

Instytut Chemii Bioorganicznej PAN  
Zakład Genomiki Roślin

Poznań, 05.02.2022 r.

**Recenzja rozprawy doktorskiej pana mgr. Juana Camilo Ochoa Cabezas  
pt. „Elucidation of genetic factors determining resistance or susceptibility to clubroot disease  
through genome wide association study, transcriptome profiling and functional genetics in  
*Arabidopsis natural accessions*”**

Przedstawiona mi do recenzji rozprawa doktorska pana mgr. Juana Camilo Ochoa Cabezas została wykonana w Zakładzie Zintegrowanej Biologii Roślin Instytutu Genetyki Roślin PAN, pod kierunkiem Profesora tego Instytutu, dr hab. Roberta Malinowskiego. Promotorem pomocniczym jest dr William Truman. Praca przedstawia wyniki badań poświęconych identyfikacji genetycznych czynników związanych z wykazywaniem zwiększonej odporności na infekcję pierwotniakiem *Plasmodiophora brassicae* przez niektóre ekotypy rzodkiewnika pospolitego (*Arabidopsis thaliana*).

*P. brassicae* jest obligatoryjnym biotrofem, który infekuje głównie rośliny z rodziny kapustowatych (*Brassicaceae*), powodując u nich chorobę zwaną kiłą kapusty. Charakterystycznymi symptomami tej choroby są: powstawanie narośli na podziemnej części roślin, zahamowanie wzrostu, więdnienie oraz chloroza. Wynikają one z szeregu zmian metabolicznych i fizjologicznych, wywołanych obecnością patogena. Wiele podatnych na zakażenie roślin kapustowatych ma istotne znaczenie ekonomiczne. Ogromna trwałość zarodników przetrwalnikowych *P. brassicae*, która w odpowiednich warunkach może wynosić wiele lat, stwarza realne zagrożenie dla uprawy tych roślin. Rodzi to potrzebę opracowywania skutecznych rozwiązań ograniczających rozprzestrzenianie się kiły kapusty, do których należy zaliczyć tworzenie odmian odpornych na tę chorobę. To z kolei wymaga pogłębienia wiedzy na temat naturalnych mechanizmów chroniących rośliny przed atakiem patogenu bądź ograniczających negatywne skutki infekcji.

Należący do kapustowatych rzodkiewnik jest podatny na infekcję *P. brassicae*. Jest to jednocześnie roślina modelowa w badaniach molekularnych i genetycznych. Rzodkiewnik jest od szeregu lat wykorzystywany jako model w badaniach odpowiedzi roślin na infekcję *P. brassicae* w Zakładzie Zintegrowanej Biologii Roślin IGR PAN. Praca mgr Juana Camilo Ochoa Cabezas bardzo dobrze wpisuje się w nurt tych ważnych poznawczo, a potencjalnie także aplikacyjnie, badań.

Przedstawiona rozprawa doktorska jest napisana w języku angielskim i ma klasyczny układ dla prac w swojej dziedzinie. Liczy 158 stron, na co składają się, oprócz zasadniczych rozdziałów, które omawiam poniżej, również streszczenia w języku polskim i angielskim, wykaz stosowanych w pracy skrótów, spisy tabel (których jest 10) i rysunków (których jest 50), oświadczenia autora, promotora i promotora pomocniczego oraz informacje o finansowaniu doktoranta i jego badań. W tym miejscu chciałabym zauważyć, że na część z opisanych w pracy badań Doktorant zdobył finansowanie z Narodowego Centrum Nauki, w ramach trwającego nadal grantu PRELUDIUM. Z porównania opisu projektu umieszczonego na stronie NCN z zawartością rozprawy doktorskiej można wnioskować, że główne zadania badawcze tego grantu zrealizowano. Wskazuje to na umiejętność Doktoranta w planowaniu i prowadzeniu projektu badawczego, a tym samym na jego dużą samodzielność naukową.

