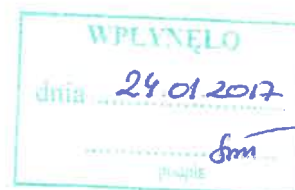


Dr. hab. Jacek Leluk, prof. nadzwyczajny
Uniwersytet Zielonogórski
Wydział Nauk Biologicznych
65-516 Zielona Góra, ul. Szafrana 1 (bud. A-8)
tel. (+48) 68 328 7870
fax. (+48) 68 328 7323
Tel. kom.: (+48) 501466847

Zielona Góra, 19.01.2017



Dziedkanat Wydziału Elektrotechniki,
Automatyki, Informatyki i Inżynierii
Biomedycznej,
Akademia Górniczo-Hutnicza im. St.
Staszica w Krakowie
Al. Mickiewicza 30
30-059 Kraków

Recenzja rozprawy doktorskiej mgr Krzysztofa Sarapaty

Mgr Krzysztof Sarapata wykonał pracę doktorską zatytułowaną "Model biocybernetyczny w analizie genomu człowieka" wykonaną pod kierunkiem Prof. dr hab. med. Marka Sanaka oraz Prof. dr hab. inż. Ryszarda Tadeusiewicza na Wydziale Elektrotechniki, Automatyki, Informatyki i Inżynierii Biomedycznej Akademii Górniczo-Hutniczej w Krakowie.

Wybrany temat nie jest tematem łatwym. Wymaga dużego zaangażowania oraz szerokiej wiedzy o charakterze interdyscyplinarnym. Autor przedstawił w pracy wyniki swoich badań teoretycznych, których głównym celem było rozszerzenie modelu TargetScore oraz jego weryfikacja w oparciu o dostępny materiał kliniczny,

Imponująca jest zastosowana w pracy metodologia badawcza. Wymagała szeroko opanowanej wiedzy w zakresie nauk medycznych, biologii molekularnej, genetyki oraz bioinformatyki.

Autor podchodził do swoich wyników z dużą dozą ostrożności i samokrytycyzmu, co jest bardzo ważną cechą dobrego badacza.

Praca jest obszerna (zawiera 131 stron) i napisana z dużą starannością. Jedyne drobne zastrzeżenia można mieć do charakteru stylistyki użytego języka w niektórych fragmentach tekstu. Jednakże nie umniejsza to w żadnym stopniu wysokiej wartości merytorycznej i poznawczej pracy doktorskiej mgra Krzysztofa Sarapaty.

Z tego powodu nie będę nawet cytował tych fragmentów, które mogłyby być napisane nieco dojrzałszym językiem naukowym. Przytoczę jedynie, że niepotrzebnie nadużywane są takie określenia jak *target* (używane w liczbie pojedynczej i mnogiej oraz odmienianej przez przypadki) oraz *predykcja*. Słowa te nie występują w słowniku polskim i powinny zostać zastąpione właściwymi określeniami, np. przewidywanie i sekwencje docelowe (specyficzne). Cytowanych w pracy 188 pozycji literaturowych świadczy o dobrym i rzetelnym opanowaniu wiedzy w wybranej tematyce. Osobiście

jestem pełen uznania dla Autora, który przedstawił pracę na bardzo wysokim poziomie. Rozprawa doktorska oprócz swoich walorów ujawnia także pewne niedoskonałości.

Przedstawiony model biocybernetyczny mechanizmu interferencji RNA (Rys. 4.2 Model regulacji genów z wyróżnionym mechanizmem interferencji RNA, s.61) na podstawie aktualnej wiedzy o mechanizmie funkcyjnych cząsteczek miRNAs ujawnia, jakie są aktualne możliwości modelowania tego procesu. Ustalenie rzeczywistych relacji ilościowych między poziomami transkryptów i cząsteczek miRNAs oraz ustalenie jakościowej relacji między tymi cząsteczkami dopiero pozwoli na potwierdzenie zaproponowanego modelu tego procesu biologicznego.

