



Instytut „Centrum Zdrowia Matki Polki”

Zakład Genetyki

93-338 Łódź; ul. Rzgowska 281/289

tel.: (0-42) 271-12-71; genetyka@iczmp.edu.pl

Kierownik: Dr hab. n. med. Agnieszka Gach, profesor ICZMP

Łódź, 31.09.2023

**Recenzja rozprawy na stopień doktora nauk biologicznych
mgr Marcina Słomki**

**p.t. „Zmienność genów kodujących białka z nadrodziny ABC (ATP-binding cassette)
w populacji polskiej w odniesieniu do innych populacji”**

Promotor: Prof. dr hab. Grzegorz Bartosz

Promotor pomocniczy: Dr Dominik Strapagiel, prof. UŁ

Medycyna precyzyjna nazywana też medycyną spersonalizowaną jest definiowana jako strategia zarządzania zdrowiem oparta na zmienności genetycznej, danych środowiskowych i stylu życia danej osoby, w celu określenia najlepszego podejścia do zapobiegania lub leczenia chorób. W ostatnich kilkunastu latach to właśnie nowe osiągnięcia w genetyce warunkowane rozwojem technik wysokoprzepustowych umożliwiły ogromny postęp w medycynie. W tym przypadku postęp zależy nie tylko od samej identyfikacji często złożonego podłoża genetycznego choroby, ale także genetycznych czynników prognostycznych rozwoju choroby czy skuteczności terapii.

Ze względu na swoją fundamentalną rolę w procesach związanych z detoksykacją organizmu rodzina białek ABC ma istotne znaczenie w odpowiedzi na zastosowaną farmakoterapię. Odpowiedź osobnicza na przyjmowane leki może być bardzo zróżnicowana, jest konsekwencją m.in. transportu i metabolizmu leków, punktu uchwytu leku i patofizjologicznych cech organizmu. Każdy etap jest bardzo zmienny i zależy także od profilu genetycznego pacjenta. Stąd spersonalizowana farmakoterapia opiera się na międzyosobniczej zmienności genetycznej, a więc i różnym rozpowszechnieniu wariantów genetycznych w poszczególnych populacjach.

Recenzowana rozprawa doktorska jest poświęcona analizie zmienności genów ABC w populacji polskiej w kontekście innych populacji, zarówno o odrębnym pochodzeniu etnicznym, jak i o pochodzeniu europejskim. Praca wpisuje się w nurt badań obejmujących

genetyczne podstawy zróżnicowania międzypopulacyjnego w genach kodujących transportery o znaczeniu farmakogenetycznym i epidemiologicznym. Ze względu na kluczowe znaczenie tej rodziny białek w rozwoju medycyny precyzyjnej, podjęty przez Doktoranta temat jest bardzo aktualny w aspekcie naukowym i potrzebny w praktyce klinicznej.

Przedstawiona do oceny dysertacja ma typowy dla rozpraw doktorskich układ redakcyjny. Zawiera 260 stron, w tym spis treści, 97 stron tekstu pracy, streszczenia w języku polskim i angielskim, publikacje wchodzące w skład rozprawy, suplementy zawierające wykresy i tabele dokumentujące opublikowane wyniki, zestawienie dorobku Doktoranta niewchodzącego w skład rozprawy. W rozprawie zamieszczono także informację o źródłach finansowania i współpracy oraz oświadczenia współautorów wskazujące na wiodący udział mgr Słomki w publikacjach. Pracę cechuje tradycyjny podział na wstęp, cel pracy i hipotezy badawcze, materiały i metody, wyniki badań, dyskusję, wnioski i weryfikację hipotez badawczych oraz bibliografię. W zasadniczej części pracy umieszczono 12 tabel i 47 rycin, w suplementach dodatkowo 87 tabel i 86 rycin. W skład rozprawy wchodzi zarówno dwie prace oryginalne opublikowane w czasopiśmie o wysokim współczynniku oddziaływania, jak i dane jeszcze nieopublikowane.

