



Poznań, 04.10.2023

Prof. dr hab. Piotr Ziółkowski  
Instytut Biologii Molekularnej i Biotechnologii  
Uniwersytet im. Adama Mickiewicza w Poznaniu

**Recenzja rozprawy doktorskiej mgr inż. Małgorzaty Marszałek-Zeńczak  
pt. „Udział polimorfizmu liczby kopii w kształtowaniu wewnątrzgatunkowej zmienności  
strukturalnej metabolicznych klastrów genów u *Arabidopsis thaliana*”**

Przedstawiona mi do oceny praca doktorska Pani Małgorzaty Marszałek-Zeńczak dotyczy szeroko badanego w zespole Pani dr hab. Agnieszki Żmieńko zagadnienia zmienności liczby kopii u *Arabidopsis thaliana*. Zagadnienie to do niedawna było poznane tylko w bardzo niewielkim stopniu, ponieważ w przeciwieństwie do polimorfizmu pojedynczych nukleotydów (SNP), polimorfizm liczby kopii jest znacznie trudniejszy w identyfikacji w oparciu o dominującą obecnie technologię sekwencjonowania wysokoprzepustowego opartą na krótkich odczytach. W przypadku pracy Pani Marszałek-Zeńczak akcent został postawiony na analizę zmienności strukturalnej metabolicznych klastrów genów, co jest szczególnie interesujące z naukowego punktu widzenia, dostarcza bowiem nowej wiedzy na temat genezy, ewolucji funkcjonalnej i zmienności populacyjnej bardzo specyficznej klasy genów roślinnych.

Rozprawa ma formę zbioru trzech artykułów oryginalnych, w których Pani Marszałek-Zeńczak odgrywa rolę pierwszego (w jednej pracy) bądź drugiego (w dwóch publikacjach) autora. Wszystkie prace ukazały się w renomowanych czasopismach naukowych z listy filadelfijskiej: dwa z nich we *Frontiers in Plant Science* a jeden w *Plant Cell*. Ponieważ publikacje są wieloautorskie, Kandydatka do stopnia doktora opatrzyła całą rozprawę autoreferatem, w którym pokrótce opisane i przedyskutowane są uzyskane przez nią wyniki. Poza omówieniem wyników badań, autoreferat obejmuje także Wprowadzenie, klarownie zdefiniowany Cel Pracy, a także Podsumowanie i perspektywy. Uważam, że autoreferat jest przygotowany znakomicie: napisany jest bardzo przejrzystym i ładnym językiem, z dużą dbałością o styl wypowiedzi, w sposób logiczny przedstawiając poszczególne zagadnienia i łącząc je w jedną spójną całość. Co warto podkreślić, Doktorantka w autoreferacie nie stroni od licznych odniesień do nowych trendów w badaniach genomów eukariotycznych, szeroko dyskutując najnowsze publikacje a nawet preprinty dotyczące sekwencjonowania przy użyciu technologii opartych na długich odczytach, takich jak PacBio czy Oxford Nanopore Sequencing. Pozwala jej to na spojrzenie na swoje badania w sposób krytyczny, dostrzegając ich ograniczenia i jasno definiując dalsze perspektywy postępu nauki w tym zakresie. Świadczy to o dużej dojrzałości naukowej Doktorantki i nabycia przez nią umiejętności wnikliwej analizy literatury naukowej.

1

Chociaż w autoreferacie trudno znaleźć błędy rzeczowe czy logiczne, z racji mojej roli recenzenta rozprawy pozwolę sobie zwrócić uwagę na kilkukrotnie pojawiające się stwierdzenia o charakterze deterministycznym: Na przykład:

Strona 30: „W rezultacie grupowanie genów może być cechą skrajnej selekcji, dążącej do współdziedziczenia i optymalizacji najlepszych kombinacji genów dla nowych szlaków metabolicznych, które zapewniają selektywne korzyści (...).”

Strona 39: „Różnice te mogą być skutkiem pozytywnej selekcji w kierunku utrzymania ciągłego klastra genów.”

Selekcja naturalna, podobnie jak ewolucja, nie są procesami ukierunkowanymi, ani też nie mają z góry określonego celu. Dlatego w przyszłości Kandydatka powinna unikać stosowania deterministycznych sformułowań wypowiadając się w podobnych kontekstach.

