SZKOŁA GŁÓWNA
GOSPODARSTWA
WIEJSKIEGO

PODPIS ZAUFANY
PIOTR
GAWROŃSKI
30.10.2025 16:17:11 GMT+1
Dokument podpisany elektronicznie
podpisem zaufanym

Piotr Gawroński

Katedra Genetyki, Hodowli i Biotechnologii Roślin

Warszawa, 30.10.2025

Recenzja pracy doktorskiej Pani mgr. Marii Nuc zatytułowanej „Metodologia statystyczna badania zmian jakościowych transkryptomu roślin w aspekcie interakcji genotypu i środowiska”.

Przedstawiona mi do recenzji praca doktorska została wykonana w Zakładzie Biometrii i Bioinformatyki Instytutu Genetyki Roślin Polskiej Akademii Nauk. Promotorem rozprawy jest prof. dr hab. Paweł Krajewski, od wielu lat prowadzący badania nad rozwojem narzędzi statystycznych, informatycznych i bioinformatycznych służących analizie eksperymentów biologicznych. Niniejsza rozprawa wpisuje się w ten nurt badawczy i dotyczy opracowania metodologii statystycznej badania zmian jakościowych transkryptomu roślin. Badania nad transkryptomem roślin w kontekście interakcji genotyp–środowisko stanowią obecnie jeden z kluczowych kierunków genomiki funkcjonalnej. Dynamiczny rozwój technologii sekwencjonowania RNA, w tym metod długich odczytów, generuje ogromne ilości danych wymagających opracowania nowych narzędzi analitycznych. W tym kontekście rozprawa mgr Marii Nuc wnosi istotny wkład metodologiczny, proponując podejście umożliwiające ocenę jakościowych różnic pomiędzy transkryptomami roślin.

Doktorantka postawiła sobie następujące cele badawcze:

1. Opracowanie nowej metody analizy wyników doświadczeń wieloczynnikowych, w których uzyskano zbiory sekwencji transkryptomicznych.
2. Uwzględnienie w tej metodzie aspektu biologicznego.
3. Sprawdzenie, czy zastosowanie odległości Mash w analizie porównawczej może poprawić jakość wnioskowania.

Warto podkreślić, że cele te odnoszą się do zagadnień słabo rozwiniętych w dotychczasowych badaniach nad transkryptomami. O ile analiza ilościowa RNA-seq jest obecnie standardem, o tyle opracowania dotyczące charakterystyki zmian jakościowych pojawiają się stosunkowo rzadko. W

Szkoła Główna Gospodarstwa Wiejskiego w Warszawie

Instytut Biologii
Katedra Genetyki, Hodowli i Biotechnologii Roślin
ul. Nowoursynowska 159
02–776 Warszawa
+48 22 59 321 51
kghibr@sggw.edu.pl
www.sggw.pl

tym sensie praca mgr Marii Nuc ma charakter pionierski i stanowi próbę poszerzenia klasycznej analizy ekspresji o aspekt strukturalny transkryptomu.

W mojej ocenie Doktorantka przeprowadziła szeroki zakres analiz i opracowała procedury odpowiadające wszystkim założonym celom badawczym. Należy podkreślić, że w większości badań naukowych analiza transkryptomu prowadzona jest z ilościowego punktu widzenia i polega na porównaniu poziomów ekspresji genów pomiędzy próbkami. Analizy jakościowe transkryptomu są znacznie rzadsze i obejmują m.in. identyfikację nowych izoform, RNA chimerycznych, polimorfizmów czy miejsc redagowania RNA. W związku z tym opracowanie metod umożliwiających ilościowy opis różnic jakościowych pomiędzy transkryptomami stanowi istotny i oryginalny kierunek badań podjęty przez Doktorantkę.

Praca doktorska ma klasyczny układ i rozpoczyna się streszczeniem, po którym następują: wstęp, przegląd literatury (zatytułowany „Stan wiedzy"), cele pracy, opis metod, charakterystyka danych doświadczalnych, wyniki, oraz dyskusja z wnioskami. Całość uzupełnia suplement i spis literatury.

