Poznań, 20.12.2023

prof. dr hab. Paweł Krajewski
Instytut Genetyki Roślin PAN w Poznaniu

Ocena pracy doktorskiej mgr. inż. Mateusza Sypniewskiego
„Podstawy całogenomowej oceny ryzyka otyłości psa domowego"

Niniejsza ocena została wykonana na podstawie zlecenia Przewodniczącej Rady Naukowej Dyscypliny Zootechnika i Rybactwo prof. UPP dr hab. Katarzyny Szkudelskiej, zgodnego z uchwałą Rady o powołaniu recenzentów w przewodzie doktorskim mgr. inż. Mateusza Sypniewskigo nr 01/10/2023 z dnia 27.10.2023 r.

1. Podstawowe dane o kandydacie

Kandydat uzyskał tytułu magistra w dniu 14 czerwca 2018 r. na Uniwersytecie Przyrodniczym w Poznaniu, na Wydziale Rolnictwa i Bioinżynierii. Nie ubiegał się uprzednio o nadanie stopnia doktora. Jest obecnie zatrudniony w firmie Seargin dla Roche (od października 2022) na stanowisku inżyniera bioinformatyka. Jego poprzednie miejsca pracy to: Centralny Szpital Kliniczny MSWiA w Warszawie (lipiec 2022 - marzec 2023), stanowisko bioinformatyk; MNM Diagnostics (maj 2021 - czerwiec 2022), stanowisko bioinformatyk; Uniwersytet Przyrodniczy w Poznaniu (listopad 2018 - maj 2023), projekt OPUS 2016/23/B/NZ2/01762 (lipiec 2020 - sierpień 2021), stanowisko bioinformatyk. Kandydat posiada doświadczenie dydaktyczne nabyte poprzez prowadzenie zajęć dla studentów z analizy statystycznej, bioinformatyki oraz podstaw programowania w języku R.

2. Zakres pracy

Praca doktorska mgr. inż. Mateusza Sypniewskiego opisuje wyniki badań zależności pomiędzy otyłością psów rasy labrador retriever a polimorfizmem ich genomów. Celem było znalezienie nowych polimorfizmów związanych z cechami fenotypowymi psów. W badaniach zastosowano metody statystyczne analizy asocjacyjnej. Wykorzystano dane genetyczne oraz dane fenotypowe uzyskane w badaniach wcześniejszych.

3. Opis formalny pracy

Praca składa się z dwóch części: pierwszej, o układzie typowym dla prac doktorskich, zawierającej informacje wstępne, sformułowanie hipotez, opis materiałów i metod, wyniki, dyskusję, wnioski i bibliografię, oraz drugiej, zwierającej dwa raporty naukowe opisujące przeprowadzone badania, przy czym jeden z tych raportów to artykuł już opublikowany w czasopiśmie Veterinary Sciences, a drugi to artykuł przesłany do czasopisma i znajdujący się obecnie w fazie recenzowania. Współautorami obu

artykułów są kandydat oraz promotor pracy doktorskiej. Opisy przedstawione w części pierwszej pokrywają się w dużej części z odpowiednimi fragmentami artykułów przedstawionych w części drugiej. Stosując terminologię podaną w Ustawie Prawo o Szkolnictwie Wyższym i Nauce (tzw. Ustawa 2.0, art. 187 p. 3), przedstawiona praca jest "pracą pisemną" uzupełnioną o "zbiór powiązanych tematycznie artykułów naukowych", i jako całość opisuje rozwiązanie problemu naukowego z dziedziny genetyki zwierząt. Część pierwsza mieści się na 53 stronach, zawiera dwie tabele i cztery ryciny. Spis literatury zawiera 102 pozycje.

4. Ocena sposobu realizacji

Rozwijane obecnie metody sekwencjonowania genomów pozwalają na szeroko zakrojone badania asocjacyjne zmierzające do wyjaśnienia podłoża genetycznego interesujących cech fenotypowych organizmów poprzez podanie zależności tych cech od polimorfizmu gęsto rozmieszczonych w genomie markerów. Recenzowana rozprawa opisuje badania wpisujące się w ten nurt.

Do badań wybrano populację 50 osobników rasy labrador retriever o genomach zsekwencjonowanych w ramach wcześniejszych badań i o cechach fenotypowych określonych w ramach wizyt w klinikach weterynaryjnych. Przedmiotem zainteresowania były dwie cechy: masa ciała wyrażona w kilogramach i wskaźnik BCS (body condition score) opisujący stopień otłuszczenia ciała w skali pięciostopniowej. Dane genotypowe dotyczyły polimorfizmów typu SNP i indel w 41 genach określonych jako ortologi genów związanych z otyłością u człowieka. Wybrano modele statystyczne analizy asocjacyjnej odpowiednie dla badanych cech: liniowy model mieszany dla masy ciała i model regresji logistycznej dla BCS.

Uzyskano wyniki świadczące o istotnym statystycznie związku jednego markera w genie ADCY3 z masą ciała oraz dwóch markerów w tym samym genie z oceną BCS. Wszystkie znalezione markery zlokalizowane są w intronach genu. Przeprowadzona analiza nie wykazała istotnego związku z badanymi cechami polimorfizmu w genie POMC, który został uznany za istotnie związany z otyłością w badaniach innych autorów. Uwzględnienie tego polimorfizmu w modelach analizy asocjacyjnej nie zmieniło znacząco uzyskanych wyników.

