby) zawiera błąd (edytorski) dotyczący symbolów określających płeć.

Str 31 „W teście brano pod uwagę jeden patogenny wariant na gen.” -W jaki sposób selekcjonowano wariant gdy było ich więcej?

Str 31 „W tym projekcie u całkowicie zsekwencjonowanej rodziny wybrano homozygotyczne warianty dziedziczone w sposób autosomalny recesywny” dziedziczenie pokazane na rycinie nie sugeruje wprost dziedziczenia autosomalnie recesywnego, a raczej mutacje *de novo* u FAM7.3.

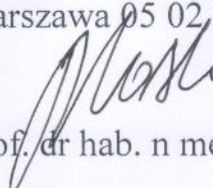
Str 33 i inne „dzielonych wariantów” – lepiej „wspólnych”

Str 48 „Średnie pokrycie we wszystkich próbkach znacznie przekraczało pokrycie potrzebne do prawidłowej identyfikacji wariantu (wynosi ono, w zależności od programu, od 5x do 20x).” Autorka pomija możliwość braku identyfikacji wariantów heterozygotycznych z powodu losowej analizy ograniczonej do jednego chromosomu. Przy pokryciu 5x szansa, że wszystkie odczyty pochodzą z jednego chromosomu to $1/2^5=1/32$. Uwzględniając dziesiątki tysięcy wariantów obecnych w eksomie można oczekiwać, że pokrycie 5x skutkuje utratą setek wariantów.

Pomimo opisanych powyżej drobnych uwag uważam, że przedstawiona praca autorstwa mgr inż. Marii Kuleckiej pt.: 'Sekwencjonowanie regionów kodujących genomu w poszukiwaniu podłoża genetycznego wybranych chorób wieloczynnikowych' spełnia wymagania stawiane rozprawie na stopień doktora nauk medycznych i w związku z tym wnioskuję o dopuszczenie Kandydatki do dalszych etapów przewodu doktorskiego.

Ze względu na wysoką rangę oryginalnego odkrycia dokonanego bezpośrednio z użyciem opracowanych przez Kandydatkę metod wnioskuję o wyróżnienie Pracy.

Warszawa 05 02 2018


Prof. dr hab. n med. Rafał Płoski