

Prof. dr hab. Henryk Bujak
Uniwersytet Przyrodniczy we Wrocławiu
Katedra Genetyki, Hodowli Roślin i Nasiennictwa
Centralny Ośrodek Badania Odmian w Słupi Wielkiej

RECENZJA

rozprawy doktorskiej mgr Ewy Borzęckiej pt. „Wykorzystanie biblioteki BAC do integracji map fizycznych i genetycznych oraz identyfikacji genów związanych z udomowieniem i odpornością na choroby żyta”

Praca została wykonana w Katedrze Genetyki, Hodowli i Biotechnologii Roślin Szkoły Głównej Gospodarstwa Wiejskiego w Warszawie pod promotorstwem Pani dr hab. Hanny Bolibok-Bragoszewskiej, profesora uczelni.

Żyto, pomimo zmniejszającej się w ostatnich latach powierzchni uprawy jest, w dalszym ciągu, jednym z najważniejszych zbóż w Polsce. Wyróżnia się stosunkowo małymi wymaganiami glebowo-klimatycznymi, dlatego jego uprawa koncentruje się głównie na glebach lekkich, które w naszym kraju przeważają. Wykazuje dużą tolerancję na niskie temperatury oraz inne rodzaje stresów abiotycznych i biotycznych. Może służyć jako model odporności dla innych roślin zbożowych, w tym pszenicy i pszenżyta. Prace hodowlane, a zwłaszcza wprowadzenie do uprawy odmian mieszańcowych, przyczyniły się do ogromnego postępu w jego plonowaniu, a ziarno żyta ze względu na swoje walory odżywcze, dietetyczne i prozdrowotne jest cennym surowcem w przemyśle spożywczym. Jest cennym źródłem błonnika pokarmowego, minerałów takich jak cynk, żelazo i potas, a także szeregu związków bioaktywnych, w tym między innymi kwasu foliowego, steroli i folianów.

W pracach hodowlanych coraz częściej wykorzystywane są wyniki badań molekularnych, które ułatwiają selekcję ogromnej liczby materiałów hodowlanych pod względem określonych, pożądanych cech, a zwłaszcza odporności na stresy wywoływane przez czynniki biotyczne i abiotyczne, w tym odporności na choroby. W ostatnich latach nastąpił intensywny rozwój metod biologii molekularnej, w tym wysokoprzepustowych technik genotypowania i narzędzi bioinformatycznych, co znacznie przyspieszyło prace nad poznaniem genomu żyta.

Pierwsza wysokoprzepustowa technologia genotypowania żyta została wdrożona wraz z utworzeniem platformy do jego genotypowania zawierającego klony DArT. Zaowocowało to

pojawieniem się początkowo map genetycznych opartych na markerach DArT, a w późniejszym czasie na markerach opracowanych na podstawie sekwencji transkryptów. Postęp prac badawczych z zakresu genetyki i genomiki tego gatunku, dostępność nowych technik genotypowania i narzędzi bioinformatycznych zaowocowało opublikowaniem 2021 roku sekwencji nukleotydowej żyta. Żyto posiada jeden z największych genomów wśród roślin zbożowych (ponad 8 Gpz), który jest wysoce zmetylowany i z dużym udziałem sekwencji powtarzalnych (około 92%).

W przedstawionej do oceny rozprawie doktorskiej mgr Ewa Brzezička na podstawie posiadanych zasobów genomowych, takich jak biblioteki BAC oraz wykorzystując dostępne nowoczesne metody genotypowania dokonała integracji mapy fizycznej i genetycznej żyta oraz uzyskała pełną sekwencję wybranych genów związanych z odpornością na choroby i udomowieniem żyta. W pracy dokonała charakterystyki klonów BAC z wykorzystaniem markerów DArTseq, następnie przeprowadziła integrację mapy fizycznej i genetycznej wykorzystując zidentyfikowane adresy klon BAC-marker DArTseq, zidentyfikowała klony BAC zawierające geny związane z odpornością na choroby oraz uzyskała pełną sekwencję potencjalnych genów związanych z odpornością na choroby, a także zidentyfikowała ontologii kilku genów udomowienia innych gatunków występujące w genomie żyta.

Zakres badań jest bardzo szeroki, a wykorzystywanie kolejnych nowoczesnych metod badawczych i konsekwencja ich stosowania doprowadziła do osiągnięcia postawionego celu pracy oraz celów szczegółowych, a także umożliwiła weryfikację postawionych hipotez badawczych.

