

Poznań, 10 października 2024r.

dr hab. Anna Philips, prof. ICHB PAN
Pracownia Bioinformatyki
Instytut Chemii Bioorganicznej PAN

Recenzja

**rozprawy doktorskiej mgra Jana Grzegorza Kosińskiego,
pt. „Identyfikacja i charakterystyka rho-niezależnych terminatorów transkrypcji u bakterii oraz ich
udziału w odpowiedzi na stres antybiotykowy”**

Przedłożona mi do recenzji praca doktorska została wykonana w Zakładzie Biologii Obliczeniowej Instytutu Biologii Molekularnej i Biotechnologii Wydziału Biologii Uniwersytetu im. Adama Mickiewicza w Poznaniu pod kierunkiem prof. UAM dr hab. Marka Żywickiego.

Rozprawa doktorska mgra Jana Kosińskiego została przygotowana w formie manuskryptu autorskiego liczącego 178 stron. Rozprawa ma nietypowy układ prac doktorskich w dziedzinie nauk ścisłych i przyrodniczych. Praca ma hybrydowy charakter, łącząc cechy klasycznej monografii naukowej oraz tematycznie powiązanego zbioru artykułów naukowych. Oprócz rozdziałów w języku polskim, Autor dołączył do rozprawy trzy publikacje naukowe, które stanowią integralną część pracy. Rozprawa rozpoczyna się obszernym *Wprowadzeniem* w problematykę badań, które wzbogacone jest o publikację przeglądową, w której Kandydat jest jednym z czworga pierwszych autorów, opublikowaną w międzynarodowym czasopiśmie naukowym *International Journal of Molecular Sciences* o 5IF: 5,6. Kolejny rozdział zawiera precyzyjnie określone cele pracy (*Cel Pracy*), a następny obejmuje opis stosowanych metod i wykorzystywanych zestawów danych (*Metody*). Wyniki i dyskusja zostały połączone w jeden rozdział, w którym Autor szczegółowo przedstawił opracowane przez siebie narzędzia, które następnie wykorzystał do realizacji celów badawczych niniejszej rozprawy. Dwa z pośród trzech omawianych narzędzi zostały zaprezentowane przy pomocy publikacji naukowej. Jedna z nich przedstawia program do adnotacji interwałów w genomie AGouTI, czasopismo *PLOS Computational Biology* (5IF: 4,3), w której Kandydat jest pierwszym autorem. Druga publikacja, dotycząca programu TERMITe, budzi jednak pewne wątpliwości, gdyż jest to jedynie preprint zamieszczony w popularnym serwisie internetowym, nie będący recenzowaną pracą naukową. W mojej opinii, preprint nie powinien być traktowany jako integralna część rozprawy, a jego dołączenie można uznać za rozwiązanie odbiegające od przyjętych standardów. **Kandydat proszony jest o przedstawienie obecnego statusu tej publikacji w trakcie obrony.** Dysertację wieńczy krótkie *Podsumowanie*. Dodatkowo manuskrypt zawiera streszczenie w języku polskim i angielskim,

informacje o źródłach finansowania, dorobek publikacyjny Autora, wykaz skrótów oraz listę literatury, obejmującą 94 pozycje.

Rozprawa została dobrze przygotowana pod względem językowym i edytorskim. Poza kilkoma uwagami wymienionymi poniżej Recenzent nie zgłasza istotnych zastrzeżeń.

1. Naprzemienne użycie jako separatora dziesiętnej kropki i przecinka, przy czym w j. polskim poprawny jest tylko przecinek.
2. Autor (zazwyczaj) stosuje znak apostrofu (') zamiast prim ('), co jest jednak powszechnie spotykane w literaturze branżowej.
3. Załączniki nie są ponumerowane podług kolejności występowania w tekście, po Zał. 1 następuje Zał. 10.
4. Str.68 "z 7 projektów" - jest 6 opisanych.
5. Str. 85 „W chwili pisanie niniejszego rozdziału” to nie jest informatywne, powinna zostać podana data.