W podjętym przez autora rozprawy zadaniu opisanego relacji ilościowych między poziomami ekspresji transkryptów i cząsteczek miRNAs wykorzystane zaplecze eksperymentalne do ustalenia tych relacji wydaje się być zdecydowanie zbyt skromne, nawet, jeśli przyjmujemy argumenty autora biblioteki TargetScan, odnośnie poprawy jakości przewidywania na podstawie integracji danych kontekstowych i pochodzących z analiz filogenetycznych. Przeprowadzony eksperyment mikromacierzowy i uzyskane wyniki nie pozwalają na stwierdzenie istotnych statystycznie poszukiwanych relacji dla celów wnioskowania medycznego. Tym bardziej, że w eksperymencie "Astma" oznaczono poziom tylko miRNAs, a nie dokonano oznaczenia poziomów samych transkryptów. Stanowiło to niestety pewne wyzwanie dla autora rozprawy, aby z częściowych danych dokonać przewidywania specyficznych sekwencji. Prawdopodobnie przyczyną takiego stanu rzeczy były ograniczenia finansowe. Doprowadzenie zaproponowanej w pracy metody "biTargetScore" do takiego stanu, aby

Leluk

1000000000

1000000000

1000000000

1000000000

umożliwiła wnioskowanie medyczne byłoby bardzo przydatnym celem pod względem aplikacyjnym. W przyszłości przedstawiony model "biTargetScore" domaga się zatem poszerzenia i zwielokrotnienia przeprowadzonych badań eksperymentalnych w celu wykazania efektywności wykorzystanego modelu matematycznego zaproponowanego w pracy na tle innych narzędzi predykcji, co zresztą zostało także zaznaczone w podsumowaniu.

Pełna procedura analizy cząsteczek miRNAs w próbce biologicznej podlega ciągłym udoskonaleniom w celu uwzględnienia różnych mechanizmów biologicznych, które na przykład modyfikują cząsteczki RNA. Zastosowana w eksperymencie "Astma" mikromacierz CURY LNA wykazuje się dużą swoistością na poziomie 1 nukleotydu przy analizie komplementarności hybrydujących nici, więc pozwala uwzględnić efekty wewnątrzkomórkowych modyfikacji sekwencji miRNAs. Natomiast, nie jest jasne, czy w pracy przeanalizowano poziom fosforylacji końców miRNA. W pracy brakuje informacji o szczegółach przeprowadzonej procedury mikromacierzowej. Czy wspomniana w rozprawie na s.15 (rozdział 2.2 Mechanizm interferencji RNA) defosforylacja końca 5' miRNA została uwzględniona w eksperymencie?

Mechanizm interferencji RNA jest relatywnie nowym odkryciem mechaniki wewnątrzkomórkowej, który w ciągu kilku ostatnich lat podlega intensywnej eksploracji. Sugerowany przez autora rozprawy kierunek znaczenia tego mechanizmu w immunologii organizmu powinien zostać potwierdzony z czasem. Wydaje się to bardzo obiecujący kierunek, co podkreślają także prace innych autorów, natomiast w pracy stanowi on obecnie jak na razie hipotezę.

Można zauważyć nieścisłości w rozdziale 2. Biologiczne podstawy regulacji genów, szczególnie na rysunku 2.1. Centralny dogmat biologii molekularnej, gdzie wykazano możliwość syntezy nici DNA lub RNA na matrycy białkowej. Definicje pojęć "gen", "splicing" użyte zostały dość ogólnikowo, chociaż wystarczająco na potrzeby tejże rozprawy.

Nie sprecyzowano wyraźnie charakteru pracy – jest to praca teoretyczno-empiryczna.

Nie podano wyraźnie rodzaju metody weryfikacji postawionej hipotezy. Stosowana w pracy metoda weryfikacji jest metodą empiryczną, która wykorzystuje statystykę w analizie uzyskanych danych i wyników.

Niedobory danych eksperymentalnych, które są podstawową bolączką weryfikacji zaproponowanego modelu wyszukiwania sekwencji docelowych biTargetScore można byłoby zredukować poprzez wykorzystanie publicznie dostępnych danych mikromacierzowych. Dlaczego wzorem autora biblioteki TargetScore autor rozprawy nie wykorzystał np. archiwum ArrayExpress (<http://www.ebi.ac.uk/arrayexpress/>)?

Część obliczeniowa doktoratu obejmuje procedurę wyznaczania rozpoznawanych sekwencji. Oprócz tego pewna część rozprawy poświęcona jest rozważaniom o jakości bioinformatycznych danych (roz. 2.5, roz. 2.6), a właściwie to poszukiwaniu innego, logicznego wyjaśnienia mechanizmu interferencji RNA, który nie ogranicza się tylko do ilościowej regulacji transkryptów w cytoplazmie (roz. 2.2, roz. 2.4). Autor rozprawy próbuje wspierać hipotezę o kontroli jakościowej transkryptów w RNAi.

