



INSTYTUT GENETYKI ROŚLIN POLSKIEJ AKADEMII NAUK

ul. Strzeszyńska 34, 60-479 Poznań

Tel.: centrala 616550200, sekretariat 616550255, e-mail: office@igr.poznan.pl www.igr.poznan.pl

Dr hab. Grzegorz Koczyk
Instytut Genetyki Roślin
Polskiej Akademii Nauk
Zakład Biometrii i Bioinformatyki
ul. Strzeszyńska 34, 60-479, Poznań

Poznań, 30.09.2023

RECENZJA

w postępowaniu doktorskim
mgr inż. Małgorzaty Marszałek-Zeńczak
opracowana na podstawie informacji
Rady Naukowej Instytutu Chemii Bioorganicznej PAN

1. Podstawy formalne recenzji

Podstawą formalną recenzji jest decyzja Rady Naukowej Instytutu Chemii Bioorganicznej PAN z dnia 1 sierpnia 2023 r. powołująca recenzentów pracy pt. "Udział polimorfizmu liczby kopii w kształtowaniu wewnątrzgatunkowej zmienności strukturalnej metabolicznych klastrów genów u *Arabidopsis thaliana*".

Przedstawiona praca doktorska została zrealizowana w Instytucie Chemii Bioorganicznej PAN pod kierunkiem dr hab. Agnieszki Żmieńko, prof. IChB PAN przy wykorzystaniu środków finansowych z projektu NCN/OPUS/2014/13/B/NZ2/03837 pt. "Poszukiwanie mechanizmów indukujących polimorfizm liczby kopii dużych fragmentów genomu oraz związków tego zjawiska z tworzeniem i ewolucją metabolicznych klastrów genów". Wyniki realizacji pracy doktorskiej włączono do trzech wieloautorskich publikacji recenzowanych w czasopiśmie *Frontiers in Plant Sciences* oraz *Plant Cell*. Wszystkie prace są pracami oryginalnymi, w jednej z nich Kandydatka jest autorem wiodącym. Dorobek Doktorantki obejmuje ponadto współautorstwo dwu publikacji niewchodzących w skład pracy.

Rozprawie towarzyszy zestaw oświadczeń współautorów prac wchodzących w skład cyklu. Z oświadczeń tych można wnioskować, że mgr inż. Małgorzata Marszałek-Zeńczak odegrała istotną rolę w przeprowadzeniu opisanych badań jak również brała udział w przygotowaniu samych publikacji na poszczególnych etapach edycji. W świetle przedstawionych danych, dobór tych prac nie budzi moich wątpliwości. Wobec braku istotnych niedociągnięć procesu recenzenckiego opublikowanych prac recenzowanych, nie widzę zatem konieczności omawiania ich strony edytorskiej w dalszej części mojej oceny.



INSTYTUT GENETYKI ROŚLIN POLSKIEJ AKADEMII NAUK

ul. Strzeszyńska 34, 60-479 Poznań

Tel.: centrala 616550200, sekretariat 616550255, e-mail: office@igr.poznan.pl www.igr.poznan.pl

2. Ocena

Złożona praca została opatrzona obszernym autoreferatem stanowiącym wstęp do analizowanych problemów (zarówno identyfikacji wariantów strukturalnych, w tym wariantów kopijności; jak i zróżnicowania pod tym względem klastrów metabolicznych). Autorka w liczącym 35 stron autoreferacie (nie licząc sekcji technicznych), w sposób przekrojowy podsumowuje uzyskane wyniki badań, jak również ich podstawy praktyczne i teoretyczne oraz znaczenie. Pewien niedosyt budzi sekcja podsumowująca, w której chętnie przeczytałbym więcej na temat perspektyw dalszych badań oraz nowych narzędzi w niezwykle dynamicznie rozwijającej się dziedzinie jaką są studia nad zróżnicowaniem strukturalnym i sekwencyjnym (pan)genomów roślinnych.

W cyklu zgłoszonych prac, pierwsze dwie stanowią fundament warsztatu odpowiednio eksperymentalnego i bioinformatycznego wykorzystanego przez Doktorantkę najpełniej w ostatniej publikacji cyklu.

W pierwszej z prezentowanych prac “MLPA-Based Analysis of Copy Number Variation in Plant Populations” (Frontiers in Plant Science, 2017), rola Doktorantki obejmowała optymalizację ilości matrycy do rozwijanego zastosowania technik MLPA w badaniach *Arabidopsis thaliana*. Udział Autorki mimo zwięzłego opisu, oceniam jako kluczowy dla efektywnego i powtarzalnego zastosowania techniki w badaniach kopijności u *Arabidopsis* (patrz poniżej). Należy podkreślić, że użycie MLPA w analizach kopijności u rzodkiewnika miało w dużym stopniu charakter pionierski. Co niezwykle istotne, zweryfikowany i zoptymalizowany protokół stanowi podstawę eksperymentalnej weryfikacji wyników bioinformatycznych w trzeciej pracy osiągnięcia (analizy MLPA prowadzone dla 13 genów w 232 liniach, uzupełnione dodatkowo o analizy *droplet digital PCR*).