W dalszej części swojej pracy Doktorantka zamieściła kopie artykułów, które wchodziły w skład jej pracy doktorskiej, wraz z dołączonymi do nich suplementami. Pierwszy artykuł, opublikowany we *Frontiers in Plant Science*, jest pracą metodyczną, której celem była optymalizacja metody multipleksowej amplifikacji zależnej od ligacji (MLPA) na potrzeby eksperymentalnej weryfikacji zmienności liczby kopii genów. Przedstawiony w tej pracy protokół był następnie wykorzystywany w kolejnych publikacjach Zespołu. Rola Pani Małgorzaty Marszałek-Zeńczak w jego powstaniu była stosunkowo skromna, polegała bowiem głównie na określeniu optymalnych zakresów stężeń matrycy DNA używanych do analiz.

Drugi artykuł, w którym Doktorantka występuje jako drugi autor, została opublikowana w jednym z najważniejszych czasopism z zakresu biologii molekularnej roślin, tj. w *Plant Cell*. Praca obejmuje wielopoziomową i bardzo rozbudowaną analizę zmienności liczby kopii genów u *A. thaliana*, przeprowadzoną w oparciu o ponad tysiąc zsekwencjonowanych linii tego gatunku. Analiza opierała się o złożony „pipeline”, który w oparciu o kilka różnych narzędzi bioinformatycznych identyfikował regiony CNV. Łącznie odkryto ponad 70 tys. dużych indeli i ponad 34 tys. regionów CNV. Przeprowadzone przez zespół badania pozwoliły na dokonanie szeregu odkryć dotyczących struktury genomu Arabidopsis pozwalając na wyciągnięcie ważnych wniosków odnośnie jego funkcjonowania i ewolucji w skali populacyjnej. Rola Doktorantki w tej pracy, zgodnie z informacjami zawartymi w autoreferacie jak i oświadczeniami autorów, była bardzo duża: Kandydatka brała czynny udział w planowaniu badań i w wykonaniu znacznej części analiz bioinformatycznych, zwłaszcza w optymalizacji parametrów i kryteriów filtracji dla niektórych zastosowanych narzędzi, a także, co wydaje się szczególnie istotne, opracowaniu autorskiego podejścia do integracji wyników uzyskanych z różnych programów. Pozwoliło to na stworzenie algorytmu do identyfikacji CNV, którego wysoką użyteczność potwierdzono eksperymentalnie. Pani Marszałek-Zeńczak przeprowadziła także analizę statystyczną wszystkich wyników ujętych w publikacji, włączając w to analizy PCA, oraz zaangażowała się w przeprowadzanie analizy GWAS – wkrótce stała

się ekspertem w tym zakresie a nabyte umiejętności z powodzeniem wykorzystwała także w kolejnej pracy.

Ostatnia praca, wchodząca w skład rozprawy doktorskiej, została opublikowana w bieżącym roku w czasopiśmie *Frontiers in Plant Science* i w niej Pani Marszałek-Zeńczak jest pierwszym autorem. Biorąc pod uwagę tytuł dysertacji oraz udział Doktorantki w powstaniu tej publikacji należy uznać ją za najważniejszą w jej dorobku. Praca jest swoistą kontynuacją analiz przeprowadzonych w publikacji z *Plant Cell*, ale z wyraźnym ukierunkowaniem na zmienność metabolicznych klastrów genowych. Pomimo lokalizacji wszystkich czterech klastrów w obrębie najbardziej zmiennych regionów chromosomowych, poszczególne klastry wykazują duże zróżnicowanie w zakresie polimorfizmu strukturalnego. Doktorantka wiąże te różnice w zmienności ze znaczeniem biologicznym poszczególnych szlaków dla funkcjonowania rośliny. Tak więc najważniejszy funkcjonalnie klaster marneralu charakteryzuje się niemal całkowitą konserwacją strukturalną, podczas gdy klaster thalianolu występuje w dwóch wersjach strukturalnych, związanych z zajęciem niewielkiej inwersji w części linii. Najważniejszy moim zdaniem wynik został uzyskany dla klastru arabidiolu/baruolu, gdzie Doktorantka odkryła nową, niewystępującą w genomie referencyjnym i nieznaną wcześniej parę genów, występującą jednakże w znacznej liczbie linii. W związku z tym odkryciem Kandydatka przeprowadziła szeroką charakterystykę efektów wyodrębnienia się tej nowej podjednostki strukturalnej w kształtowaniu zmienności fenotypowej i adaptacji do warunków środowiskowych. Są to istotne wyniki, które rzucają nowe światło na ewolucję genomów roślinnych w zakresie powstawania i ewolucji genów.