Przechodząc do oceny formalnej pracy doktorskiej można stwierdzić, że podział treści na rozdziały i podrozdziały jest powszechnie stosowany w tego typu opracowaniach i nie budzi zastrzeżeń. Rozprawa zawiera następujące rozdziały: streszczenia w języku polskim i angielskim, wykaz stosowanych skrótów, wstęp, cel i hipotezy badawcze, przegląd literatury z podrozdziałami, materiały i metody z wydzielonymi podrozdziałami, wyniki z podrozdziałami, dyskusję, wnioski, aneks oraz spis literatury. Brak jest jedynie spisu tabel, rysunków i załączonych tabel w aneksie, co jest drobnym uchybieniem technicznym i nie umniejsza to wartości merytorycznej pracy. Przedstawiony do oceny maszynopis rozprawy posiada 135 stron, w tym 19 tabel, 13 rysunków oraz 3 obszerne tabele zawarte w aneksie.

Wprowadzenie do badań przedstawione we wstępie i przeglądzie literatury jest szczegółowe, wielowątkowe, a równocześnie syntetycznie i jasno informuje czytelnika o zagadnieniach, które są celem dysertacji. Doktorantka omawia, na podstawie danych literaturowych, systematykę i filogenezę, znaczenie gospodarcze i zalety żyta, ewolucję jego

genomu, a następnie metody badawcze i techniki wykorzystywane w poznaniu genomu żyta. Upowszechnianie wysokoprzepustowych technologii genotypowania u żyta, w tym wykorzystanie jednego z wariantów GBS (genotypowanie przez sekwencjonowanie), jakim jest technologia DArTseq pozwoliło na konstrukcję pierwszej wysokorozdzielczej mapy genetycznej. Niestety ze względu na swoje ograniczenia opublikowana mapa genomu linii wsobnej żyta nie miała praktycznego zastosowania. Dopiero nowa wersja sekwencji genomu tej samej linii, która jest dostępna od 2021 roku umożliwiła oszacowanie liczby genów na 34 441 (geny o wysokim poziomie ufności), które najprawdopodobniej stanowią 97,9% wszystkich genów żyta.

Doktorantka informuje, że dotychczas scharakteryzowano niewiele genów warunkujących ważne użytkowo cechy żyta. Przy wykorzystaniu klasycznych metod analizy genetycznej i mapowania QTL ustalono, m. in., lokalizację niektórych genów warunkujących odporność na patogeny i stesy abiotyczne, w tym dwa geny związane z odpornością żyta na rdzę brunatną. W przytoczonych obszernych danych literaturowych można odnaleźć informacje o osiągnięciach polskich naukowców w poznawaniu genomu żyta, jak chociażby, uzyskanie sekwencji genów kontrolujących syntezę głównych wtórnych metabolitów, czyli biosyntezę benzoksazynoidów (BX) u żyta, czy określenie pozycji loci cech ilościowych (QTL) dla aktywności α -amylazy, wczesności kwitnienia i porastania ziarna na mapie genetycznej. Doktorantka charakteryzuje także badania nad określeniem QTLi zaangażowanych w kontrolę cech o dużym znaczeniu agronomicznym takich jak wysokość rośliny, długość kłosa i liczba ziaren w kłosie, które były rozmieszczone na wszystkich chromosomach żyta, a zidentyfikowano je dzięki wykorzystaniu map genetycznych o wysokiej gęstości.

Następnie Autorka rozprawy analizuje dostępne informacje na temat wykorzystania sztucznych chromosomów bakteryjnych (BAC), które dzięki swojej stabilności, braku chimeryzmu i możliwości rearanzacji insertów stały się cennym narzędziem genomiki strukturalnej. Są relatywnie łatwe w manipulacji i późniejszej aplikacji. Biblioteki BAC są przydatne w badaniach genomicznych, zwłaszcza u gatunków z niepoznaną jeszcze sekwencją genomu. Mogą służyć do izolacji genów, w wysokoprzepustowym fizycznym mapowaniu i sekwencjonowaniu genomów, rozwoju nowych markerów czy w cytogenetyce

Biblioteki BAC są cennym narzędziem genomiki strukturalnej umożliwiającym zidentyfikowanie rejonu genomu zawierającego wybraną sekwencję, poznanie sekwencji sąsiadujących z interesującym nas fragmentem i w efekcie zidentyfikowanie pełnej sekwencji genu. Ponadto kotwiczenie pojedynczych klonów BAC i kontigów na mapach genetycznych o

wysokiej rozdzielczości ma kluczowe znaczenie przy konstrukcji i eksploatacji map fizycznych, co zostało wykorzystane u wielu gatunków roślin.