W drugiej pracy pt. “AthCNV: A Map of DNA Copy Number Variations in the Arabidopsis Genome” (Plant Cell, 2020), ilościowy wkład Doktorantki w stworzenie bazy i powstanie publikacji był zdecydowanie większy. Utworzony przez autorów katalog dużych indeli (o długości 50-499 pz) oraz regionów o zmiennej liczbie kopii w ponad 1000 naturalnych linii *Arabidopsis* jest niezwykle przydatnym zasobem w porównawczych analizach zmienności strukturalnej genomu krzyżowych. W pracy, Kandydatka wykonała zarówno większość analiz wariantów strukturalnych z wykorzystaniem 4 z 7 opisywanych narzędzi, jak i opracowała podsumowania całości wyników z wykorzystaniem autorskiego środowiska integracji danych. Wykonała również analizy GWAS i PCA, zawarte zarówno w tej jak i ostatniej publikacji. W kontekście całości cyklu, jej wkład stanowi fundament niezbędnego warsztatu bioinformatycznego oraz (jako udostępniona społeczności naukowej baza danych) istotny przyczynek do zrozumienia zakresu zróżnicowania strukturalnego genomu różnych linii rzodkiewnika, również w kontekście ewolucyjnym.

Trzecia z przedstawionych prac “Analysis of Arabidopsis non-reference accessions reveals high diversity of metabolic gene clusters and discovers new candidate cluster members” (Frontiers in Plant



INSTYTUT GENETYKI ROŚLIN POLSKIEJ AKADEMII NAUK

ul. Strzeszyńska 34, 60-479 Poznań

Tel.: centrala 616550200, sekretariat 616550255, e-mail: office@igr.poznan.pl www.igr.poznan.pl

Science, 2023) zasługuje na szczególne wyróżnienie. Synteza wykorzystywanych uprzednio technik bioinformatycznych i eksperymentalnych umożliwiła Doktorantce szczegółowe określenie zróżnicowania strukturalnego klastrów metabolicznych *Arabidopsis thaliana*. Warto podkreślić, iż dla szczegółowo przebadanych genów, zaobserwowano wysoką zgodność wyników analiz bioinformatycznych i eksperymentalnych (98,8%). Szczegółowe wnioski dotyczą czterech klastrów biosyntezy terpenoidów (tirucalladienolu, marneralu, thalianolu i arabidiolu/baruolu). Spośród wymienionych, klastry biosyntezy tirucalladienolu i marneralu wykazały jedynie minimalne zróżnicowanie. W przeciwieństwie do powyższych, skupiska genów związanych z biosyntezą odpowiednio thalianolu i arabidiolu/baruolu charakteryzowało wysokie zróżnicowanie strukturalne.

W tym ostatnim klastrze, zidentyfikowano długą insercję zawierającą dywergentne duplikaty genów CPY705A2 oraz BARS1 - identyfikując w ten sposób nową, prawdopodobnie funkcjonalną syntazę oksydoskwalenową BARS2. Wyniki uzyskane przez Doktorantkę wskazują na intensywność zmian ewolucyjnych klastra datowaną na okres po rozdzieleniu linii gatunkowych *Arabidopsis thaliana* i *A. lyrata* i odzwierciedloną w zróżnicowaniu obecnych populacji rzodkiewnika.

Wyniki badań potwierdziły ponadto zmienioną ekspresję klastra w liniach rzodkiewnika zawierających insercję. Co znamienne, wspomniane linie wydaje się charakteryzować odmienny fenotyp (dynamika wzrostu systemu korzeniowego) oraz preferencje środowiskowe (ciepłe strefy klimatyczne). Mimo braku silniejszego potwierdzenia statystycznego, badania te ze wszech miar zasługują na dalsze rozwinięcie. Na uwagę, moim zdaniem, zasługują również sprawne wykorzystanie dyskretnego kodowania zmian typu CNV w analizie asocjacyjnej oraz identyfikacja prawdopodobnej subfunkcjonalizacji pary BARS1/BARS2 na drodze zmian profilu ekspresji.

Wobec wieloautorskiej natury prac, należy podkreślić iż autorski wkład Doktorantki jest jasno określony. Tym samym wpisuje się w naturę osiągnięcia jako syntezy współtworzonych warsztatów eksperymentalnego i bioinformatycznego, w celu lepszego poznania ewolucji klastrów metabolicznych u roślin. Zwłaszcza ostatnia praca, obszerna i syntetyczna, łącząca techniki eksperymentalne i bioinformatyczne, w udanej analizie skomplikowanego zagadnienia ewolucyjnego, zasługuje moim zdaniem na szczególne wyróżnienie.

Oprócz omówionych powyżej prac wchodzących w skład rozprawy doktorskiej, Kandydatka jest również współautorką dwu kolejnych prac (International Journal of Molecular Sciences, 2021; Functional & Integrative Genomics, 2023) dotyczących odpowiednio polimorfizmów genu EGFR w kontekście chemioterapii drobnokomórkowych guzów tkanki płucnej oraz zróżnicowania kopijności genów mikroRNA u *Arabidopsis thaliana*. Rosnąca aktywność publikacyjna w ostatnich latach dodatkowo świadczy o dojrzałości naukowej i interdyscyplinarnym potencjale Autorki.



