

Recenzja rozprawy doktorskiej  
**mgr inż. Tomasza Szmatoły**

pt. „Charakterystyka ciągów homozygotyczności u wybranych ras bydła”

## **Wprowadzenie**

W erze postgenomiczej, dzięki dysponowaniu pełnymi sekwencjami nukleotydowymi wielu gatunków zwierząt gospodarskich, pojawiają się nowe możliwości analizy genetycznej. Odkrycie jednonukleotydowych polimorfizmów (SNP) rozproszonych wzdłuż całego genomu w połączeniu z wysokowydajnymi technikami genotypowania (mikromacierze) stworzyło nową sytuację, w której genetyczna wartość cech może być szacowana bez znajomości fenotypu. Dostęp do tej informacji nie jest ograniczony ani do wieku, ani płci czy stanu fizjologicznego. Ta rewolucyjna zmiana najpełniej wyraża się w selekcji genomowej bydła, która stała się rutynową metodą stosowaną przez hodowców w wielu krajach, w tym także w Polsce.

Poza szacowaniem genomowej wartości hodowlanej, genotypowanie zwierząt umożliwia pozyskanie danych przydatnych do dodatkowych typów analiz, np. kontroli pochodzenia czy wykrywania nosicieli defektów genetycznych. Ponadto, możliwa jest ocena stopnia homozygotyczności określonej populacji zwierząt poprzez użycie tzw. ciągów homozygotyczności (ROH – runs of homozygosity), tj. identycznych serii homozygotycznych SNP oddziedziczonych od obojga rodziców. Przyjmuje się, że długie ROH świadczą o bliskim spokrewnieniu osobników, a krótkie są wynikiem większej liczby crossing-over i dlatego reprezentują dalsze spokrewnienie. Poznanie długości i liczby ROH może być zatem użyteczne w badaniach filogenetycznych, w ocenie chowu wsobnego czy w śledzeniu procesów selekcyjnych, w tym także typowaniu subregionów genomu, w których ujednoczenie genetyczne wyrażone przez ROH świadczy o obecności genów warunkujących cechy użytkowe podlegające selekcji.

W tym właśnie nurcie naukowych eksploracji mieści się dysertacja doktorska Pana mgr inż. Tomasza Szmatoły, który podjął się próby kompleksowej charakterystyki ROH dla 11 ras bydła hodowanego w Polsce. Ze względu na oryginalny cel pracy, unikalność zastosowanych danych oraz nowoczesny charakter metod badawczych, pracę tę należy uznać za interesującą i wartościową w dyscyplinie naukowej zootechnika.

Całość dysertacji jest napisana zwięzłym i precyzyjnym językiem naukowym, w sposób przejrzysty i logiczny. Praca ma typowy układ klasycznej rozprawy doktorskiej liczącej łącznie 102 strony znormalizowanego maszynopisu zawierającego wstęp, cel pracy, materiały i metody, wyniki badań, dyskusję oraz wnioski. Zawiera także wykaz skrótów, streszczenie oraz spis piśmiennictwa, na który składa się 95 pozycji, ściśle odnoszących się do tematu dysertacji.

Ponadto zaopatrzona jest w materiały uzupełniające ukazujące kształt wzorców ROH dla poszczególnych chromosomów i każdej z badanych ras oraz listę genów leżących w regionach o wysokiej częstości występowania ROH w każdej z ras.

Czytając pracę odnosi się bardzo pozytywne wrażenie wynikające z umiejętnego połączenia aktualnej tematyki i nowoczesnego warsztatu badawczego.

## **Uwagi i komentarze**

### *Tytuł pracy*

Tytuł pracy koresponduje z czterema pierwszymi celami pracy. Ostatnie dwa cele, tj. dotyczące identyfikacji genów kandydujących w regionach o szczególnie wysokim stopniu występowania ROH oraz obliczenie współczynnika inbredu genomowego, nie są zapowiedziane w tytule pracy. Można to zrozumieć jako chęć uniknięcia zbyt długiego tytułu, jednak zasadą nadrzędną powinna być maksymalna adekwatność tytułu do treści. Przykładowo, alternatywny tytuł o brzmieniu „Charakterystyka ciągów homozygotyczności i ich wykorzystanie do szacowania współczynnika inbredu i identyfikacji sygnatur selekcji u wybranych ras bydła”, pełniej oddawałby zawartość pracy.

