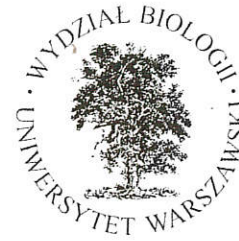




UNIwersytet
Warszawski

Wydział Biologii
Instytut Genetyki i Biotechnologii
prof. dr hab. Paweł Golik



Warszawa, 08.09.2021

Recenzja rozprawy doktorskiej mgr Magdaleny Kukli-Bartoszek pt. „*Genetic prediction of human pigmentation characteristics – implementation of the HirisPlex-S predictive tool and searching for improvements*”

Pomimo tego, że od opublikowania pierwszej wersji sekwencji genomu człowieka minęły już dwie dekady, a od kilkunastu lat dysponujemy technikami wysokowydajnego sekwencjonowania i charakteryzowania zmienności genetycznej, pozwalającymi na gromadzenie danych o różnorodności genotypowej na skalę genomu dla coraz większych prób populacyjnych, wciąż daleko jesteśmy od zrealizowania podstawowego programu badawczego genetyki, jakim jest zrozumienie tego, jak genotyp determinuje fenotyp. Praktycznie wszystkie cechy zmienności prawidłowej człowieka, podobnie jak większość często występujących chorób, są to cechy wieloczynnikowe, których odziedziczalna składowa obejmuje zmienność dziesiątków, a częściej nawet setek różnych genów. Zrozumienie genetyki cech wieloczynnikowych jest obecnie największym wyzwaniem, przed którym stoi genetyka, wpisując się w ogólny problem nauk przyrodniczych XXI. wieku, jakim jest opisanie systemów złożonych, o dużej liczbie składowych połączonych licznymi, nieliniowymi interakcjami.

Tematyka, której dotyczy recenzowana praca doktorska mgr Magdaleny Kukli-Bartoszek dotyczy zatem bardzo aktualnego i dalekiego wciąż od rozwiązania zagadnienia stanowiącego rdzeń problematyki badań genetycznych. Problematyka badawcza została zatem wybrana właściwie, i dotyczy zagadnienia niezwykle istotnego zarówno od strony poznawczej, jak i praktycznej. Tytuł pracy jest precyzyjnie sformułowany i dobrze oddaje istotę jej treści.

Zgodnie z nowelizacją ustawy o stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki z dnia 18 marca 2011 (art. 13, ust. 2), praca ma formę zbioru sześciu publikacji z lat 2018 – 2021 opublikowanych w czasopismach ujętych w wykazie czasopism naukowych Ministerstwa Nauki i Szkolnictwa Wyższego (część A). Pięć z nich opublikowano w czasopiśmie *Forensic Science International. Genetics* (I.F. 4,882), zaś jedną w *International Journal of Legal Medicine* (I.F. 2,686). Są to bardzo dobre czasopisma specjalistyczne, lokujące się w pierwszym

kwartylu współczynnika oddziaływania. Taki dorobek publikacyjny na etapie studiów doktoranckich należy uznać za wybitny, porównywalny wręcz z niejednym osiągnięciem przedstawianym przy ubieganiu się o wyższe stopnie naukowe. Są to prace wieloautorskie, częściowo powstałe także we współpracy międzynarodowej, co jest standardem w genomice człowieka, w której to dziedzinie oryginalnych prac jednoautorskich praktycznie się nie spotyka. Jest to też kolejny dowód na to, że badania Doktorantki mają poziom światowy i mieszczą się w głównym nurcie podejmowanej tematyki. W czterech spośród tych artykułów Doktorantka jest pierwszą (wiodącą) autorką, a udział pozostałych autorów jest udokumentowany odpowiednimi oświadczeniami. Doktorantka uczestniczyła w planowaniu i przeprowadzaniu doświadczeń, analizie danych, formułowaniu wniosków, oraz przygotowywaniu manuskryptów. Udział Doktorantki w przedstawionych publikacjach nie budzi zatem żadnych wątpliwości, a włączenie ich do rozprawy doktorskiej jest całkowicie uzasadnione. Jest to bez wątpienia oryginalny dorobek z nawiązką wystarczający by stanowić podstawę do nadania stopnia doktora. Na podstawie przedstawionej rozprawy mogę stwierdzić, że Doktorantka wykazała się ogólną wiedzą teoretyczną i umiejętnościami praktycznymi w dziedzinie genetyki sądowej i genomiki człowieka, a także umiejętnością samodzielnego prowadzenia oryginalnych badań naukowych.