W dalszej części tego rozdziału Autorka dysertacji przedstawia rozwój metod sekwencjonowania DNA i ich wpływ na sposób prowadzenia badań z zakresu genomiki ze względu na możliwość uzyskiwania obszernych danych dotyczących sekwencji genomu i transkryptomu w znacznie krótszym czasie i przy stosunkowo niskich kosztach, co przyczyniło się do rozwoju metod sekwencjonowania genomów. Sekwencjonowanie nowej generacji (NGS) zastosowano do genotypowania dużych genomów u różnych gatunków roślin uprawnych, takich jak kukurydza, ryż, jęczmień, pszenica i soja. Markery GBS są bardzo użyteczne w mapowaniu genetycznym i umożliwiają konstrukcję map genetycznych o wysokiej gęstości, a jedną z wersji GBS jest metoda DArTseq. Zaletami techniki DArTseq są wysoka przepustowość, atrakcyjna cena, uzyskiwanie dużej liczby markerów oraz krótki czas oczekiwania na wygenerowanie danych. Technologia ta jest optymalizowana dla każdego gatunku, w tym także dla żyta.

Postępy w technologii sekwencjonowania oraz rozwój metod analizy bioinformatycznej DNA dają coraz większe możliwości zrozumienia mechanizmów procesu udomowienia na poziomie genów i całego genomu organizmu. Doktorantka charakteryzuje zidentyfikowane dotychczas geny związane z udomowianiem różnych gatunków, co związane jest głównie z genami odpowiedzialnymi za zwiększenie wielkości części jadalnych roślin, jak np. liczba ziaren, formowanie osłonki ziarniaków, opadanie nasion, łamliwość osadki kłosa, osypywanie się nasion, półkarłowość, czy czas kwitnienia.

W przypadku genów kontrolujących odporność roślin na porażenie przez patogeny wskazanie przypuszczalnej funkcji homologów tych genów u spokrewnionych gatunków jest często utrudnione, ze względu na ich zdolność do szybkiego ewoluowania. W ostatnich dekadach sklonowano i scharakteryzowano ponad 100 genów odporności u różnych gatunków roślin, jednak, jak zauważa Doktorantka, dotychczas scharakteryzowano niewiele genów związanych z odpornością na choroby u żyta. Zidentyfikowano i zmapowano pięć dominujących genów odporności na rdzę brunatną u żyta, która jest jedną z najważniejszych chorób tego gatunku, wywoływaną przez grzyba *Puccinia recondita* f. sp. *secalis* (Prs): *Pr1*, *Pr2*, *Pr3*, *Pr4* i *Pr5*, stąd jednym z celów dysertacji jest poszerzenie wiedzy na ten temat.

Podsumowując stwierdzam, że Doktorantka bardzo dobrze przygotowała tę część dysertacji i zademonstrowała bardzo dobrą znajomość literatury z zakresu prowadzonych badań. Piśmiennictwo jest bogate i obejmuje 244 pozycje literatury bardzo precyzyjnie naświetlające problematykę badawczą podjętą w rozprawie, a o aktualności i ważności

podjętego w pracy problemu świadczy cytowanie najnowszej, opublikowanej w ostatnich latach literatury.

W rozdziale „Materiały i metody” Doktorantka charakteryzuje materiał badawczy sposób prowadzenia poszczególnych etapów badań, namnażania i analizy klonów BAC, szczegółowo opisuje metody izolacji DNA, zastosowane techniki genotypowania oraz narzędzia statystyczne i bioinformatyczne używane do obróbki uzyskanych danych.

W swoich badaniach wykorzystwała dostępną bibliotekę BAC linii wsobnej żyta L318, ze średnią długością insertu 121 kpz i z około 1,5-krotnym pokryciem genomu. Biblioteka ta składa się z 105 216 pojedynczych klonów, z której wydzieliła superpule (SP), a następnie wyizolowała DNA. Szczegółowo opisuje zastosowane procedury izolacji DNA, zarówno z pul zbiorczych, a jak i pojedynczych klonów BAC. Genotypowanie pul zbiorczych klonów BAC wykonała z wykorzystaniem jednej z metod Genotyping by Sequencing (GBS) stosując technikę DArTseq. W kolejnych podrozdziałach opisuje analizy bioinformatyczne markerów DArTseq oraz sposób kotwiczenie klonów BAC za pomocą tych DArTseq z wykorzystaniem zintegrowanej mapy genetycznej wygenerowanej z użyciem markerów DArTseq. Do sprawdzenia wiarygodności przyporządkowania markerów DArTseq do klonów BAC zastosowała algorytm BLAST oraz uzyskane informacje o położeniu markerów na mapie genetycznej.

Natomiast sekwencjonowanie indywidualnych klonów BAC przeprowadzone zostało z wykorzystaniem dwóch metod sekwencjonowania nowej generacji: Illumina oraz Oxford Nanopore. Dwa z klonów BAC zostały zsekwencjonowane obiema metodami.