### *Wstęp*

Wstęp napisany jest w sposób bardzo przejrzysty i kompetentny o czym świadczy trafny i aktualny dobór literatury. Szczególnie docenić należy ten fragment, w którym Autor opisuje techniczne kryteria identyfikacji ROH, takich jak liczba SNP, minimalna długość ROH, minimalne odstępstwa między homozygotycznymi SNP, maksymalna liczba heterozygotycznych SNP w obrębie ROH. Znajomość tych ograniczeń świadczy o dobrym przygotowaniu Autora do wykonania analiz, które przy zbyt dużej dowolności przyjętych kryteriów, mogą wygenerować zupełnie inne wyniki charakteryzujące ROH.

We wstępie Autor nie zamieścił omówienia trzech prac, które z jego udziałem zostały już opublikowane, a które wyraźnie korespondują z wybranymi elementami ocenianej pracy. Chodzi o prace:

Gurgul i wsp. (2016), która ukazała się w *Journal of Animal Breeding and Genetics* pod tytułem “Identification of genome-wide selection signatures in the Limousin beef cattle breed”;

Gurgul i wsp. pt.: “The use of runs of homozygosity for estimation of recent inbreeding in Holstein cattle” (*Journal of Applied Genetics* 2016);

Szmatola i wsp. (2016) pt.: Characteristics of runs of homozygosity in selected cattle breeds maintained in Poland (Livestock Science).

Pierwsza z prac nie została w ogóle zacytowana w dysertacji, do drugiej Autor krótko odniósł się w rozdziale Dyskusja, a trzecia jest zdawkowo przytoczona na stronie 16.

We wszystkich przypadkach jednak brakuje wyraźnego sprecyzowania, w jakim zakresie wyniki już opublikowane są ponownie przytoczone przez Autora w ocenianej pracy, a w jakim stopniu stanowią autonomiczną analizę.

### *Cel pracy*

W rozdziale tym precyzyjnie sformułowano 6 szczegółowych celów. Cele te są nowatorskie i zapowiadają uzyskanie nowej wiedzy natury głównie poznawczej, zwłaszcza w przypadku ras zachowawczych bydła, to jest dla bydła rasy polskiej czerwonej, polskiej czarno-białej, polskiej czerwono-białej i białogrzbietej.

### *Material i metody*

Badania wykonano na 1931 osobnikach reprezentujących 11 ras bydła utrzymywanego w Polsce. Zauważalna jest znaczna dysproporcja w liczebności poszczególnych ras, wahająca się od 77 sztuk u bydła rasy Hereford do 299 sztuk u bydła rasy holsztyńsko-fryzyjskiej odmiany czarno-białej. Ten fragment powinien być nieco lepiej opisany aby uwiarygodnić losowość próby, np. poprzez liczbę stad, z których pochodziły badane zwierzęta, liczbę ojców po których pochodziły badane osobniki czy opis metody losowego wyboru (z jakiej puli i w jaki sposób losowano). Wiadomo, że liczba osobników wpływa na frekwencję genotypów i alleli, a tym przypadku na liczbę i długość ROH. Rasa Hereford, w porównaniu do innych ras, reprezentowana jest najmniej licznie i jednocześnie bardzo wyróżnia się zarówno w liczbie ROH jak i sumie długości ROH. Ta koincydencja wzbudza naturalne podejrzenie, czy rzeczywiście jest to próba reprezentatywna.

W legendzie tabeli 1 powinny być rozwinięte skróty ras. Choć dla większości z nich skrót domyślnie sugeruje pełną nazwę, nie dotyczy to skrótów ZR czy ZB.

Prawidłowo przyjęto wskaźniki filtracji jakości genotypów CallRate (95%), GenCall (0,7) i GenTrain (>0,4), choć w przypadku CallRate jest on inny niż w pracy Szmatola i wsp. (2016). Podobnie inna jest też ostateczna liczba SNPs zastosowana w analizach statystycznych (42787 w niniejszej pracy i powyżej 46 tysięcy w publikacji).

Przyjęta wielkość pięciu kategorii długości ROH jest typowa dla tego typu prac, podobnie jak poziomy dopuszczenia heterozygot oraz liczba nieoznaczonych genotypów w poszczególnych

kategoriach ROH. Za adekwatną także należy uznać metodę obliczania genomowego współczynnika inbredu zaczerpniętą z pracy McQuillana i wsp. (2008), gdyż jest ona najczęściej stosowana przez innych autorów.