Żadnych wątpliwości nie budzi też to, że włączone do rozprawy doktorskiej artykuły stanowią zbiór spójny tematycznie. Podejmują szereg zagadnień związanych z przewidywaniem cech pigmentacji człowieka na podstawie danych genetycznych. W pracy Kukła-Bartoszek i wsp. z 2020 r. główny problem badawczy dotyczy wstępnego, lecz fundamentalnego dla genetyki sądowej zagadnienia, jakim jest uzyskanie wiarygodnych danych z materiału o różnym, często znacznym stopniu degradacji. Uzyskane wyniki wykazały, że metody sekwencjonowania NGS mają tu istotną przewagę nad wcześniej stosowanymi technikami opartymi na kapilarnej elektroforezie amplikonów. Artykuł ten koncentruje się, podobnie jak pozostałe w dorobku Doktorantki, na zastosowaniach w predykcji dla celów sądowych (FDP), warto jednak choćby wspomnieć o tym, że podobne problemy i metody ich rozwiązywania mają ogromne znaczenie dla badań kopalnego DNA (aDNA) w paleogenomice i archeologii genetycznej. Ogólnie, we wszystkich artykułach nieco brakuje dyskusji tego, jakie może być znaczenie uzyskanych rezultatów poza ściśle aplikacyjnym polem predykcji w kryminalistyce, jest to jednak zapewne związane z wymaganiami narzuconymi przez profil czasopism, w których zostały opublikowane.

Rdzeniem pracy są artykuły poświęcone udoskonalaniu panelu HIrisPlex-S, stosowanego do przewidywania koloru oczu, włosów i skóry. W pierwszej chronologicznie publikacji (Chaitanya i wsp. 2018) celem było wdrożenie, walidacja i praktyczna aplikacja tego narzędzia, oryginalnie wykorzystującego kapilarną elektroforezę amplikonów. Już w kolejnej pracy (Breslin i wsp. 2019)

narzędzie to zostało zaadaptowane do zapewniających większą przepustowość metod sekwencjonowania NGS. Jak wykazano to we wspomnianej już pracy z 2020 r., pozwoliło to nie tylko na zwiększenie przepustowości analiz, ale także na skuteczniejsze pozyskiwanie danych z kości o znacznym stopniu degradacji. W tym miejscu, nie będąc specjalistą w dziedzinie genetyki sądowej, zastanawiam się, dlaczego wprowadzenie metod NGS do tych narzędzi zajęło tak wiele czasu, podczas gdy w (lepiej mi znanej) paleogenomice ich skuteczność w analizowaniu trudnego, mocno zdegradowanego materiału aDNA została dobitnie wykazana już sukcesem programu sekwencjonowania genomu neandertalczyka w 2006 r.

Praca Kukla-Bartoszek i wsp. z 2018 r. dotyczy kolejnego fundamentalnego problemu genetyki. Fenotyp nie jest w prosty sposób determinowany przez genotyp, istotną rolę odgrywa też szeroko pojmowane środowisko i efekty związane z plastycznością fenotypową. Oczywistym przykładem będzie wpływ ekspozycji na słońce na pigmentację skóry (czy uwzględniany w przewidywaniach systemu HIrisPlex-S?), a także będące przedmiotem omawianego artykułu efekty związane z wiekiem. Wyniki pracy wykazują, że ciemnienie włosów z wiekiem istotnie wpływa na dokładność przewidywania fenotypu koloru włosów. W tym miejscu nasuwa się pytanie, czy połączenie systemu HIrisPlex-S z przewidywaniem wieku na podstawie markerów epigenetycznych nie byłoby optymalne w przypadku próbek osób, których wiek nie jest znany.