Dysertację otwiera „Wstęp”, w którym Doktorant szczegółowo scharakteryzował aktualny stan wiedzy nt. nadrodziny białek ABC, w tym budowy poszczególnych rodzin, charakterystycznych domen i złożonej funkcji białek. W dalszych podrozdziałach wstępu Autor omówił znaczenie wariantów genów kodujących białka ABC w aspekcie związku z rzadkimi chorobami genetycznymi. Na wybranych przykładach, jak m.in. choroba tangerska, choroba Stargarta, rybia łuska arlekinowa czy mukowiscydoza przybliżył patomechanizm choroby. Kolejno z dużą starannością i licznymi odniesieniami do piśmiennictwa scharakteryzował znaczenie białek ABC w odpowiedzi na farmakoterapię, krótko omawiając zagadnienie zjawiska oporności wielolekowej MDR, z zaznaczeniem ugruntowanej pozycji białek MDR1 oraz BCRP. Dalej opisał założenia farmakoterapii spersonalizowanej i przybliżył badania nad profilowaniem genów związanych z farmakogenetką w populacji polskiej i w populacjach o innym pochodzeniu etnicznym. W ostatniej części wprowadzenia Pan mgr Marcin Słomka odniósł się do roli technik wysokoprzepustowych jak mikromacierze czy sekwencjonowanie następnej generacji (NGS) w badaniach populacyjnych nad zmiennością genetyczną farmakogenów. Przytoczył wyniki badań wskazujące na duże międzyosobnicze i międzypopulacyjne zróżnicowanie genów ABC.

W mojej ocenie wprowadzenie stanowi szczegółowe i rzetelne omówienie poruszanych zagadnień w świetle bogatego i aktualnego piśmiennictwa. Autor w przystępny sposób pozwala czytelnikowi zapoznać się z problematyką, której dotyczy Jego rozprawa naukowa. Dodatkowo, starannie przygotowana na potrzeby wprowadzenia obszerna tabela porządkuje informacje i ułatwia zapoznanie się z funkcją fizjologiczną i znaczeniem w patogenezie chorób człowieka bądź farmakoterapii wszystkich 48 genów kodujących białka nadrodziny ABC. Wstęp

czyta się z dużą przyjemnością, a jego lektura dowodzi, że Doktorant ma bardzo dobre przygotowanie teoretyczne do realizacji podjętych celów badawczych.

Głównym celem rozprawy doktorskiej Pana mgr Marcina Słomki była charakterystyka zmienności genów nadrodziny ABC w populacji polskiej w kontekście danych z innych populacji o innym pochodzeniu etnicznym, jak i pochodzeniu europejskim. Realizację celu oparto na postawieniu trzech hipotez badawczych:

H1. Obserwowaną zmienność sekwencji genów kodujących białka ABC stanowi głównie zmienność o charakterze rzadkim.

H2. Populacja polska pod względem zmienności genów kodujących białka ABC istotnie różni się od populacji o innym pochodzeniu etnicznym.

H3. Populacja polska pod względem zmienności genów kodujących białka ABC istotnie różni się od innych populacji europejskich.

Po weryfikacji i odrzuceniu hipotezy nr 3 postawiono kolejny cel badawczy, w którym założono identyfikację czynników odpowiadających za wynik klastrowania w populacji polskiej i w referencyjnych populacjach europejskich.

Cele pracy zostały sformułowane w sposób jasny i przejrzysty i wskazują na ambitne założenia pracy badawczej. Co warte podkreślenia, badania takie nie były dotąd przeprowadzone w populacji polskiej. Szeroki zakres analizy genetycznej i imponująca wielkość grupy badanej wskazuje na ogrom podjętej przez Doktoranta pracy.