Z metodyki pracy wynika, że przyjęto własną metodę identyfikacji regionów genomu o wysokiej częstości występowania ROH, czyli tzw. „ROH hotspots”. Za wyznaczniki takich miejsc uznano te SNPs, które stanowiły zaledwie 1% najwyższych wartości ich wystąpień w danej populacji. To dość wymagające kryterium miało decydujący wpływ na wyniki realizacji celów nr 4 i 5 pracy, a więc identyfikację sygnatur selekcji i genów kandydujących podlegających kierunkowej selekcji. Trzeba w tym miejscu szczególnie docenić logikę i adekwatność zastosowanych metod, szczególnie w zakresie użycia programów komputerowych do bioinformatycznej obróbki danych.

### *Wyniki*

Wyniki badań należy uznać za ciekawe i wartościowe. Po raz pierwszy otrzymaliśmy tak obszerną charakterystyką ROH u większości ras bydła hodowanych w Polsce. Wykresy 1 i 2 oraz tabele 2 i 3 pokazują zarówno różnice między rasami, ale także między typami użytkowymi. Generalnie różnice te nie są zaskakujące, gdyż rasy mleczne znane z większego inbredu charakteryzują się wyższą liczbą ROH a rasy dwukierunkowe, w tym rasy zachowawcze – mniejszą liczbą ROH. Grupa trzech ras typu mięsnego wykazuje wartości pośrednie, z wyjątkiem rasy Hereford, która znacząco odbiega od wszystkich innych.

W tabeli 6 pokazano brak statystycznie istotnych różnic dla rozkładów średnich ilości ROH między badanymi rasami, nawet między rasami Hereford i polskiej czerwonej, gdzie różnice w rozkładzie ilości ROH były wielokrotne (Tabela 5). W przypadku jednak średniej sumy długości ROH w poszczególnych kategoriach ROH (Tabela 7), statystycznie istotne różnice między wieloma rasami zostały stwierdzone (Tabela 8). Wizualne wrażenie powstałe z porównania tych tabel nasuwa pytanie – co jest przyczyną w ocenie statystycznej istotności tych różnic dotyczących zależnych od siebie cech ROH.

Jeśli chodzi o wartości genomowego współczynnika inbredu korespondują one z wcześniejszymi charakterystykami ROH, co klarownie pokazuje Wykres 3 i szczegółowo dokumentuje Tabela 9. Zgodnie z oczekiwaniami najwyższą wartość inbredu zaobserwowano u rasy Hereford, rasy holsztyńsko-fryzyjskiej odmiany czarno-białej i rasy Montbeliarde co potwierdza się dla kilku kategorii ROH (>1 Mb, >2Mb i >4 Mb). Najniższy współczynnik inbredu stwierdzono dla pozostałych ras mięsnych oraz ras zachowawczych. Na stronie 33 wkraśl się prawdopodobnie

błąd – dla tych ras podano zbyt szeroki zakres tego współczynnika od 0,025 do 0,40; powinno chyba być od 0,025 do 0,04.

Przyjęta metoda typowania regionów genomu o najwyższej częstości ROH pozwoliła na stworzenie tzw. wzorców ROH dla każdej z ras, których pozycje wskazano orientacyjnie (Wykres 4 i 5) oraz precyzyjnie (Tabela 10). Wykres 4 pokazuje w zasadzie brak wyraźnych takich regionów u ras zachowawczych, natomiast u ras mlecznych i mięsnych pojawiają się wyraźne piki, szczególnie widoczne u ras Limousin i Charolaise. Trudno dostrzec, aby wszystkie rasy należące do jednego typu użytkowego miały jakiś charakterystyczny wspólny wzorec ROH, chociaż pewne podobieństwo widać dla obu odmian rasy holsztyńsko-fryzyjskiej.

Wykres 5 pokazuje indywidualne chromosomy, dzięki czemu możemy dostrzec czy występują regiony o najwyższej częstości występowania ROH wspólne dla kilku ras. Takie regiony występują na wielu chromosomach, m. in. na chromosomie 6, 14 i 16. Są to regiony stosunkowo dużej wielkości, stąd liczba genów znajdujących się w ich obrębie jest znaczna (od 98 do 250).

Ciekawa wydaje się analiza wskazująca na występowanie genów leżących w obrębie wspólnych ROH hotspots dla kilku ras (Wykres 6). Pewne zaskoczenie budzi fakt braku takich genów dla 4 ras mlecznych i 3 ras mięsnych.

W Tabeli 11 w legendzie użyto terminu „Procesy metaboliczne” zamiast „Procesy biologiczne”. Procesy metaboliczne stanowią jedną z klas procesów biologicznych, wymienioną zresztą w przedostatniej pozycji. Zastanawia także to, dlaczego odmiany tej samej rasy (HO i RW) tak znacząco różnią się liczbą genów zaangażowanych w procesy komórkowe (2-krotnie), lokalizację (ponad 3-krotnie) czy procesy metaboliczne (prawie 2-krotnie).