Kolejna praca (Kukla-Bartoszek i wsp. 2019) poświęcona jest stworzeniu modelu pozwalającego na predykcję występowania piegów. Praca ta zasługuje na uwagę, gdyż w jej przypadku nie wystarczyło udoskonalenie wcześniej znanego narzędzia takiego, jak HIrisPlex – opisany model jest oryginalnym osiągnięciem tej pracy, w której Doktorantka jest wiodącą autorką.

Ostatni artykuł, opublikowany w tym roku, jest również bardzo interesujący, gdyż podejmuje się poszukiwania ulepszeń w finalnym etapie tworzenia narzędzia przewidywania fenotypu – matematycznej i statystycznej analizie danych. To właśnie ten etap jest kluczowy dla osiągnięcia postępów w badaniu genetyki cech wieloczynnikowych. Współczesne techniki NGS pozwalają bez większego trudu zbierać dane dotyczące zmienności genetycznej, obecnie na skalę kompletnych sekwencji. Wyzwaniem jest właśnie etap analizy – przetworzenia danych na model pozwalający na przewidywanie fenotypu, a może i zrozumienie mechanizmów jego powstania. Zastosowanie algorytmów uczenia maszynowego (ML) jest obecnie bardzo dynamicznie rozwijanym wątkiem w tym polu badań. Kolejnym interesującym wątkiem tego artykułu jest zastosowanie sekwencjonowania eksomowego dla identyfikacji nowych wariantów. Wyniki wykazały, że do pełnego wykorzystania potencjału metod ML konieczne jest zastosowanie większych zbiorów danych, nie mam jednak wątpliwości, że przyszłością genetyki cech wieloczynnikowych jest właśnie zastosowanie zaawansowanych technik algorytmicznych w połączeniu z kompletnymi sekwencjami

eksomu, a nawet genomu. W tym miejscu narzuca się też pytanie, czy algorytmy ML pozwoliłyby na dalsze ulepszenie przewidywań opartych na omawianych w pozostałych artykułach danych pochodzących z zastosowania narzędzia HRisPlex-S (albo opisanego modelu przewidywania występowania piegów), które posługuje się bardziej tradycyjnymi modelami wielomianowymi.

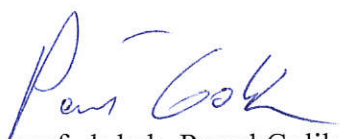
Ogólnie poziom naukowy wchodzących w skład recenzowanej pracy publikacji jest bardzo wysoki, a po przejściu przez sito recenzji w cenionych światowych czasopismach specjalistycznych nie zawierają one niedociągnięć, które byłbym w stanie wychwycić w mojej lekturze. Poza dodatkowymi pytaniami, które zawarłem we wcześniejszych akapitach tej recenzji, nasuwa mi się wyłącznie jedna refleksja natury ogólnej. Przedstawione wyniki bez wątpienia przyczyniły się do znacznego udoskonalenia metod przewidywania fenotypu pigmentacji u człowieka, a niekiedy wręcz do stworzenia nowych modeli tego typu. Pod tym względem praca niewątpliwie stanowi rozwiązanie istotnego problemu naukowego. Jednak we mnie, genetyku nie związanym bezpośrednio z dziedziną genetyki kryminalistycznej, budzą one pewien niedosyt. Przedstawione analizy pozwalają osiągnąć praktyczny, aplikacyjny cel, jakim jest coraz dokładniejsze przewidywanie cech wyglądu zewnętrznego człowieka na podstawie analiz DNA. Mam jednak wrażenie, że nadal bardzo niewiele wiemy i rozumiemy w kwestii tego, jak ten fenotyp powstaje.