INSTYTUT GENETYKI ROŚLIN POLSKIEJ AKADEMII NAUK

ul. Strzeszyńska 34, 60-479 Poznań

Tel.: centrala 616550200, sekretariat 616550255, e-mail: office@igr.poznan.pl www.igr.poznan.pl

3. Pytania i uwagi

Z recenzenckiego obowiązku, pragnę zwrócić uwagę Autorki na kilka punktów, które wzbudziły w trakcie lektury moje zastrzeżenia. Jako czytelnikowi, sekcja 1.2 wprowadzenia (Identyfikacja wariantów strukturalnych w genomie człowieka i ich związek z fenotypem) wydaje się stosunkowo luźno związana z pozostałymi i relatywnie zbędna. Pozostałe sekcje tekstu wydają się wyczerpujące, napisane zwięzłym i klarownym językiem i opatrzone licznymi odniesieniami literaturowymi.

W sekcji 1.6 wprowadzenia Autorka definiuje metabolity wtórne jako związki, które “nie są niezbędne dla wzrostu i rozwoju samej rośliny” i “odgrywają dużą rolę w komunikacji między rośliną a środowiskiem”. Ze względu na nasze rosnące zrozumienie roli tych metabolitów w środowisku naturalnym, sugerowałbym dokładniejszą kwalifikację tego rodzaju definicji. Związki te nie są wprawdzie niezbędne w kontekście przeżycia w warunkach ściśle kontrolowanych, jednak w warunkach naturalnych (tj. konkurencji o ograniczone zasoby środowiskowe) mogą odgrywać jak najbardziej kluczową rolę. Kwalifikacja powyższa jest tym istotniejsza ze względu na udokumentowany wpływ związków takich jak marneral na istotne cechy rozwojowe komórki (m.in. cytowana przez autorkę praca Guo i in. 2012).

Jak wspomniałem wcześniej, pewien niedosyt budzi podsumowanie autoreferatu (rozdział 4), byłbym więc Kandydatce wdzięczny za podzielenie się swoją oceną perspektyw i wyzwań stojących przed analizami pangenomu roślinnego podczas obrony rozprawy.

Podczas lektury pracy nasunęły mi się następujące dodatkowe pytania:

- 1) W ostatnich kilku latach zidentyfikowano istotnie większą liczbę przypuszczalnych klastrów biosyntezy w genomach roślinnych. W pracy opisującej narzędzie plantiSMASH służące identyfikacji klastrów na podstawie sekwencji (Kautsar i in. 2017), w genomie Arabidopsis zidentyfikowano między 31 a 114 klastrów kandydackich (w zależności od przyjętych poziomów odcięcia). Czy wykonano jakieś analizy wstępne (np. bioinformatyczną ocenę z wykorzystaniem AthCNV, co sugeruje treść trzeciej pracy) na pozostałych klastrach kandydackich rzodkiewnika? Jeśli tak, jakie były ich wyniki?
- 2) W odniesieniu do wniosku pracy dotyczącego zmienności genów kodujących syntazy terpenowe lub cytochromy P450 (w kontraście do loci niesparowanych). Czy wykonywano również analizy presji selekcyjnej (np. z wykorzystaniem stosunkowo odpornej na błędy nadzacowania metodyki, opisanej w pracy Davydow i in. 2019; <https://doi.org/10.1093/molbev/msz048>)? Analizy selekcji mogłyby w tym kontekście dostarczyć cennych informacji o funkcji adaptacyjnej poszczególnych zmian SNP.

Powyższe uwagi i pytania, w najmniejszym stopniu nie obniżają oceny przedstawionej pracy, stanowiąc wyraz ciekawości naukowej oraz przemyśleń recenzenta.



INSTYTUT GENETYKI ROŚLIN POLSKIEJ AKADEMII NAUK

ul. Strzeszyńska 34, 60-479 Poznań

Tel.: centrala 616550200, sekretariat 616550255, e-mail: office@igr.poznan.pl www.igr.poznan.pl

6. Wnioski końcowe

Przedstawiona do recenzji rozprawa doktorska spełnia warunki 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. 2003, nr 65 poz. 595, ze zm.) oraz w Rozporządzeniu Ministra Nauki i Szkolnictwa Wyższego z dnia 19 stycznia 2018 r. w sprawie szczegółowego trybu i warunków przeprowadzania czynności w przewodzie doktorskim, w postępowaniu habilitacyjnym oraz w postępowaniu o nadanie tytułu profesora (Dz.U. 2018 poz. 261) i wnioskuję do Rady Naukowej Instytutu Chemii Bioorganicznej PAN o dopuszczenie Małgorzaty Marszałek-Zeńczak do dalszych etapów przewodu doktorskiego. Jednocześnie, wnoszę o rozpatrzenie możliwości wyróżnienia pracy.

dr hab. Grzegorz Koczyk