### *Dyskusja*

Dyskusja własnych wyników wskazuje, że są one ogólnie zbieżne z uzyskanymi przez innych autorów. Brakuje w jej jednak odniesienia do własnych wyników opublikowanych wcześniej. Dyskusja taka jest konieczna, jeśli uzyskane rezultaty mają charakteryzować poszczególne rasy w możliwie aktualny i obiektywny sposób.

Szczególnie konieczne jest szersze omówienie wyników pracy Szmatoła i wsp. (2016) pt.:

“Characteristics of runs of homozygosity in selected cattle breeds maintained in Poland”

(Livestock Science), której tytuł jest niemal identyczny z tytułem dysertacji. Porównanie liczby osobników w dysertacji i rzeczonyj publikacji ujawnia różnice w przypadku rasy holsztyńsko-fryzyjskiej (853 sztuki bez podziału na odmiany wobec 299 dla HO i 231 dla RW), co jak przypuszczam, wpłynęło na uzyskane wyniki badań. Widać to zarówno w tabelach jak i

wykresach. Zachodzi zatem pytanie, które wyniki są bliższe prawdy. Zastanawiające jest to, że Autor nie omawia w Dyskusji tych różnic i nie próbuje ich wyjaśnić.

Podobna uwaga dotyczy pracy Gurgul i wsp. (2016) pod tytułem "Identification of genome-wide selection signatures in the Limousin beef cattle breed". Pracy tej Autor nie zacytował i nie omówił. Czy regiony nazwane przez Autora w tej publikacji jako „selection footprints” pokrywają się z regionami o najwyższej częstości występowania ROH dla rasy Limousin zamieszczonymi w tabeli nr 10 dysertacji? Z metodyki obu prac wynika, że w analizie użyto prawie takiej samej liczby osobników (197 vs 201). Czy były to te same zwierzęta czy inne? Na czym polega użycie ROH do wytypowania ROH hotspots, a na czym tzw. core haplotypes? Odpowiedzi na takie pytania powinny być omówione w Dyskusji w odpowiednim rozdziale ocenianej pracy. Jeśli Autor, poniekąd słusznie, w swojej dysertacji zawarł dużo szerszy zakres analiz obejmujący 11 ras, powinien różnice te omówić uwypuklając oryginalność wyników dotychczas niepublikowanych czy ewentualne korzyści płynące z łącznej analizy.

W związku z tym trudno się zgodzić z ostatnim zdaniem sformułowanym w „Podsumowaniu wstępu”, że „Badania przeprowadzone w niniejszej pracy doktorskiej są pierwszymi wielkoskalowymi badaniami ROH przeprowadzonymi u bydła utrzymywanego w Polsce”.

Można je uznać za słuszne w przypadku 7 ras, natomiast w przypadku 4 innych ras zostały one już opublikowane i niezależnie od różnic trudno uznać je za „pierwsze”.

W podrozdziale dotyczącym ROH jako miary inbredu, przytoczone są prace innych autorów świadczące o dość wysokim wskaźniku korelacji między genomycznym współczynnikiem inbredu a tradycyjnym współczynnikiem inbredu opartym na danych rodowodowych. Nie znajduje to potwierdzenia we wcześniejszej pracy z udziałem Autora (Gurgul i wsp. 2016), gdzie te wskaźniki były dużo niższe. Co prawda Autor podaje prawdopodobną przyczynę tej rozbieżności, ale nie konkluduje, która metoda obliczania korelacji jest bardziej właściwa co byłoby pomocne w ocenie praktycznej przydatności genomicznego współczynnika inbredu. Ma to znaczenie w sytuacji upowszechniającej się oceny genomowej bydła, stwarzającej możliwości do rutynowego stosowania ROH w ocenie inbredu. W tym miejscu cenna byłaby opinia Autora, czy inna metoda obliczania korelacji, jest w stanie ewentualnie skorygować w literaturze naukowej wrażenie, że u polskiego bydła holsztyńsko-fryzyjskiego inbred oparty o badania ROH i rodowody jest dość dalece rozbieżny, co może sugerować niską jakość rodowodów bydła w Polsce.

Dyskusja tej części wyników, która dotyczy tzw. ROH hotspot jest przekonująca i w wielu przypadkach znajduje potwierdzenie w badaniach innych autorów. Najciekawszym i nowatorskim rezultatem wydaje się stwierdzenie w obrębie ROH hotspots 13 genów wspólnych dla czterech ras zachowawczych (Wykres 6), które Autor nieco bliżej charakteryzuje na str. 57.