Ocena Merytoryczna

Ogólnym celem niniejszej rozprawy było zbadanie rho-niezależnych terminatorów transkrypcji u bakterii, ze szczególnym uwzględnieniem ich roli w odpowiedzi na stres antybiotykowy u gronkowca złocistego. Rho-niezależna terminacja transkrypcji to mechanizm zakończenia transkrypcji, który opiera się na specyficznych sekwencjach DNA i strukturach RNA. Mechanizm ten wykorzystuje stabilną strukturę spinki do włosów oraz sekwencję poli-uracylową, które powodują zatrzymanie polimerazy RNA i odłączenie transkryptu od matrycy DNA. Badanie rho-niezależnej terminacji transkrypcji u bakterii jest istotne, ponieważ ten mechanizm odgrywa kluczową rolę w regulacji ekspresji genów, co ma bezpośredni wpływ na procesy życiowe bakterii, w tym ich adaptację do środowiska oraz odpowiedź na stres, taki jak działanie antybiotyków. Zrozumienie rho-niezależnej terminacji pozwala identyfikować kluczowe elementy regulacyjne w genomie bakterii, co może prowadzić do odkrycia nowych celów terapeutycznych. W kontekście wzrastającej oporności na antybiotyki, badania te mogą dostarczyć fundamentalnej wiedzy potrzebnej do opracowania nowych klas leków, które precyzyjnie zakłócałyby bakteryjne procesy transkrypcyjne.

Szczegółowe cele pracy obejmowały opracowanie oprogramowania bioinformatycznego umożliwiającego identyfikację i adnotację terminatorów transkrypcji na podstawie danych z sekwencjonowania 3' końców RNA (Term-seq), oraz RNA (RNA-seq), a także na adnotację interwałów w transkryptach z użyciem referencyjnej adnotacji genomowej. Narzędzia te zostały następnie zastosowane do identyfikacji i adnotacji rho-niezależnych terminatorów transkrypcji u różnych gatunków bakterii, ze szczególnym uwzględnieniem gronkowca złocistego. Kolejnym celem było scharakteryzowanie rho-niezależnych terminatorów transkrypcji u gronkowca oraz analiza ich różnic pomiędzy poszczególnymi, spokrewnionymi taksonami. Ostatnim szczegółowym celem była identyfikacja i analiza rho-niezależnych terminatorów transkrypcji w odpowiedzi na stres antybiotykowy u gronkowca.

Podjęcie tematyki rho-niezależnej terminacji transkrypcji u gronkowca jest bardzo uzasadnione, ponieważ wpisuje się w aktualne potrzeby badań nad mechanizmami regulacyjnymi u bakterii, zwłaszcza w kontekście oporności na antybiotyki. Badanie tego procesu pozwala na

pogłębienie wiedzy o sposobach, w jakie bakterie kontrolują ekspresję genów w odpowiedzi na stres, co ma fundamentalne znaczenie w mikrobiologii i medycynie. Praca nad tym zagadnieniem wymagała interdyscyplinarnego podejścia, obejmującego informatykę, statystykę, biologię molekularną i genomikę, co stanowiło wyzwanie naukowe, ale jednocześnie umożliwiło rozwój cennych kompetencji badawczych Kandydata. W perspektywie długoterminowej, wyniki tego doktoratu mogą dostarczyć nowych informacji użytecznych w poszukiwaniu innowacyjnych strategii terapeutycznych, odpowiadając na wyzwania związane ze wzrastającą opornością bakterii na dostępne leki.