Umieszczone na początku pracy streszczenia zawierają klarowny opis treści, co mogłam potwierdzić po przeczytaniu całej rozprawy. Zwróciłam jednak uwagę, że tytuły obu wersji językowych nie są tożsame. Zastosowane w wersji angielskiej sformułowanie „Elucidation of genetic factors” jest dużo bardziej ogólne od stwierdzenia „wyjaśnienia udziału poszczególnych genów”, którym posłużono się w wersji polskiej. Ponadto po przeczytaniu obu wersji otrzymujemy pełniejsze informacje, niż ograniczając się do jednej z nich. Przykładowo, tylko w wersji angielskiej podano, że analizy transkryptomyczne przeprowadzono w 19-tym dniu po inokulacji. Brak w niej natomiast pełnej nazwy pierwotniaka *P. brassicae* czy też rozwinięcia skrótu nazwy białek typu TIR-NB-LRR.

Rozdział 1 (Wstęp) wprowadza czytelnika w problematykę przedmiotu. Przedstawiono w nim cykl życiowy *P. brassicae*, obejmujący etapy pierwotnej i wtórnej infekcji komórek korzenia rośliny gospodarza oraz tworzenie i uwalnianie zarodników przetrwalnikowych. Obrazuje to Rysunek 1, którego jednak nie uznałam za szczególnie pomocny. Badania Doktoranta skupiają się na ocenie zmian zachodzących w roślinie, dlatego zabrakło mi tu zestawienia faz rozwoju patogenu z fazami proliferacji i ekspansji komórek korzenia, o których dużo mówi się w dalszej części pracy. We Wstępie zaprezentowano także systemy klasyfikacji patotypów patogenu, oparte o ocenę symptomów choroby w wybranych gatunkach i odmianach roślin. Dalej, opisano stan wiedzy na temat genetycznych podstaw odporności roślin kapustowatych na kiłę kapusty, w kontekście pokrewieństwa ich genomów. Genomy *Brassica oleracea* i *Brassica rapa* mogą służyć jako źródła introgresji genów odporności w ulepszeniach rzepaku (*Brassica napus*), który jest amfidiploidem, powstałym na drodze hybrydyzacji pomiędzy tymi gatunkami. Jednak, jak wskazuje Doktorant, pomimo intensywnych badań, tylko w kilku przypadkach udało się zidentyfikować takie geny w naturalnie odpornych odmianach. Kolejną część Wstępu stanowi wyczerpujący opis stanu badań nad udziałem fitohormonów i zmian w cyklu komórkowym w postępie infekcji pierwotniakiem *P. brassicae* u rzodkiewnika. Ujęto tu również wyniki wcześniejszych badań zespołu Profesora Malinowskiego. Dalej następuje krótkie omówienie historii badań asocjacji całego genomu (GWAS) u rzodkiewnika. Część teoretyczną kończy przedstawienie dwóch podstawowych mechanizmów odporności roślin na biotyczne czynniki stresowe: odporności nieswoistej (tzw. PTI, czyli *pattern-triggered immunity*) oraz swoistej (ETI, czyli *effector-triggered immunity*), wyzwalanej w wyniku obecności specyficznych efektorów – białek produkowanych przez określone szczepy patogenne. W mojej ocenie, ta ostatnia sekcja powinna poprzedzać omówienie badań nad identyfikacją genów odporności na kiłę kapusty. Niemniej uważam, że Autorowi udało się zawrzeć we Wstępie wszystkie zasadnicze elementy potrzebne do zrozumienia prezentowanych dalej wyników oraz ich dyskusji.

