



dr hab. Michał Szcześniak, prof. UAM
Uniwersytet im. A. Mickiewicza w Poznaniu
Wydział Biologii
miszcz@amu.edu.pl

Poznań, 7 listopada 2022

Recenzja rozprawy doktorskiej mgr Katarzyny Nowis pt.
„Podejścia bioinformatyczne w badaniach kolistych RNA u roślin”

Rozprawa doktorska przedstawiona do recenzji została wykonana pod kierunkiem prof. dr hab. Marka Figlerowicza (promotor) oraz dr hab. Anny Philips (pełniącej rolę promotora pomocniczego) w Zakładzie Biologii Molekularnej i Systemowej Instytutu Chemii Bioorganicznej PAN w Poznaniu. Można w niej wyróżnić rozdziały: *Lista publikacji*, *Streszczenie*, *Abstrakt*, *Krótkie omówienie pracy doktorskiej*, *Publikacje wchodzące w skład rozprawy doktorskiej* oraz *Oświadczenia dotyczące prac wchodzących w skład rozprawy doktorskiej*. Rdzeń merytoryczny rozprawy oraz podstawę do ubiegania się o stopień naukowy doktora stanowią trzy recenzowane publikacje naukowe, których tematem przewodnim jest identyfikacja małych kolistych RNA (circRNA) u roślin, a także badanie ich właściwości, w tym procesu biogenezy. Poniżej przedstawiono zwięzłą charakterystykę tych trzech prac badawczych, a następnie ewaluację całego cyklu publikacji w kontekście postępowania o nadanie stopnia naukowego doktora.

1. Krótkie omówienie poszczególnych prac w cyklu publikacji

Publikacja nr 1

Tytuł: Expression Landscape of circRNAs in *Arabidopsis thaliana* Seedlings and Adult Tissues

Krótką charakterystyka wyników:

Opracowano już różne metody eksperymentalne i sprzężone z nimi podejścia bioinformatyczne służące do identyfikacji circRNA, jednak często prowadzą one do uzyskiwania rozbieżnych wyników przez różnych badaczy przy wykorzystaniu tego samego materiału lub danych wejściowych. Oznacza to, że istnieje potrzeba standaryzacji, a także wypracowania podejścia, które pozwoli zoptymalizować uzyskiwane wyniki pod kątem czułości i specyficzności. Mając to na względzie, autorzy pracy przetestowali sześć podejść *in silico* oraz dwa eksperymentalne. Od strony laboratoryjnej, wypróbowane dwa sposoby przygotowywania biblioteki do sekwencjonowania transkryptomów (wariant z wykorzystaniem RNazy R oraz bez niej). Jeśli chodzi o analizy *in silico*, przetestowano różne metody normalizacji liczebności odczytów

pochodzących z sekwencjonowania wysokoprzepustowego. Dokładna analiza danych oraz wspomagające ją eksperymenty pozwoliły wytypować optymalną kombinację wyżej wspomnianych podejść, która to została użyta do opracowania finalnego zestawu circRNA u *Arabidopsis thaliana*. Zestaw ten stał się przedmiotem dalszych badań. Między innymi, celem lepszego zrozumienia biogenezy circRNA i procesów zawiadujących tym procesem, badano korelację ekspresji między circRNA a liniowymi transkryptami powstającymi z tych samych genów. Zbadano też powtarzalność występowania circRNA pomiędzy powtórzeniami biologicznymi oraz pomiędzy badanymi organami *A. thaliana* i nasionami. W ten sposób wykazano, że zdecydowana większość circRNA powstaje specyficznie w pojedynczych próbkach, co wskazuje na silnie stochastyczny charakter ich biogenezy. Z drugiej strony, wyszczególniono zestaw circRNA, które występują w szeregu próbek, co może wskazywać na ich znaczenie funkcjonalne.

Uwagi merytoryczne:

- W analizie danych przyjęto, że odczyty z sekwencjonowania RNA-Seq po obróbce mogą mieć minimalną długość 20 nt – czy to nie za mało?
- 99.00% odczytów przeszło etapy obróbki (tj. usuwanie adaptorów, filtrowanie względem jakości). Jest to dość wysoka wartość, a zatem można by zadać pytanie: jakość danych była niezwykle wysoka czy też należałoby przyjąć nieco ostrzejsze kryteria obróbki danych?
- Jakie są różnice algorytmiczne pomiędzy *CIRI2* i *find_circ*, które uzasadniają użycie obu narzędzi w tandemie?
- W przypadku analizy korelacji (*Pearson correlation coefficient*) nie podano wartości *P-value*.
- Tylko 1.3% do 5.0% circRNA występuje we wszystkich czterech powtórzeniach biologicznych - czy u zwierząt obserwuje się podobną tendencję?