W ramach prowadzonych prac Kandydat stworzył narzędzie TERMITE, służące do identyfikacji rho-niezależnych terminatorów transkrypcji w danych typu Term-seq. Do tej pory jedynym narzędziem do tego typu analiz był *termseq-peaks*. TERMITE rozwija algorytm zaimplementowany w *termseq-peaks*, wprowadzając istotne modyfikacje w procedurze wykrywania pików (peak calling) oraz dodając nowy moduł do anotacji zidentyfikowanych końców 3' RNA. Uważam, że modyfikacje wprowadzone w TERMITE są istotne i przyczyniają się do bardziej precyzyjnej identyfikacji rho-niezależnych terminatorów transkrypcji. TERMITE redukuje wpływ genów o wysokiej ekspresji. To podejście może prowadzić do odkrycia bardziej subtelnych sygnałów, które mogłyby zostać przeoczone przy użyciu metod faworyzujących piki o najwyższej wyrazistości. Poważnym ograniczeniem jest jednak możliwość wykorzystania jedynie danych typu Term-seq, które zdają się nie być tak popularne jak RNA-seq. **Jakie ograniczenia związane z bazowaniem TERMITE na danych Term-seq zauważył Pan podczas analizy? Czy uważa Pan, że istnieje potrzeba dalszego rozwinięcia narzędzia w kierunku integracji z wynikami RNA-seq, szczególnie, że są one dostępne dla jednego z protokołów Term-seq?** Częściową odpowiedzią Autora na to ograniczenie jest program *termiRNAtor* który wykorzystuje wyniki uzyskane z TERMITE (jako bardziej wiarygodne) do identyfikacji rho-niezależnych terminatorów transkrypcji w popularnych danych RNA-seq. Autor dokonał walidacji stworzonego narzędzia obliczając jego czułość i precyzję. Tutaj recenzent musi zwrócić uwagę na kilka kwestii: W Tabeli 1 widoczny jest nieporządek – występuje mieszana nomenklatura dla prawdziwych pozytywów (PP i TP), oraz prawdziwych negatywów (oznaczanych zarówno PN, jak i TN). Niemniej jednak czułość i precyzja narzędzia zostały policzone w prawidłowy sposób i zamieszczone na adekwatnym Wykresie 1, w opisie którego zauważalna jest zamiana terminów: zamiast czułości i precyzji pojawia się „specyficzność” (oznaczająca rzeczywiste negatywne przypadki). Wszystkie miary są ze sobą ściśle związane, wprowadza to jednak pewne trudności w interpretacji wyników. *termiRNAtor* przy precyzji na poziomie 0,73 osiąga czułość 0,61. Popularną miarą modelu klasyfikacyjnego jest wskaźnik F1, który to jednak nie został zastosowany przez Autora. Wskaźnik F1 jest średnią harmoniczną precyzji i czułości, co pozwala na zrównoważoną ocenę modelu pod względem jego zdolności do prawidłowego identyfikowania pozytywnych przypadków oraz unikania tych fałszywych. Wysoki wskaźnik F1 wskazuje na dobry balans między precyzją a czułością. Dla wskazanego modelu, jak policzyłam, wartość F1 wynosi 0,66 i wskazuje na umiarkowaną wydajności modelu. **W związku z tym, chciałabym zapytać Doktoranta, jakie podejścia mógłby rozważyć, aby zwiększyć czułość modelu.**

Meritum pracy była identyfikacja i charakterystyka rho-niezależnych terminatorów u gronkowca złocistego w stresie antybiotykowym. Kandydat wykorzystał do tego stworzone przez siebie oprogramowanie i publicznie dostępne dane, zarówno typu Term-seq, jak i RNA-seq. Łącznie zidentyfikował i adnotował 1128 terminatorów transkrypcji, co stanowi obecnie najbardziej

kompleksowy atlas rho-niezależnych terminatorów transkrypcji u gronkowca. 1083 terminatory wykryto w danych Term-seq, a 412 metodą RNA-seq. Obie techniki pozwoliły wspólnie zidentyfikować 367 terminatorów, co oznacza, że 45 terminatorów zostało wykrytych wyłącznie w danych RNA-seq. **Jakie mogą być przyczyny tego, że kilkadziesiąt terminatorów zostało zidentyfikowanych wyłącznie za pomocą RNA-seq, mimo iż główną techniką dla identyfikacji rho-niezależnych terminatorów transkrypcji jest Term-seq? Czy różnice w wynikach między obiema metodami mogą sugerować, że pewne typy terminatorów są specyficzne dla jednej z tych metod, a jeśli tak – co to może oznaczać dla dalszych badań nad atlasem terminatorów gronkowca złocistego i w ogóle bakterii?**