Rozdziały *Wstęp* oraz *Stan wiedzy* starannie wprowadzają czytelnika w tematykę badań. Autorka przedstawia aktualny stan wiedzy dotyczący sekwencjonowania transkryptomów, omawia naturę zmian sekwencji nukleotydowych oraz metody badania zmienności nukleotydowej. Uważam jednak, że część ta mogłaby zostać rozszerzona, szczególnie o opis technologii sekwencjonowania trzeciej generacji, które w ostatnich latach rozwijają się bardzo dynamicznie. Na końcu rozdziału Doktorantka wprowadza pojęcie odległości, ze szczególnym uwzględnieniem odległości Mash. Temat ten został zaprezentowany w sposób zrozumiały nawet dla czytelnika niezajmującego się tą dziedziną, co zasługuje na uznanie. Z obowiązku recenzenta muszę jednak zwrócić uwagę na brak ryciny 2.2, mimo jej opisu w tekście. Uwaga ta nie wpływa jednak istotnie na pozytywną ocenę tej części pracy.

Cele pracy zostały jasno i precyzyjnie sformułowane. W części metodycznej Kandydatka opisuje wykorzystane oprogramowanie oraz zaimplementowane metody statystyczne. Zabrakło jednak informacji o zastosowanych parametrach programów, które mogą mieć znaczący wpływ na uzyskane wyniki – chyba że wszystkie analizy wykonano z użyciem ustawień domyślnych. W tym kontekście nasuwa się pytanie, dlaczego Doktorantka zdecydowała się użyć programu SOAPdenovo, a nie jego wersji dedykowanych analizom transkryptomicznym, takich jak SOAPdenovo-Trans lub Trinity, które są szeroko stosowane do składania odczytów w kontigi lub skafoldy. Wybór tego oprogramowania można jednak uznać za świadomy i ciekawy eksperyment metodologiczny, ponieważ umożliwia porównanie wyników uzyskanych przy użyciu narzędzia pierwotnie opracowanego do analiz genomowych w kontekście danych transkryptomicznych.

Szkoła Główna Gospodarstwa Wiejskiego w Warszawie

Instytut Biologii
Katedra Genetyki, Hodowli i Biotechnologii Roślin
ul. Nowoursynowska 159
02–776 Warszawa
+48 22 59 321 51
kghibr@sggw.edu.pl
www.sggw.pl

Dane doświadczalne pochodziły z eksperymentu trójczynnikowego (czynniki: genotyp, dzień, temperatura) przeprowadzonego na jęczmieniu jarym (*Hordeum vulgare*), opisanego w pracy Mikołajczak i in. (2022). Dane uzyskano za pomocą sekwencjonowania na platformie Illumina.

Najistotniejsza część pracy dotyczy analizy wyników. W pierwszym etapie Doktorantka przeprowadziła ocenę jakości uzyskanych odczytów sekwencjonowania, nie wykazując istotnych nieprawidłowości w danych. Zwróciłem jednak uwagę na zaburzenie zawartości GC na początku odczytów (Ryc. 6.1), co często wynika z obecności sekwencji adapterowych. Warto byłoby wyjaśnić, czy po filtracji i przycięciu odczytów problem ten ustąpił.

W kolejnym kroku Doktorantka przeprowadziła składanie odczytów w kontigi i skafoldy, uzyskując bardzo niewielkie różnice między ich długościami. Zastanawia mnie przyczyna tego zjawiska i czy nie należałoby oczekiwać większych różnic. Następnie Doktorantka obliczyła odległości Mash pomiędzy zbiorami danych, osobno dla każdego etapu przetwarzania. Zaobserwowała, że odległości Mash dla skafoldów i kontigów były niemal identyczne, lecz wyraźnie mniejsze niż dla surowych odczytów. Wynik ten sugeruje utratę części k-merów w procesie składania, ponieważ nie wszystkie odczyty mogą być wykorzystane do budowy kontigów i skafoldów. Doktorantka potwierdziła tę interpretację w dyskusji. Warto byłoby także rozważyć wpływ różnej głębokości sekwencjonowania pomiędzy próbkami na wartości odległości Mash — próbki o mniejszej głębokości mogą nie obejmować transkryptów o niskiej ekspresji, co może wpływać na ocenę podobieństwa.

W dalszej części Doktorantka przeprowadziła analizę PCoA oraz jedno- i wieloczynnikową analizę wariancji, wykazując, że największy wpływ na różnice jakościowe między zbiorami transkryptomów miały czynniki *temperatura* i *dzień*. Zaobserwowano również, że etap składania odczytów wyraźnie różnicował zbiory między sobą.