W sekcji „Materiały i metody” Doktorant charakteryzuje grupę badaną w dwóch podrozdziałach, w pierwszym podrozdziale omawiając grupę badaną i metody wykorzystane w dwóch publikacjach będących częścią rozprawy, a w drugim podrozdziale dla badania jeszcze nieopublikowanego. We włączonych publikacjach grupę badaną stanowili niespokrewnieni ochotnicy z populacji polskiej, których materiał biologiczny w postaci śliny zebrano w ramach projektu TESTOPLEK i zarchiwizowano w Pracowni Biobanku Uniwersytetu Łódzkiego pod akronimem POPULOUS. Do badań użyto losowo wybrane próbki spośród 10 000 zdeponowanych. I tak w pierwszym etapie włączono 190 próbek do badania zmienności genu *ABCC1* metodą skanowania z wykorzystaniem HMR. Następnie w puli innych 380 próbek oceniono częstość występowania specyficznych wariantów metodą celowanego genotypowania. Następną pulę 190 próbek wykorzystano do badania zmienności genu *ABCG2* metodą skanowania HMR. W podrozdziale tym Doktorant szczegółowo opisuje metody biologii molekularnej wykorzystane w pracy badawczej, jak: metoda topnienia produktów PCR o wysokiej rozdzielczości HRM wykorzystana do skanowania i celowanego genotypowania oraz sekwencjonowanie metodą Sangera. Z dużą starannością opisane zostały poszczególne etapy pracy eksperymentalnej, jak i analiza bioinformatyczna i statystyczna zmienności genetycznej oraz nierównowagi sprzężeń bloków haplotypowych *ABCC1* i *ABCG2*.

W drugim podrozdziale Doktorant opisuje grupę badaną i użyte metody do badań jeszcze nieopublikowanych. W tym przypadku do badań włączono imponującą grupę 6481 próbek z populacji POPULOUS oraz wyjątkową grupę 179 ochotników z wytypowanych regionów

geograficznych Polski reprezentujących mniejszości etniczne, grupy kulturowe: Karaimów, polskich Tatarów, polskich Ormian, Podhalan, Kurpiów, Ślązaków i Hrubieszowian. Pobrane próbki wykorzystano jako źródło DNA do badań całogenomowych metodą mikromacierzy SNP. Ze względu na specyfikę badanej grupy Doktorant sięgnął po dwa eseje do wysokoprzepustowego genotypowania, zestawy mikromacierzy różniły się pulą i liczbą wariantów, pozwalając na zgenotypowanie ponad pół miliona wariantów w przypadku Infinium HTS Human Core Exome oraz ponad 1,7 miliona wariantów w przypadku Infinium Multi-Ethnic. W pracy wykorzystano dane z populacji referencyjnych z 5 regionów geograficznych świata reprezentowanych przez 26 różnych populacji, pozyskane z bazy danych *1000 Genomes Project*. W dalszej części podrozdziału Autor opisuje szczegółowo metodę imputacji genotypu przy użyciu serwera TopMed oraz wykonane zaawansowane analizy bioinformatyczne i statystyczne, jak m. in.: analiza skalowania wielowymiarowego, analiza IBS, analiza liczby klastrów, analiza asocjacyjna całego genomu oraz całego transkryptomu.

Podsumowując tę część rozprawy stwierdzam, że dobór metod analitycznych, bioinformatycznych i statystycznych do badań jest właściwy. W mojej ocenie, ten rozdział dysertacji dowodzi, że Doktorant wykonał ogromną pracę laboratoryjną, projektując analizę molekularną, a następnie wykonując poszczególne oznaczenia, a także pracę naukową, opracowując i podsumowując wyniki badań przy zastosowaniu złożonych i zaawansowanych metod bioinformatycznych i statystycznych. Świadczy to o naukowej dojrzałości Autora i jednoznacznie dowodzi Jego przygotowania do samodzielnego prowadzenia badań naukowych.