W trakcie studiowania rozprawy doktorskiej Pani Marszałek-Zeńczak nasunęło mi się kilka pytań, które chciałbym przedyskutować w trakcie obrony rozprawy:

1. W schemacie przedstawionym na Rys. 1.1 pokazano, że zmienność genetyczna wpływa na zmienność epigenetyczną. Czy jest możliwa relacja odwrotna i czy można zaproponować mechanizmy, które mogłyby ją wywoływać?
2. We Wprowadzeniu Doktorantka forsuje hipotezę, że organizacja genów szlaku metabolicznego w klastry sprzyja koregulacji ich ekspresji. Jednak fakt, że dla zdecydowanej większości genów oddzielna lokalizacja nie jest przeszkodą w synchronizacji transkrypcji sugeruje, że nie jest to element kluczowy dla istnienia takich klastrów. Zatem czy organizacja klastrowa ma jakiegokolwiek zalety w genomach eukariotycznych? A może jest raczej efektem szczególnego sposobu powstawania tego typu szlaków syntezy metabolitów wtórnych?
3. Kontynuując w tym temacie: opisując zmienność strukturalną klastra thalianolu Doktorantka proponuje, że wersja ciągła klastra znajduje się w trakcie utrwalania w populacjach *A. thaliana* w związku z ich domniemaną przewagą ewolucyjną. Czy jednak fakt, że wszystkie opisane u *Arabidopsis* klastry są specyficzne jedynie dla tego gatunku,

ewentualnie rodzaju, brak natomiast klastrow o silnie konserwowanym układzie genów, nie sugeruje jednak, że w dłuższej perspektywie czasowej ulegną one rozpadowi, podobnie jak większość innych genów? A może znane są przypadki klastrow metabolicznych o zachowanej strukturze?

4. Poza klastrami metabolicznymi inną grupą genów, która często występuje w postaci rodzin tandemowych u roślin są geny odporności na patogeny. Jakie są podobieństwa i różnice dotyczące ewolucji i funkcjonowania tych dwóch typów klastrow genowych?
5. Lynch and Conery w 2000 roku wykazali bardzo wysoką częstość zdarzeń duplikacji genów przy jednoczesnym relatywnie szybkim zanikaniu ich zduplikowanych kopii w ciągu kilku milionów lat po powstaniu (Lynch and Conery, Science, 2000). Zaledwie niewielki ułamek nowo powstałych genów ma szansę na utrwalenie w genomie w toku subfunkcjonalizacji lub neofunkcjonalizacji. W jakim stopniu Pani wyniki dotyczące zmienności strukturalnej klastrow metabolicznych potwierdzają te obserwacje?
6. Czy może Pani zaproponować mechanizmy genetyczne, które prowadzą do powstawania i ewolucji nowych regionów CNV?

Chociaż pierwsze dwie publikacje są tematycznie znacznie szersze niż ostatnia praca, koncentrująca się na zmienności klastrow metabolicznych, nie mam wątpliwości, że wszystkie prace stanowią spójny i wartościowy naukowo cykl. Podsumowując moją recenzję chciałbym podkreślić bogaty i zaawansowany warsztat bioinformatyczny zastosowany przez Doktorantkę w jej badaniach oraz głęboką znajomość najnowszej literatury w zakresie analizy genomów eukariotycznych. Osobiście jestem bardzo ciekawy dyskusji z Doktorantką na temat ewolucji strukturalnej rodzin genowych, którą odbędziemy w trakcie obrony pracy!

Nie mam wątpliwości, że w swojej rozprawie Pani mgr inż. Marszałek-Zeńczak podjęła istotny problem badawczy oraz przedstawiła w sposób satysfakcjonujący jego rozwiązanie. Dlatego z pełnym przekonaniem stwierdzam, że przedstawiona mi do recenzji praca Pani Małgorzaty Marszałek-Zeńczak pt. „Udział polimorfizmu liczby kopii w kształtowaniu wewnątrzgatunkowej zmienności strukturalnej metabolicznych klastrow genów u *Arabidopsis thaliana*” 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 Pani Małgorzaty Marszałek-Zeńczak do dalszych etapów przewodu doktorskiego.