W całej pracy brakuje mi elementu pokazującego jak w czasie mogły zmieniać się zarówno liczba jak i długość ROH u tej samej rasy, co stanowiłoby jeszcze bardziej przekonujący dowód na istnienie sygnatur selekcji w postaci ROH. Taka analiza wymagałaby przebadania subpopulacji osobników tej samej rasy z różnych okresów selekcji. W materiale jakim Autor dysponuje takie analizy są potencjalnie możliwe w dwóch przypadkach. Zważywszy na fakt, że bydło polskie holsztyńsko-fryzyjskie obu odmian powstało poprzez dolew krwi HF do rasy polskiej czarno-białej i polskiej czerwono-białej trwający od ponad 40 lat – nasuwa się sugestia, że być może głębsza analiza porównawcza tych grup rasowych doprowadziłaby do wskazania miejsc w genomie, w których doszło do najbardziej znaczących zmian heterozygotyczności.

#### *Podsumowanie i wnioski*

W końcowej części dysertacji znajdujemy siedem przekonująco sformułowanych konkluzji, które w pełni odzwierciedlają uzyskane wyniki badań.

We wniosku nr 1 brakuje odniesienia, wobec którego Autor ocenia, że występowanie ROH jest „częstym” zjawiskiem w genomach analizowanych ras bydła. Jak pokazują wyniki ta częstość jest zróżnicowana w zależności od rasy, czy grup ras. Wniosek drugi to potwierdza. Oba wnioski można by połączyć w jeden.

Unikałbym sformułowania „rasy komercyjne”, a pozostawił pojęcie – rasy produkcyjne.

#### *Piśmiennictwo*

W wykazie piśmiennictwa nie można znaleźć dwóch zacytowanych w dysertacji pozycji: Willimas i wsp. (2015) na str. 17 i Meszaros i wsp. (2015) na str. 55.

#### **Wniosek końcowy**

Przedstawione w niniejszej recenzji uwagi i komentarze stanowią element naukowego dyskursu. Za nieuzasadnione uznaję zubożenie dyskusji wyników przedstawionych w pracy o wyniki uzyskane przez Autora we wcześniej opublikowanych swoich pracach. Przypuszczam, że to nie pomniejszyłoby wartości merytorycznej ocenianej rozprawy doktorskiej, lecz przeciwnie, pozwoliłoby na lepsze poznanie różnic pojawiających się w tego typu analizach, wynikających z przyjętej metodyki czy nieco innego materiału. Biorąc to pod uwagę, skłaniam się do opinii, że najbardziej oryginalną i wartościową częścią dysertacji jest charakterystyka ROH u kilku ras zachowawczych utrzymywanych w Polsce, ocena ich genomowego współczynnika inbredu oraz funkcjonalna analiza ROH hotspot pozwalająca na lepsze zrozumienie molekularnego podłoża

szczególnych cech tych ras. Niezaprzeczalnymi walorami dysertacji jest nowoczesny warsztat badawczy i precyzyjny opis wyników. Na bazie osiągniętych wyników, skonfrontowanych z najnowszymi badaniami z tego samego zakresu, wyciągnięto przekonujące i trafne wnioski. Dlatego, biorąc pod uwagę całą zawartość pracy stwierdzam, że **przedstawiona do oceny dysertacja doktorska Pana mgr inż. Tomasza Szmatoły jest wartościowym dziełem naukowym, zawierającym szereg wyników rozszerzających wiedzę z zakresu zootechniki, a w szczególności genomiki bydła. W związku z powyższym stwierdzam, że rozprawa doktorska Pana mgr inż. Tomasza Szmatoły pt. „Charakterystyka ciągów homozygotyczności u wybranych ras bydła” spełnia wymagania określone w ustawie z dnia 14 marca 2003 r. o stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki (Dz. U. Nr 55, poz. 595, z późniejszymi zmianami) oraz w Rozporządzeniu Ministra Nauki i Szkolnictwa Wyższego w sprawie szczegółowego trybu i warunków przeprowadzania czynności w przewodzie doktorskim, w postępowaniu habilitacyjnym oraz postępowaniu o nadanie tytułu profesora. Dlatego wnioskuję do Rady Naukowej Instytutu Zootechniki – Państwowego Instytutu Badawczego w Krakowie o dopuszczenie Pana mgr inż. Tomasza Szmatoły do dalszych etapów przewodu doktorskiego.**

Olsztyn, 21 stycznia 2019 r.

Prof. dr hab. Stanisław Kamiński