W pierwszej części rozdziału „Wyniki” Doktorant dokonał syntetycznego omówienia publikacji wchodzących w skład rozprawy. Manuskrypty zostały opublikowane w recenzowanych pismach o zasięgu międzynarodowym i wysokim współczynniku oddziaływania. Łączny IF wynosi 6,428. Publikacje są opracowaniami zbiorowymi, w obydwu pracach Doktorant jest pierwszym autorem. Wkład Doktoranta w przygotowanie każdej publikacji oryginalnej został określony jako 60%, co zostało potwierdzone w oświadczeniach współautorów, zawartych na końcu rozprawy. Doktorant współtworzył koncepcję pracy i projektowanie eksperymentu skanowania genu metodą HRM, wykonał pracę eksperymentalną, dokonał analizy bioinformatycznej, statystycznej oraz interpretacji wyników, odegrał także główną rolę w opracowaniu manuskryptów do publikacji.

Publikacja nr 1 dokumentuje badanie zmienności genetycznej genu *ABCC1* w populacji polskiej. W badanej próbie metodą skanowania znaleziono łącznie 46 wariantów, w tym 30 wariantów rzadkich. Siedem z nich nie zostało wcześniej odnotowanych w bazach danych. Obserwowaną zmienność genetyczną *ABCC1* porównano z zareportowaną zmiennością tego genu w innych populacjach. Dokonano także analizy *in silico* wszystkich niesynonimicznych wariantów. Podobną strategię badania przyjęto w analizie zmienności genetycznej *ABCG2* opisanej w publikacji nr 2. W badaniu zidentyfikowano 17 wariantów. Dwa ze zidentyfikowanych wariantów mogą mieć istotne znaczenie kliniczne, m.in. dla odpowiedzi na

leczenie przeciwnowotworowe. Zmienność genetyczną *ABCG2* w populacji polskiej porównano ze zmiennością tego genu w innych populacjach.

W kolejnym podrozdziale Doktorant omówił wyniki nieopublikowanych jeszcze badań będących częścią rozprawy. Porównano zmienność genów ABC wszystkich populacji referencyjnych z bazy *1000 Genomes Project* z wynikami liczącej 1035 ochotników kohorty polskiej. Następnie przeanalizowano zróżnicowanie populacji europejskich i populacji polskiej. Dokonano analizy klastrowania wyznaczając dziewięć odrębnych klastrów i dalej analizy wariantów związanych z klastrowaniem. Ta część rozdziału Wyniki została wzbogacona o ryciny dokumentujące klastrowanie i analizę nierównowagi sprzężeń oraz obszerną tabelę, w której zaraportowano 647 wariantów istotnie związane z grupowaniem próbek na dziewięć klastrów. Dalej Autor przytacza wyniki badania potencjalnego efektu wariantów na poziom ekspresji genów z bazy danych GTEx. Stwierdzono wpływ na ekspresję czterech genów kodujących białka ABC. Najwięcej wariantów oddziaływało na ekspresję genu *ABCC8* w wybranych tkankach układu nerwowo-mięśniowego. Obszerne dane dotyczące efektu 647 wariantów na poziom ekspresji genów w 54 strukturach/narządach zestawiono w skrupulatnie przygotowanej, czytelnej tabeli. Kolejno Autor zamieszcza dane z grupowań próbek na podstawie procesów biologicznych lub lokalizacji w poszczególnych kompartmentach komórkowych z bazy *Gene Ontology*. W ostatniej części Wyników Doktorant podaje dane dotyczące zróżnicowania genów ABC w populacji polskiej na podstawie badań w grupie 6660 próbek i analizy prawie 25 tysięcy współdzielonych wariantów. Przytacza dane z analizy asocjacyjnej genomu i transkryptomu. Wyniki wieńczy podrozdział o związku analizy klastrowania w populacji polskiej z fenotypem. Analiza ta była możliwa dzięki danym ankietowym dotyczącym chorób cywilizacyjnych zebranych od 5656 osób z badanej kohorty. Stwierdzono, że dystrybucja wariantów była istotnie związana z cukrzycą.