Poza krótkimi wzmiankami w autorskim wstępie, brakuje mi próby systemowego powiązania tego, co wiadomo o funkcji genów, których zmienność pozwala na przewidywanie badanych cech, z tym jak cechy te się wykształcają. Opis architektury genetycznej cechy wieloczynnikowej ogranicza się obecnie do wyliczenia genów, które na fenotyp tej cechy wpływają. Elementy analiz mechanistycznych dotyczą pojedynczych genów, głównie *MC1R* i kilku innych (np. *TYR* czy *HERC2*). Wciąż brakuje systemowego modelu, pozwalającego na zrozumienie powiązań pomiędzy funkcjami poszczególnych genów, a fenotypem. Nie jest to, w żadnym wypadku, zarzut wobec prac Doktorantki – jest to opis aktualnej sytuacji całej dziedziny genetyki cech wieloczynnikowych.

Wspominając we wstępie o problemie niedostatecznego wyjaśnienia odziedziczalności fenotypu koloru oczu przez znane geny, Autorka sugeruje (ślądem wielu innych publikacji), że istotną rolę odgrywać mogą nieznanne jeszcze warianty, być może rzadkie, albo mające niewielką cząstkową penetrację, które mogą zostać ujawnione przy zastosowaniu sekwencjonowania całogenomowego na bardzo dużą skalę. Mam przeczucie, że kolejnym brakującym elementem tej układanki są epistatyczne interakcje pomiędzy różnymi genami. Biorąc pod uwagę, że architektura genetyczna cech zmienności człowieka obejmuje dziesiątki, a nawet setki genów dla każdej z nich (około 700 w przypadku wzrostu), oraz to że całkowita liczba kodujących białka genów w genomie nieznacznie przekracza 20 tysięcy, oczywiste jest, że mamy do czynienia z usieciowanym systemem, w którym powszechna będzie plejotropia i epistaza (*sensu lato*). Bez próby stworzenia modelu takiej sieci nie

zbliżymy się do **zrozumienia** genetyki, niezależnie od postępów w praktycznych aplikacjach. Nie jest to w żadnym wypadku zarzut wobec recenzowanej pracy, bardziej wskazanie wyzwania, z którym szczególnie uzdolnione badaczki i badacze, jak Doktorantka, będą musieli się zmierzyć w przyszłości tej dyscypliny.

W podsumowaniu stwierdzam, że przedstawiona mi do recenzji praca doktorska Pani mgr Magdaleny Kukli-Bartoszek pt. „*Genetic prediction of human pigmentation characteristics – implementation of the HirisPlex-S predictive tool and searching for improvements*” stanowi oryginalne rozwiązanie istotnego problemu naukowego oraz została przedstawiona w formie spójnego tematycznie zbioru artykułów opublikowanych w czasopismach naukowych, opatrzonego streszczeniem w języku angielskim i polskim, z dołączonymi oświadczeniami współautorów publikacji. Praca spełnia zatem wymogi art. 13 ustawy o stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki z dnia 18 marca 2011, i może stanowić podstawę do nadania stopnia doktora. Zwracam się zatem do Rady Dyscypliny Nauki Biologiczne Uniwersytetu Jagiellońskiego o dopuszczenie mgr Magdaleny Kukli-Bartoszek do dalszych etapów postępowania o nadanie stopnia doktora w dziedzinie nauk ścisłych i przyrodniczych w dyscyplinie nauki biologiczne. Ze względu na znakomity dorobek publikacyjny i bardzo wysoki poziom merytoryczny recenzowanej pracy zgłaszam też wniosek o wyróżnienie jej w przyjęty przez Radę sposób.



prof. dr hab. Paweł Golik