Zasadniczo autor rozprawy opiera się na dwóch faktach: zauważa niedoskonałości metod procedury pozyskiwania odpowiednich danych oraz wskazuje mocne przesłanki ewolucyjne np. odmienne funkcjonowanie tego mechanizmu w świecie roślin. Stąd pojawia się owa hipoteza kontroli jakościowej transkryptów w RNAi. Autor rozprawy próbuje rozwinąć teorię Sergey Petoukhov organizacji kodu genetycznego do wsparcia tej hipotezy (roz. 7.2). Można tutaj dodać, że proces ewolucji charakteryzuje ciągłość czasowa, więc obecny stan biologii upoważnia do stawiania nie tylko hipotez dotyczących stanu obecnego, ale także przewidywania w kwestii "przygotowań" środowiska do nowej funkcyjności. Hipoteza autora rozprawy wymaga pewnie jeszcze wielu badań i analiz. Na przykład dokładnej lokalizacji w obrębie istniejących mechanizmów biologicznych procedury dekompozycji informacji genetycznej na macierzy Hadamarda. Skoro mechanizm organizacji redundancji kodu genetycznego występuje u wszystkich organizmów, dlaczego dopiero u zwierząt występuje podział funkcyjny cząsteczek miRNA na seed i pozostały fragment o najbardziej tajemniczej funkcji? W jaki sposób patologiczny transkrypt miałaby indukować wzmożoną interakcję z cząsteczkami miRNAs? Zaproponowany w pracy model biTargetScore realizuje przewidywanie funkcyjnych sekwencji zgodnie z kategoryzacją narzędzi zaproponowaną na s. 42. Wykorzystuje w tym celu informację o zmienności ekspresji miRNAs oraz równocześnie mRNAs w porównywanych próbach referencyjnej i badanej. Gdzie badana próba może mieć cechy patologiczne lub być hodowana w jedynie odmiennych warunkach. To podejście pozyskiwania danych niewątpliwie wykazuje przewagę nad metodami transfekcji, które bezpośrednio wprowadzają nowe cząsteczki miRNAs do hodowli, powodując zaburzenie spontanicznej zależności endogennych miRNAs jakie tam występują. Trzeba jednak podkreślić, że nadal nierozwiązany pozostał problem wynikający z izolacji próby od całościowego, fizjologicznego systemu w organizmie dawcy. Niewątpliwie otwartym pozostaje pytanie, jakie właściwości mechanizmu RNAi zostają wówczas utracone i pominięte, albo, jakie pojawiają się nowe

Celula

relacje – artefakty w tak przeprowadzonej procedurze. Brak kontekstu całego organizmu w analitycznej hodowli wyabstrahowanej próbki niewątpliwie mocno wpływa w rozumienie mechanizmów rozwoju organizmu, co zresztą zauważył autor rozprawy (s.25). Czy to uproszczenie utrudnia poznanie faktycznego mechanizmu RNAi?

Podstawowe niewiadome procesu RNAi pozostają: brak stabilności wiązania dupleksów miRNA/mRNA oraz niska komplementarność miejsca wiązania i sekwencji miRNA. Te niejasności wymagają rewizji biochemicznego modelu parowania się zasad nukleotydowych co jest już bardzo dobrze poznane. Zamiast alfabetu czteroliterowego można posługiwać się alfabetem odmierności strukturalnych czterech podstawowych zasad azotowych. Chodzi o podział nukleotydów na jedno i dwupierścieniowe pirymidyny/puryny, podatność na mutację kwasem azotowym, liczba wiązań wodorowych. Właśnie to podejście zostało uwzględnione przez wspomnianego powyżej Sergeya Petoukhova.

Niezależnie od moich krytycznych drobnych uwag przedstawiona praca posiada bardzo wysoką wartość poznawczą i aplikacyjną. Pod względem merytorycznym jest bardzo dojrzałym opracowaniem naukowym.

Podsumowując oceniam pracę doktorską mgra Krzysztofa Sarapaty jako **bardzo dobrą** i wnioskuję o dopuszczenie do dalszych etapów przewodu doktorskiego. Jednocześnie sugeruję przedstawić tę pracę do konkursu na najlepszą pracę doktorską organizowanego przez Polskie Towarzystwo Bioinformatyczne.

Z poważaniem



dr hab. Jacek Leluk

prof. nadzwyczajny Uniwersytetu Zielonogórskiego

10

C

C