W mojej opinii wyniki przedstawione są w sposób uporządkowany i logiczny. Wielkość grupy badanej jest imponująca, a zastosowane metody analizy i umiejętnie wykorzystane zasoby ogólnodostępnych baz danych sprawiają, że uzyskane wyniki są unikalne.

Do tej części pracy miałabym trzy pytania:

- 1. Jakie czynniki wzięto pod uwagę przy oszacowaniu wielkości poszczególnych grup badanych na etapie projektowania badania?*
- 2. Nie znalazłam w pracy numerów zgłoszeń czy informacji o zamieszczeniu w ogólnodostępnych bazach nowych wariantów SNV zidentyfikowanych w ramach recenzowanej pracy doktorskiej. Dzielenie się ze środowiskiem naukowym i klinicznym informacjami o nieodnotowanych wcześniej wariantach jest ważnym aspektem prowadzenia pracy badawczej. Rozwój baz danych zmienności genetycznej człowieka warunkowany jest bowiem zgłoszeniami z lokalnych laboratoriów na całym świecie. Proszę Doktoranta o wyjaśnienie, czy takie zgłoszenia zostały wykonane lub są zaplanowane.*
- 3. Zastosowanie macierzy całogenomowych SNP pozwoliło na wygenerowanie ogromnych zasobów danych dotyczących zmienności tysięcy genów. Choć to absolutnie*

zrozumiałe (i konieczne), że w rozprawie wykorzystano zdefiniowaną część danych, w tym przypadku dotyczącą rodziny genów ABC, z czystej ciekawości zapytam czy zaplanowano na przyszłość analizę także innych genów istotnych z punktu widzenia rozwoju medycyny precyzyjnej?

W „Dyskusji”, obejmującej 8 stron, Autor przystąpił do interpretacji uzyskanych rezultatów w odniesieniu do danych z piśmiennictwa. Doktorant odniósł się do liczby wariantów częstych i rzadkich identyfikowanych zarówno w populacji polskiej, jak i na podstawie bazy *1000 Genom Project*. Skomentował wzbogacenie w rzadkie warianty w genach kodujących ABC w świetle dotychczas opublikowanych badań. Odniósł się do możliwego podłoża obserwowanego klastrowania identyfikując jako istotny czynnik zmienność genów ABC kodujących białka zaangażowane w homeostazę kwasów żółciowych. W dyskusji analizy fenotypowej dla poszczególnych klastrów w populacji polskiej podkreślił rolę niektórych białek ABC w rozwoju chorób cywilizacyjnych związanych z cukrzycą i otyłością.

Dyskutując wyniki własnych badań Doktorant zachował stosowny krytycyzm i z dużą ostrożnością odniósł je do danych z piśmiennictwa. Na uwagę zasługuje solidne przedyskutowanie ograniczeń badania, z jakimi mierzył się Autor rozprawy. Między innymi słusznie zauważył, że zastosowana metodyka w zakresie identyfikacji wariantów była ograniczona zdefiniowaną przez producenta mikromacierzy pulą specyficznych wariantów. Stąd identyfikowana zmienność o charakterze rzadkim może mieć wartość zaniżoną. Autor wskazuje więc na zasadność reanalizy pełnych sekwencji wybranych genów ABC metodą HRM. Całokształt dyskusji dowodzi, iż Doktorant wykazał się dużą dojrzałością naukową i potrafił krytycznie i z dystansem spojrzeć na wyniki swojej pracy. Jestem przekonana, że wyniki zaprezentowanych w rozprawie badań, mają szansę na opublikowanie w dobrym czasopiśmie o wysokim współczynniku oddziaływania (IF).

Uzyskane w toku badań wyniki oraz ich analiza w oparciu o dane z ogólnodostępnych baz danych i piśmiennictwa były podstawą do sformułowania przez Doktoranta wniosków i weryfikacji hipotez badawczych. Wnioski zostały sformułowane poprawnie i w całości odpowiadają powziętym do realizacji celom i postawionym hipotezom.