Następnie Autorka nadała swoim analizom wymiar biologiczny, dzieląc zbiory danych według terminów GO Slim opisujących funkcje genów. Wykazała, że zastosowanie odległości Mash w połączeniu z ontologią GO Slim pozwala na wzbogacenie analiz transkryptomicznych o aspekt funkcjonalny. Ten etap pracy stanowi ważny krok w kierunku integracji metod obliczeniowych z interpretacją biologiczną. Przeniesienie punktu ciężkości z porównania czysto technicznego (na poziomie sekwencji) na analizę funkcjonalną pozwala ocenić, czy obserwowane różnice mają znaczenie biologiczne, co jest kluczowym celem współczesnych analiz omicznych. W tej części zabrakło mi jednak bardziej szczegółowego omówienia biologicznych implikacji uzyskanych wyników.

Ostatnia część wyników, moim zdaniem szczególnie cenna, dotyczy analizy różnicowej ekspresji genów z uwzględnieniem różnic jakościowych transkryptomu. Takie podejście pozwala na analizę ilościową ekspresji z korektą uwzględniającą podobieństwo sekwencji, co można zrealizować, wykorzystując algorytm DESeq2 oraz współrzędne główne obliczone na podstawie odległości Mash.

Szkoła Główna Gospodarstwa Wiejskiego w Warszawie

Instytut Biologii
Katedra Genetyki, Hodowli i Biotechnologii Roślin
ul. Nowoursynowska 159
02–776 Warszawa
+48 22 59 321 51
kghibr@sggw.edu.pl
www.sggw.pl

Uważam, że metoda ta może być niezwykle użyteczna w analizach RNA-seq pochodzących z różnych odmian, gdzie różnice sekwencyjne mogą maskować rzeczywiste zmiany ekspresji. Ciekawym uzupełnieniem byłaby analiza genów odfiltrowanych w tym podejściu, by sprawdzić, czy rzeczywiście wykazują one różnice sekwencji między próbami.

Dyskusja jest zwięzła, logiczna i spójna. Doktorantka trafnie odniosła się do zagadnień, które pojawiają się podczas lektury wyników, i w sposób kompetentny porównała swoje obserwacje z aktualnym stanem wiedzy. Wskazała również ograniczenia pracy, co świadczy o jej dojrzałości naukowej. Analiza sposobu prowadzenia badań i prezentacji wyników wskazuje, że mgr Maria Nuc posiada bardzo dobrą znajomość metod statystycznych i bioinformatycznych, a także umiejętność ich krytycznej oceny. W pracy widoczna jest samodzielność badawcza, umiejętność formułowania wniosków oraz świadomość ograniczeń zastosowanych metod.

Podsumowując, uważam, że recenzowana rozprawa stanowi wartościowy wkład w rozwój metod statystycznych do analizy jakościowej transkryptomów. Zastosowane podejście ma również potencjał aplikacyjny. Może znaleźć zastosowanie w badaniach nad aklimatyzacją roślin do stresu środowiskowego, różnorodnością genetyczną odmian czy w analizach porównawczych transkryptomów różnych gatunków. Praca ta dostarcza więc nie tylko wartości teoretycznych, lecz także praktycznych narzędzi analizy danych RNA-seq. Wyniki badań zostały opublikowane w pracy, w której Doktorantka jest pierwszą współautorką. Zgłoszone w recenzji uwagi i pytania nie wpływają na moją ogólnie pozytywną ocenę rozprawy.

Stwierdzam, że przesłana mi do recenzji praca doktorska Pani mgr Marii Nuc stanowi oryginalne rozwiązanie problemu naukowego i spełnia wymagania stawiane rozprawom doktorskim jak i wymagania ustawy z dnia 20 lipca 2018 r. „Prawo o szkolnictwie wyższym i nauce" oraz wnoszę do Rady Naukowej Instytutu Genetyki Roślin Polskiej Akademii Nauk o dopuszczenie Kandydatki do dalszych etapów przewodu doktorskiego.

Dr hab. Piotr Gawroński

Szkoła Główna Gospodarstwa Wiejskiego w Warszawie

Instytut Biologii
Katedra Genetyki, Hodowli i Biotechnologii Roślin
ul. Nowoursynowska 159
02–776 Warszawa
+48 22 59 321 51
kghibr@sggw.edu.pl
www.sggw.pl