Przedstawiony w Rozdziale 2 cel pracy jest jasno i precyzyjnie sformułowany, a jego osiągnięcie oparto na założeniach logicznie wynikających z dotychczasowego stanu wiedzy. Zastosowane i wypisane w Rozdziale 3 metody badawcze są różnorodne, od prowadzenia kultur rośliny i patogenu, poprzez obserwacje makro- i mikroskopowe, do analiz genetycznych, molekularnych i bioinformatycznych. Metody są przedstawione w sposób ogólny, stosowany w opisach do artykułów badawczych, co wydaje mi się wystarczające. Niemniej w niektórych miejscach brak jest precyzji w ich opisie. Przykładowo, w sekcji 3.3 napisano, że reakcja trawienia RNazą (zwracam uwagę na niepoprawny zapis tej nazwy w pracy) została zatrzymana przy pomocy „chloroform:isoamyl 24:1”, połączoną z wytrącaniem mieszaniną etanolu i octanu sodu. Jak rozumiem, chodzi o ekstrakcję mieszaniną chloroform: **alkohol izoamylowy** 24:1, a wytrącano fazę wodną. Z kolei w opisie warunków reakcji qPCR w tejże sekcji chodziło chyba o końcowe stężenie starterów 0.25  $\mu$ M każdy, a nie o 0.25  $\mu$ moła użyte w reakcji? Wreszcie nie rozumiem, jak i po co miano by obliczać logarytm

względnej ekspresji genów na podstawie qPCR („log<sub>2</sub> relative expression was calculated”), skoro badano względny poziom genomowego DNA rośliny i patogena? Analiza tego typu zwykle stosowana jest w badaniu ekspresji genów, jednak nazwę zapożyczono tu w sposób bezkrytyczny.

Szczególnie dużo miejsca w rozdziale 3 poświęcono wytworzeniu i selekcji linii mutantów rzodkiewnika metodą mutagenyzy CRISPR/Cas9. Dla odmiany, procedura wywołania ekspresji przejściowej genu *RPBI* w liściach tytoniu została potraktowana tak zdawkowo, że czytelnik nie może się dowiedzieć właściwie niczego na temat przygotowania odpowiedniego konstruktów bądź sposobu jego pozyskania. Również analizy GWAS potraktowano bardzo skrótowo, jednak zastosowane w nich modele i transformacje zostały opisane w adekwatnej części rozdziału Wyniki. Nie wydaje mi się to szczególnie uzasadnione, bo nieco zaciemnia opis tych wyników, ale uznaję taki wybór Autora.

Tabela 4 zawiera listę wykorzystanych ekotypów rzodkiewnika. Nie wiem dlaczego uporządkowano ją według identyfikatorów z banku nasion NASC (*nota bene*, zawartych w pierwszej, a nie drugiej kolumnie, wbrew towarzyszącemu opisowi). Posortowanie według nazw ekotypów, którymi Doktorant posługuje się dalej w pracy, znacząco ułatwiłoby korzystanie z tego zestawienia. Ponadto połowa ekotypów w tabeli ma przypisane informacje określone jako GWA, a druga połowa jako 1001 genomes ID. Wydaje mi się, że w obu przypadkach chodzi o identyfikatory ekotypów wg nomenklatury projektu 1001 Genomów, niemniej poproszę o wyjaśnienie tej kwestii.