Przedstawienie wyników badań poprzedzone jest szerokim wstępem do tematyki przedstawiającym czynniki genetyczne i pozagenetyczne wpływające na otyłość psów oraz zjawiska otyłości mono- i wielogenowej. Czynione to jest na tle odpowiednich wyników badań prowadzonych na ludziach i myszach.

W dyskusji wyników stwierdza się, że rezultaty pracy mają znaczenie dla rozumienia plejotropowego działana genów w odniesieniu do różnych wskaźników otyłości. Dyskutuje się możliwe zaburzenia obserwacji cech fenotypowych, szczególnie cechy BCS, której obserwacja może być do pewnego stopnia subiektywna. Rozważa się możliwy wpływ niejednolitej struktury populacji na uzyskane wyniki. Następnie analizuje się możliwy sens biologiczny wyników od strony funkcji produktu genu ADCY3 opisanych w innych badaniach. Ponadto, rozważa się możliwość udziału wariantów intronowych w determinacji procesów, w szczególności w przebiegu splicingu.

Należy stwierdzić, że sformułowany cel pracy został osiągnięty. Podstawowa hipoteza o tym, że w genomie psów rozważanej rasy istnieją nieznane jeszcze mutacje wpływające na ryzyko otyłości i występujące z częstościami pozwalającymi na ich badanie została zweryfikowana pozytywnie. Polimorfizmy sekwencji jednego z uwzględnionych w badaniach genów wykazały związek z badanymi cechami fenotypowymi. Dla cechy masa ciała znaleziono polimorfizm w genie ADCY3 o efekcie allelicznym równym 5.56 kg, a więc znacznie większym niż efekt polimorfizmu (delecji) w genie POMC określony na 1.90 kg w badaniach innych autorów.

5. Uwagi i pytania

1. Ponieważ w pracy analizowano polimorfizm 41 wybranych genów, tytuł zawierający określenie "całogenomowa ocena" wydaje się zbyt szeroki.

2. Na str. 29 w dyskusji wyników analizy BCS stwierdza się, że postanowiono "ograniczyć wpływ struktury danych (populacji) przez zastosowanie metody LD pruning". Selekcja markerów poprzez "LD pruning" poprawia jakość oszacowania struktury populacji, natomiast nie poprawia wnioskowania o asocjacjach ze względu na tę strukturę, co stwierdza się, lecz trochę nie wprost, w zdaniu "W metodzie tej wpływ ... na jakość wnioskowania statystycznego nie jest ograniczany". Tak więc analiza cechy BCS w modelu regresji logistycznej nie brała pod uwagę struktury populacji. Jakie "trudności obliczeniowe" powodowało uwzględnienie w modelu składowych głównych? Czy istnieje jakaś metoda pozwalająca na uwzględnienie struktury populacji w modelu regresji logistycznej?

3. O istotnym związku markerów z cechą BCS w dużej mierze decydowała nierównowaga osobników heterozygotycznych w grupach przypadków i osobników zdrowych (4 / 25) czyli z jedną kopią allelu korzystnego T. Jakie ma to znaczenie dla uzyskanych wyników i ich znaczenie praktycznego?

4. Cechę BCS poddano "dualizacji" przez zaliczenie psów o wartości "5" do grupy otyłych, zaś reszty do grupy kontrolnej. Czy przeprowadzono analizę asocjacyjną przy innej dualizacji cechy i porównano wyniki obliczeń?

5. Czy analizowano - w jakimś sensie - korelację masy ciała i BCS? Jest ona decydująca dla wniosku na str. 32 "gen okazał się istotny w obydwu badaniach pomimo zastosowania innych fenotypów i metod obliczeniowych".

6. Jakie były parametry progowe przy szukaniu ortologów genów człowieka?

7. Wiele uwag można sformułować pod adresem stosowanego języka i terminologii, szczególnie w zakresie opisu metodologii statystycznej ("family wise error rate" - winno być "family-wise error rate" (str. 20); "skorygowanie p-wartości o FWER" - "wyliczenie wartości FWER" (20); "częstość mniejszego allelu" - "częstość rzadszego allelu" (19); "próg istotności $0,05 < \alpha$" - "próg istotności $\alpha = 0.05$" (22); "redukcji wielkowymiarowych danych (w tym przypadku genomowej macierzy spokrewnień)" - "redukcji wielkowymiarowych danych (w tym przypadku danych markerowych)" (28); "główne składowe jako efekty stałe" - "główne składowe jako zmienne towarzyszące (objaśniające) o efektach stałych" (29); "wpływ struktury danych" - "wpływ struktury populacji" (29).

5. Konkluzja

Recenzowana praca zawiera nowe wyniki poszerzające wiedzę o genetycznej determinacji cech fenotypowych psów. Wskazano gen o prawdopodobnym znaczeniu dla oceny ryzyka otyłości. Jakkolwiek możliwość praktycznego wykorzystania tych wyników zależy od prowadzenia dalszych badań, to jednak wskazanie analogii pomiędzy mechanizmami kształtowania się pewnych cech u ludzi i u psów ma znaczenie naukowe. Przedstawione w pracy wyniki są już częściowo opublikowane, co wskazuje na ich dobrą ocenę przez szersze grono specjalistów.

W związku z tym uważam, że praca spełnia wymagania stawiane przed rozprawami doktorskimi i wnoszę o dopuszczenie mgr. inż. Mateusza Sypniewskiego do dalszych etapów przewodu doktorskiego.

P. Krzewski