Publikacja 2

Tytuł: At-C-RNA database, a one-stop source for information on circRNAs in *Arabidopsis thaliana* in a unified format

Krótką charakterystyką wyników:

Celem przyświecającym autorom było opracowanie pierwszego repozytorium, które gromadziłoby wyniki identyfikacji circRNA przeprowadzonej według jednolitego potoku bioinformatycznego, co jest niewątpliwie wartością dodaną. Owocem starań jest At-C-RNA, największa internetowa baza danych circRNA u *A. thaliana*, przechowująca 113 327 circRNA zidentyfikowanych z wykorzystaniem publicznie dostępnych zasobów RNA-Seq.

Uwagi merytoryczne:

- Baza danych nie jest dostępna online w momencie pisania recenzji
- Skoro jednym z głównych celów było przygotowanie ujednoliconego zestawu circRNA (uzyskanego jednym protokołem bioinformatycznym), to jak poradzono sobie z faktem,

że dane wejściowe przygotowane były według różnych procedur, co prawdopodobnie wymusza zastosowanie pewnych modyfikacji w obróbce *in silico* (np. odczyty *stranded/unstranded*, odczyty sparowane/pojedyncze czy też różna zawartość odczytów mapujących się do rRNA)?

Inne uwagi:

- Szkoda, że w publikacji nie przedstawiono metod bioinformatycznych (choćby w materiałach suplementarnych), powołując się jedynie na inną, własną publikację.
- Pani Katarzyna Nowis jest pierwszym autorem pracy – i całkowicie zasłużenie, jako że przeprowadziła analizy bioinformatyczne, zbudowała od podstaw bazę danych (będącą głównym osiągnięciem publikacji) oraz uczestniczyła w pisaniu manuskryptu.

Publikacja 3

Tytuł: *Arabidopsis thaliana cbp80, c2h2, and flk* Knockout Mutants Accumulate Increased Amounts of Circular RNAs

Krótką charakterystykę wyników:

W pracy przedstawiono wyniki identyfikacji circRNA w 18 roślinach *A. thaliana* zawierających zmutowany jeden z 18 genów powiązanych ze splicingiem pre-mRNA oraz w roślinach typu dzikiego. Wykonano porównanie profilów circRNA, wykazując znaczny wzrost liczebności circRNA w mutantach *cbp80*, *flk* oraz *c2h2*. Postawiono hipotezę, że geny CBP80, C2H2 i FLK mogą być zaangażowane ustalanie prawidłowej kolejności egzonów w procesie splicingu – pod ich nieobecność (w badanych mutantach) splicing zmienia się w kierunku generowania większej ilości kolistych cząsteczek RNA. Jest to o tyle ciekawe, że poprzez regulację ekspresji czy dostępności tych białek w procesie splicingu można by modulować poziom circRNA, a zatem i sprzężonych z nimi transkryptów liniowych, co być może stanowi interesujący mechanizm regulacji ekspresji genów.

Uwagi merytoryczne:

- Podobnie jak we wcześniejszej pracy, także tutaj minimalna długość odczytu to 20 nt – czy słusznie?
- Autorzy piszą: „In total, 24 different circRNAs were identified in the leaves of the wild-type plant.” – nie jest od razu jasne, że w tym paragrafie (i w kolejnych) mowa o tych circRNA, które wykryto we wszystkich powtórzeniach biologicznych
- Potwierdzenie eksperymentalne obecności a także poziomów ekspresji wybranych circRNA (na przykład techniką qPCR) znacznie podniosłoby wartość naukową pracy

2. Podsumowanie osiągnięć mgr Katarzyny Nowis w przedłożonym cyklu publikacji

Koliste RNA (circRNA) zostały już stosunkowo dobrze opisane u człowieka i powiązano je z szeregiem chorób i procesów komórkowych, choć trzeba pamiętać, że jedynie ułamek z nich doczekał się dogłębnej charakterystyki funkcjonalnej. Możliwe, że większość z nich nie pełni istotnych funkcji molekularnych, stanowiąc jedynie tło dla innych procesów zachodzących na poziomie transkryptomu. Niemniej jednak w świecie roślin nasza wiedza jest jeszcze bardziej ograniczona: poza bardzo nielicznymi przykładami, nie znamy ani też nie domyślamy się, jakie funkcje mogłyby pełnić w komórce roślinnej. Wciąż tajemnicę stanowią szczegóły dot. procesu ich biogenezy, w tym odpowiednie mechanizmy regulatorowe. Co więcej, trudno o wiarygodny zestaw tych cząsteczek, nawet w takim organizmie modelowym, jak *A. thaliana*. Wychodząc naprzeciw wspomnianym ograniczeniom, pani Katarzyna Nowis i współautorzy prac poczynili wartościowe postępy w tej dziedzinie, które pokrótce scharakteryzowano wyżej.