W kolejnych dwóch rozdziałach, Autor opisuje kolejno przeprowadzone eksperymenty i ich wyniki (Rozdział 4 - Wyniki) oraz omawia ich znaczenie na tle istniejącego stanu wiedzy (Rozdział 5 - Dyskusja). Ta część rozprawy pozwala mi ocenić, że Doktorant bardzo dobrze poradził sobie z osiągnięciem postawionego celu. Badania rozpoczął on od obserwacji zmian morfologicznych i dynamiki namnażania patotypu PIB *P. brassicae* w wysoce wrażliwym ekotypie Col-0, w celu wyznaczenia optymalnego czasu do zbierania danych fenotypowych – w tym wypadku wybrano 19 dzień od momentu inokulacji. Mam pytanie do Doktoranta, czy zgodnie z Jego wiedzą były jakieś szczególne wskazania, dla których użyto akurat systemu Somé do oznaczenia patotypu? Czy na terenie Polski ten system dobrze różnicuje rasy *P. brassicae*? Następnie Doktorant wykonał pomiary względnego poziomu DNA patogenu oraz określił wskaźniki nasilenia symptomów choroby w 142 liniach rzodkiewnika. Oba typy danych wykazywały dużą zbieżność, co pozwoliło wyselekcjonować do dalszych badań ekotypy najbardziej odporne na infekcję *P. brassicae* (Est-1 oraz Uod-1), jak również jeden ekotyp o nietypowo wysokim poziomie DNA patogenu, przy stosunkowo słabych objawach choroby (Pro-0). Zebrane dane fenotypowe wykorzystano niezależnie od siebie w badaniach GWAS, wykonanych przy pomocy dwóch dostępnych narzędzi internetowych, dostarczających dane genotypowe dla rzodkiewnika. W tym wypadku były to informacje o występowaniu polimorfizmów pojedynczego nukleotydu (SNP). W efekcie Doktorant wytypował dwa regiony w genomie, dla których wykazane asocjacje były statystycznie istotne, bądź poniżej progu istotności, ale wyróżniające. W tym miejscu chciałabym zapytać, czy Doktorant rozważał bądź podejmował próby łącznego zastosowania obu fenotypów w badaniu typu GWAS? Czy jego zdaniem mogłoby to zwiększyć istotność statystyczną bądź liczbę uzyskanych asocjacji?

Jeden z wytypowanych regionów pokrywał się z locus genomowym, zidentyfikowanym we wcześniejszych badaniach, jednak dotąd nie zweryfikowanym przy pomocy analiz funkcjonalnych. Poprzez analizy sekwencji genomowych, w tym szczegółowe scharakteryzowanie tego regionu i ekspresji zakodowanych w nim ramek odczytu w odpornych ekotypach Est-1 i Uod-1, Autor zawęził

obszar zainteresowania do dwóch genów, *RPBI* oraz *RPBI-like-4*. W dalszej części pracy Doktorant wykonał serię eleganckich eksperymentów, obejmujących wyprowadzenie linii pojedynczych i podwójnych mutantów z delecjami w genach *RPBI* i *RPBI-like-4* w obu ekotypach, analizę ich ekspresji oraz makro- i mikroskopową ocenę symptomów choroby w zainfekowanych roślinach. Wykazały one jednoznacznie, że funkcjonalny gen *RPBI*, w przeciwieństwie do *RPBI-like-4*, jest niezbędny dla odporności ekotypów Est-1 i Uod-1 na zakażenie *P. brassicae*. *RPBI* koduje krótkie białko. Nie jest on zaangażowany bezpośrednio w rozpoznanie patogenu i chociaż nie udało się określić jego funkcji, to Doktorant wykazał, że jest on niezbędny dla indukcji genów związanych z odpowiedziami obronnymi, w tym wywołaniem reakcji nadwrażliwości. Uważam jednak, że analiza porównawcza sekwencji badanego regionu mogłaby być głębsza i objąć więcej ekotypów, np. można by wykorzystać tu krótkie odczyty sekwencyjne z Projektu 1001 Genomów w celu dokładniejszego określenia stopnia jego zmienności. Doktorant zaobserwował, że wszystkie naturalne ekotypy z prawdopodobną delecją genu *RPBI* są wrażliwe na infekcję *P. brassicae*. Proszę o skomentowanie faktu, że szereg ekotypów wrażliwych jednak ten gen posiada, zwłaszcza, że na poziomie sekwencji aminokwasowej wykazuje on wysoki stopień identyczności w obrębie gatunku. Autor wskazuje ponadto, że *RPBI* jest zduplikowany w niektórych ekotypach, na przykład w odpornym Tsu-0. Na podstawie Rysunków 13 i 16 nie jestem w stanie wywnioskować, na ile prowadzone analizy objęły również obecność i rolę tej duplikacji. Czy w świetle zebranych wyników polimorfizm liczby kopii genu *RPBI* może mieć znaczenie dla fenotypu odporności? Wreszcie proszę o uściślenie, czy matrycą do sekwencjonowania ekotypów Est-1 i Uod-1 był pojedynczy fragment obejmujący cały region o długości 11 kbp, czy też złożono wyniki sekwencjonowania krótszych amplikonów, prezentowanych na Rysunku 16?