Kandydat porównał wyniki uzyskane dla gronkowca złocistego z wynikami dotyczącymi sześciu innych gatunków bakterii, zarówno gram-dodatnich, jak i gram-ujemnych. Zwrócił uwagę na różnice w liczbie zidentyfikowanych terminatorów, wyjaśniając, że mogą one wynikać z różnic w budowie i wielkości genomów, co Recenzent uważa za trafne spostrzeżenie. Kandydat proponuje w pracy by zatem odnosić liczbę terminatorów do liczby genów danego gatunku, co jest się uzasadnione. Zauważalne są jednak znaczne różnice zarówno w liczbie zidentyfikowanych terminatorów, jak i w ich stosunku do liczby genów w obrębie tych samych gatunków, co rodzi obawy o spójność wyników. Szczególnie wyraźna rozbieżność występuje w przypadku *B. subtilis*, gdzie liczba zidentyfikowanych terminatorów waha się od 635 do 1214 w zależności od użytego zbioru danych (stosunek do liczby genów wynosi odpowiednio 0,15 i 0,29). Sugeruje to, że różnice te mogą być determinowane różnorodnością danych. **Proszę o komentarz Kandydata na temat wpływu tych czynników – co mogło spowodować tak znaczną rozbieżność wyników w obrębie jednego gatunku?**

Następnie Doktorant przeprowadził szczegółową analizę miejsc rho-niezależnych terminacji transkrypcji, porównując skład nukleotydowy tych obszarów oraz otaczających je sekwencji w genomach sześciu różnych gatunków bakterii. Wyniki jego badań wskazują, że zarówno pod względem struktury terminatora, jak i zawartości uracylu w sekwencji poniżej spinki terminacyjnej, kompozycja u gronkowca jest podobna do gatunków blisko spokrewnionych. **Z obliczeń wynika, że u gronkowca w obszarze poniżej spinki terminacyjnej znajduje się średnio 7–8 nukleotydów uracylowych. Natomiast przedstawiony w kolejnym rozdziale model spinki dla gronkowca, ilustrujący zachowawczość tej struktury, ma ich jedynie 4 (Wykres 14). Proszę o komentarz dotyczący tej rozbieżności.**

Ostatnim etapem badań Kandydata była analiza zestawu siedmiu publicznie dostępnych zbiorów danych RNA-seq, w których gronkowiec złocisty był poddawany działaniu subletalnych dawek różnych antybiotyków. Doktorant wykazał, że terminator transkrypcji o ID 2018 może brać udział w odpowiedzi tej bakterii na działanie linezolidu i nafciliny. Terminator ten zlokalizowany jest w operonie *pyr* odpowiedzialnym za biosyntezę pirymidyn (zwracam tu uwagę na niezgodność w nazewnictwie – na odpowiednich wykresach operon oznaczono jako *pyrBI*). Poza terminatorem 2018, doktorant zidentyfikował w regionie *pyr* jeszcze jeden terminator o ID 2015. Doktorant zauważył również, że terminator 2018 oraz sąsiedni terminator 2015 zajmują pozycje bardzo zbliżone do tych opisanych w literaturze jako terminatory regulujące ekspresję operonu *pyr* u *B. subtilis*, co potwierdza wiarygodność uzyskanych wyników. Wyniki sugerują, że stres antybiotykowy może wywoływać przedwczesną terminację transkrypcji operonu *pyr*, co w efekcie ogranicza ekspresję jego białek.

Podsumowując, wyniki przedstawione w rozprawie jednoznacznie dowodzą, że Pan mgr Jan Grzegorz Kosiński w pełni zrealizował wszystkie założone cele badawcze. Praca doktorska została przygotowana na wysokim poziomie merytorycznym, a zawarte w niej wyniki wnoszą oryginalny i istotny wkład do nauki. Przedstawione rozwiązanie naukowego problemu świadczy o solidnej wiedzy teoretycznej Kandydata w dziedzinie nauk biologicznych oraz o jego zdolności do samodzielnego prowadzenia badań. Uwagi sformułowane powyżej mają charakter dyskusyjny i nie wpływają na wartość naukową rozprawy.

Stwierdzam, że przedstawiona mi do oceny rozprawa doktorska spełnia warunki określone w obowiązującej Ustawie o stopniach i tytule naukowym i wnoszę do Rady Dyscypliny Nauki Biologiczne Uniwersytetu im. Adama Mickiewicza w Poznaniu o dopuszczenie Pana mgra Jana Grzegorza Kosińskiego do dalszych etapów przewodu doktorskiego.



PODPIS ZAUFANY

ANNA
PHILIPS

14.10.2024 12:43:19 (GMT+2)

Dokument podpisany elektronicznie
podpisem zaufanym