3. Wnioski

Zgodnie z obowiązującymi przepisami, wliczając *art. 187 ust. 3 ustawy z dnia 20 lipca 2018 r. Prawo o szkolnictwie wyższym i nauce*, recenzja rozprawy doktorskiej powinna zawierać jednoznaczną odpowiedź na poniższe trzy pytania:

1. Ocena wraz z uzasadnieniem, czy rozprawa doktorska prezentuje ogólną wiedzę teoretyczną osoby ubiegającej się o nadanie stopnia doktora w określonej dyscyplinie albo dyscyplinach

Przedstawiona rozprawa doktorska w dyscyplinie „nauki biologiczne” w sposób jednoznaczny pokazuje wystarczającą wiedzę pani Katarzyny Nowis w zakresie bioinformatycznej obróbki danych oraz powiązanych aspektów biologii molekularnej, co wnioskuje na podstawie trzech publikacji, jak również wprowadzenia teoretycznego, stanowiącego wstęp do rozprawy.

2. Ocena wraz z uzasadnieniem, czy rozprawa doktorska wykazuje umiejętność samodzielnego prowadzenia pracy naukowej lub artystycznej przez osobę ubiegającą się o nadanie stopnia doktora

Autorka rozprawy doktorskiej miała znaczący wkład merytoryczny w powstanie wszystkich trzech publikacji w cyklu; szczególnie wyraźne jest zaangażowanie w bioinformatyczną obróbkę danych, co – uwzględniając walory merytoryczne tej części – pozwala sądzić, że pani Nowis jest gotowa prowadzić samodzielne badania na tym polu. Pewien niedosyt jednak pozostawia raczej techniczny charakter udziału doktorantki w dwóch pracach (na podstawie oświadczeń współautorów dołączonych do rozprawy doktorskiej wiadomo, że oprócz analiz bioinformatycznych, doktorantka uczestniczyła w planowaniu eksperymentów, ale w samych publikacjach brak o tym wzmianki).



3. Ocena wraz z uzasadnieniem, czy rozprawa doktorska stanowi oryginalne rozwiązanie problemu naukowego, oryginalne rozwiązanie w zakresie zastosowania wyników własnych badań naukowych w sferze gospodarczej lub społecznej albo oryginalne dokonanie artystyczne.

Uzyskane wyniki i walory merytoryczne publikacji, skrótowo scharakteryzowane wyżej, pozwalają jednoznacznie stwierdzić, że rozprawa doktorska przedstawia oryginalne rozwiązanie problemu badawczego. W szczególności należy wskazać zaproponowanie wytycznych dot. identyfikacji circRNA (w tym normalizację danych NGS), opracowanie internetowej bazy danych czy też analizy bioinformatyczne prowadzące do inspirujących wniosków w kontekście biologii circRNA.

4. Podsumowanie

Przedstawiona do recenzji rozprawa doktorska, w tym wyniki uzyskane przez doktorantkę, mogą być przedmiotem zainteresowania społeczności naukowej i stanowią wartościowy wkład do badań nad circRNA u roślin, mimo pewnych drobnych niedociągnięć czy też nieścisłości merytorycznych. Stwierdzam, że rozprawa doktorska spełnia warunki określone w Ustawie z dnia 20 lipca 2018 roku prawo o szkolnictwie wyższym i nauce (Dz.U. z 2018 r. poz. 1668 ze zm.), Ustawie z dnia 3 lipca 2018 r. Przepisy wprowadzające ustawę – Prawo o szkolnictwie wyższym i nauce (Dz.U. z 2018 r. poz. 1669 ze zm.) oraz w Sposobie postępowania w sprawie nadania stopnia doktora w Instytucie Chemii Bioorganicznej PAN w Poznaniu (uchwała Rady Naukowej ICHB PAN nr 99/2022/Internet z dnia 9 czerwca 2022 r.) i wnioskuje do Rady Naukowej Instytutu Chemii Bioorganicznej PAN o dopuszczenie pani Katarzyny Nowis do dalszych etapów postępowania o nadanie stopnia doktora.

Michał Szczeniok