W dyskusji dotyczącej tej części pracy, Autor podnosi wątek potencjalnego udziału czynników epigenetycznych w regulacji ekspresji genu *RPBI*. W genomie referencyjnym w omawianym regionie znajdują się liczne retrotranspozony. Jak wiadomo, epigenetyczna kontrola aktywności elementów mobilnych w genomie, może mieć wpływ na ekspresję sąsiadujących genów. Czy Autor rozważał zatem zbadanie tego regionu u ekotypów odpornych również pod kątem zawartości transpozonów bądź wzorców metylacji?

Dalej, Doktorant przeanalizował potencjalny związek drugiego locus wytypowanego w analizach GWAS, którym był gen *RAC1*, kodujący białko z rodziny TIR-NB-LRR. Stosując jak poprzednio analizę linii mutantów niosących нефункционаlną wersję genu *RAC1*, uzyskanych metodą mutagenyzy CRISPR/Cas9, Doktorant wykazał, że gen ten nie bierze udziału w wytworzeniu odporności na *P. brassicae* u ekotypów Uod-1 i Est-1. Mam uwagę do Rysunków 18, 19 i 34, przedstawiających przewidywane sekwencje aminokwasowe odpowiednio genów *RPBI*, *RPBI-like-4* i *RAC1* w poszczególnych liniach mutantów. Niezależnie od długości genu w genomie wyjściowym, prezentowane sekwencje powinny raczej kończyć się na pierwszym kodonie STOP w ramce odczytu. Z kolei w przypadku linii Uod-1 *rpbl* L69, prezentowana sekwencja nie zawiera kodonu STOP, tym samym teoretycznie może wykraczać poza obszar genu *RPBI*, co również powinno być wykazane.

Ostatni wątek badań poświęcono ciekawemu przypadkowi, jakim był ekotyp Pro-0, wykazujący słabe symptomy choroby, pomimo wysokiego poziomu DNA patogenu. Poprzez analizy histologiczne oraz badania transkryptomyczne tego ekotypu oraz – porównawczo – ekotypu Col-0, Doktorant wykazał, że obniżona wrażliwość Pro-0 na infekcję *P. brassicae* może mieć związek z wyższą proporcją tkanki przewodzącej, w szczególności ksylemu, w hipokotylach. Może to mieć związek z późniejszym

przechodzeniem z fazy proliferacji do fazy powiększania komórek i ograniczeniem zasięgu infekcji. Wskazują na to między innymi wyniki analiz ontologii genów oraz różnicowej ekspresji genów zaangażowanych w podziały komórkowe i procesy związane z różnicowaniem ksylemu, pomiędzy ekotypami. Jednym z takich genów był *XVP*, kontrolujący proliferację i różnicowanie komórek kambium. Autor wywołał nadekspresję genu *XVP* w ekotypie Pro-0, co wpłynęło na nasilenie stopnia replikacji DNA patogenu, ale nie na skalę symptomów choroby. Autor konkluduje, że deregulacja procesu tworzenia tkanki naczyniowej jest ważnym elementem w cyklu życiowym *P. brassicae* i zależy od naturalnych różnic pomiędzy ekotypami. Doktorant przywołuje najnowsze wyniki badań innych zespołów, które na to wskazują, jednocześnie zauważając konieczność prowadzenia dalszych badań w celu zrozumienia tych bardzo ciekawych zależności. Chciałabym w tym miejscu zapytać, czy badano jaka jest proporcja ksylemu w hipokotylach innych ekotypów o znanej odporności lub wrażliwości na *P. brassicae*?