W dysertacji Doktorant odniósł się do szerokiego, trafnie dobranego piśmiennictwa z uwzględnieniem najnowszych badań populacyjnych dotyczących zmienności genetycznej. Praca została napisana poprawną polszczyzną, z dbałością o szczegóły. Cechuje ją poprawność stylistyczna, składniowa, ortograficzna i interpunkcyjna, co sprawia, że czyta się ją z łatwością i dużym zainteresowaniem. Przed publikacją w języku polskim proponuję jednak uważną redakcję tekstu pod kątem słowa „ilość” niepoprawnie użytego w kontekście rzeczowników policzalnych jak warianty czy klastry i zastąpienie go słowem liczba, czyli „liczba wariantów”, a nie „ilość wariantów”. Kolejna drobna uwaga dotyczy stosowania pojęcia „polimorfizm”, którego Autor użył kilkakrotnie w pracy, choć zdecydowanie częściej używał zgodnie z obecnie

obowiązującymi wytycznymi nomenklaturowymi HGVS słowa „wariant”. Powyższe uwagi nie wpływają jednak na zrozumiałość przekazu i ogólny odbiór tej bardzo dobrze napisanej pracy. Tekst pracy uzupełniają i wzbogacają czytelne ryciny oraz niezwykle starannie przygotowane tabele. Sposób redakcji tekstu, tabel i rycin uważam za właściwy, a wszelkie drobne uwagi krytyczne pozostają bez wpływu na całościową, pozytywną ocenę pracy.

Na zakończenie recenzji chciałabym wyraźnie podkreślić, że rozprawę doktorską Pana mgr Marcina Słomki oceniam bardzo wysoko. Uzyskane w toku realizacji pracy rezultaty mają istotne znaczenie dla nauki, a także potencjalnie dla praktyki klinicznej. Pragnę zauważyć, że rozprawa stanowi oryginalne rozwiązanie problemu naukowego, a wyniki rozprawy stanowią rzeczywisty wkład w rozwój wiedzy w reprezentowanej przez Doktoranta dziedzinie badań. Choć praca w założeniu nie ma charakteru aplikacyjnego, to wyniki badań dostarczają istotnych danych o zmienności genów ABC w populacji polskiej i mogą przyczynić się do rozwoju medycyny spersonalizowanej.

W podsumowaniu stwierdzam, że przedstawiona mi do oceny rozprawa doktorska Pana mgr Marcina Słomki spełnia wszelkie wymogi stawiane rozprawom na stopień naukowy doktora określone art. 13 ust. 1 ustawy z dnia 14 marca 2003 r. o stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki (Dz. U. 2017 poz.1789) i zwracam się do Wysokiej Komisji Uniwersytetu Łódzkiego o dopuszczenie Pana mgr Marcina Słomki do dalszych etapów przewodu doktorskiego.

Jednocześnie wnioskuję do Komisji Uniwersytetu Łódzkiego do spraw stopni naukowych o wyróżnienie pracy doktorskiej Pana mgr Marcina Słomki z uwagi na wyróżniającą się jakość badań, nowatorskie metody badawcze i szczególne walory poznawcze. Pragnę podkreślić, że badania takie nie były dotąd przeprowadzone w populacji polskiej. Doktorant z sukcesem wykonał szeroki zakres nowoczesnej analizy genetycznej przy imponującej wielkości grupy badanej. Wyniki przeprowadzonych badań są istotnym elementem charakteryzacji zmienności międzypopulacyjnej i międzypopulacyjnej dla populacji polskiej i stanowią ważny krok naprzód w rozwoju medycyny precyzyjnej. W mojej ocenie praca Pana mgr Marcina Słomki zdecydowanie wykracza ponad wymagania stawiane rozprawom doktorskim.

Z wyrazami szacunku

KIEROWNIK ZAKŁADU GENETYKI
Instytutu "Centrum Zdrowia Matki Polki"
Agnieszka Gach
dr hab. n. med. Agnieszka Gach
profesor instytutu

Dr hab. n. med. Agnieszka Gach, profesor ICZMP