Ostatnia sekcja Rozdziału 5, zatytułowana „Podsumowanie dyskusji”, jest w istocie kolejnym streszczeniem wyników pracy i wydaje mi się zbędna. Nie spotkałam się zresztą z podsumowaniem obejmującym wyłącznie dyskusję pracy. Bardziej poprawne wydaje się zamknięcie rozprawy wnioskami końcowymi. Na szczęście, taki rozdział również w niej się znalazł, a przedstawione w nim konkluzje uważam za uzasadnione w świetle przedstawionych wyników.

Od strony redakcyjnej, praca przygotowana jest dość starannie. Zamieszczone rysunki i fotografie są bardzo dobrej jakości. Jedynie w przypadku Rysunków 10, 12 i 15, które są jedynym źródłem informacji o rozkładzie alleli SNP w poszczególnych ekotypach, podpisy z nazwami ekotypów są zdecydowanie zbyt małe i nieczytelne. W tekście znalazłam pewną ilość błędów językowych i redakcyjnych. Większość z nich uznałam za mało znaczące, chciałam jednak podkreślić dwa mankamenty, które zdecydowanie obniżyły mój komfort jako czytelnika. Pierwszy to niesłychanie długie, wielokrotnie złożone zdania, które często zawierały w sobie błędy składni. Przykład takiego zdania można znaleźć np. na pograniczu stron 127-128 rozprawy – nie będę go tutaj przytaczać. Moje drugie zastrzeżenie dotyczy poważnych braków w danych bibliograficznych w Rozdziale 9 (Literatura). Dla znaczącej liczby artykułów nie podano numeru czasopisma, stron czy identyfikatora wersji elektronicznej. Ponadto w tytułach artykułów z dużą nonszalancją potraktowano pisownię nazw gatunkowych oraz nazw genów, pomijając kwestie takie jak pisownia wielką literą czy kursywa.

Powyższe uwagi nie zmieniają jednak faktu, że pod względem merytorycznym przedstawiona rozprawa jest dziełem spójnym i kompletnym. Posiada jasno postawiony cel, który jest zrealizowany w sposób logiczny i konsekwentny. Na uznanie zasługuje rozpiętość technik badawczych. Eksperymenty zawierają wymagane powtórzenia oraz są dobrze poparte statystycznymi analizami wyników. Weryfikacja wstępnych obserwacji jest zawsze wieloaspektowa, co sprawia, że końcowe wnioski są dobrze poparte dowodami. Autor ma dobre rozeznanie w problematyce tematu i umiejętnie ocenia wagę swoich badań na tle aktualnej wiedzy oraz wskazuje potencjalne kierunki do dalszych badań.

**Reasumując, w efekcie swoich badań Pan mgr Juan Camilo Ochoa Cabezas po raz pierwszy zidentyfikował w sposób jednoznaczny gen *RPBI* jako genetyczny czynnik odporności wybranych ekotypów rzodkiewnika na występując w Europie Centralnej patotyp *P. brassicae*. Wskazał również nowe ważne aspekty dotyczące związku przebiegu choroby z rozwojem ksylemu, które mogą mieć znaczenie dla łagodzenia skutków infekcji u roślin podatnych.**

**Stanowi to realizację celu pracy doktorskiej, a zarazem oryginalne i kompleksowe rozwiązanie problemu naukowego. Ponadto, jak Autor sam zauważa, jego obserwacje mogą przyczynić się do obrania nowych kierunków w poszukiwaniu źródeł podwyższonej tolerancji roślin uprawnych na zakażenie *P. brassicae*.**

Recenzowana praca spełnia ustawowe i zwyczajowe wymogi stawiane rozprawom doktorskim, stąd wnoszę do Rady Naukowej Instytutu Genetyki Roślin PAN o dopuszczenie Pana mgr Juana Camilo Ochoa Cabezas do dalszych etapów przewodu doktorskiego. Ponadto, biorąc pod uwagę wymienione powyżej wysokie walory merytoryczne pracy i istotność uzyskanych wyników, wnoszę o jej wyróżnienie.

.....*Agnieszka Żmienko*.....

Dr hab. Agnieszka Żmienko, prof. IChB PAN