Na uwagę zasługuje wykorzystanie wielu narzędzi bioinformatycznych, które umożliwiły uzyskanie informacji na temat obecności w klonach BAC homologicznych do innych gatunków genów związanych z udomowieniem, czy odpornością na choroby oraz projektowanie starterów służących od ich weryfikacji. Zwykle weryfikację obecności tych genów przeprowadzała eksperymentalnie oraz bioinformatycznie. Startery do weryfikacji spodziewanych genów w klonach BAC konstruowała posiłkując się sekwencjami z bazy danych NCBI oraz wykorzystywała sekwencje zdeponowane w bazie GSS banku genów GenBank.

Zastosowanie programu FGENESH umożliwiło wstępne określenie sekwencji i struktury nowo zidentyfikowanych genów potencjalnie związanych z odpornością na choroby u żyta oraz związanych z udomowieniem. Natomiast program Bioedit posłużył do ich edycji i wizualizacji sekwencji. Autorka rozprawy wykonała także analizę ekspresji genu odporności na porażenie przez rdzę brunatną *Puccinia recondita* f. sp. *secalis* wykorzystując w tym celu

materiał roślinny dwóch linii wsobnych żyta poddanych inokulacji zarodnikami tego patogena.

Zaproponowane i opisane w powyższym rozdziale procedury postępowania umożliwiły pełną realizację postawionych celów badawczych i nie budzą żadnych zastrzeżeń.

Wyniki ocenianej rozprawy mgr Ewy Borzęckiej są bardzo nowatorskie i dostarczają nowych informacji dotyczących genomu żyta. W pracy zaprezentowano wykorzystanie biblioteki BAC żyta do integracji map fizycznych i genetycznych. Po raz pierwszy do kotwiczenia klonów BAC na mapie genetycznej zostały wykorzystane markery typu GBS - DArTseq. W wyniku genotypowania 16 superpul klonów BAC metodą DArTseq zidentyfikowała 221 122 potencjalnych adresów marker DArTseq-klon BAC, z których wybrano 53 370 zawierających 41 438 żytnich markerów DArTseq. Długość żytnich markerów DArTseq, które znalazły się w adresach, wynosiła od 20 do 69 nukleotydów i zostały one przyporządkowane do 19 695 klonów BAC. Liczba zakotwiczonych klonów BAC była różna dla różnych chromosomów, a ich analiza wykazała ich nierównomierne rozmieszczenie wzdłuż mapy genetycznej żyta. Najwięcej klonów BAC znalazło się na chromosomach 6R i 7R, a najmniej na chromosomie 2R.

W pracy Doktorantka przedstawiła również implementację i optymalizację rozwiązań metodycznych, takich jak sekwencjonowanie pojedynczych klonów BAC metodą Nanopore - ułatwiające wykorzystanie bibliotek BAC w badaniach z zakresu genetyki i genomiki. Autorka dysertacji wykazała, że sekwencjonowanie nowej generacji Oxford Nanopore z użyciem urządzenia MinION daje w analizowanych klonach BAC pełną ciągłość sekwencji zawierającą wektor wraz z insertem, a identyczność tych sekwencji w porównaniu z sekwencją zdeponowaną w bazie danych GenBank wynosiła ponad 99%, natomiast identyczność sekwencji uzyskanych metodą Illumina i Nanopre wynosiła około 99%, a liczba wykrytych w klonach BAC polimorfizmów SNP wynosiła od 6 do 19, a liczba polimorfizmów typu InDel - od 18 do 81.

Biblioteka BAC została również wykorzystana przez Doktorantkę do identyfikacji genów związanych z udomowieniem i odpornością na choroby żyta. Analiza wszystkich sekwencji markerów DArT, skojarzonych na podstawie analiz bioinformatycznych z funkcjami związanymi z odpornością na choroby pozwoliła na identyfikację 11 klonów BAC przypuszczalnie zawierających przynajmniej dwa markery DArT powiązane z odpornością na choroby. Zaprojektowano do nich startery sprawdzone eksperymentalnie pozwoliły na potwierdzenie w 7 z nich sekwencji markerów DArT związanych odpornością na choroby, a średnie pokrycie porównywanych fragmentów wynosiło około 90%. Przeprowadzona analiza programem SoftBerry/FGENESH sekwencji klonów BAC pozwoliła na zidentyfikowanie

przypuszczalnej pełnej sekwencji i struktury obecnych w nich genów, w tym genów potencjalnie związanych z odpornością na choroby u żyta.

Bardzo cennych informacji dostarczyła analiza BLAST z wykorzystaniem bazy danych RefSeq z NCBI, która wskazała, że geny występujące w jednym z klonów BAC są podobne do genów, które mogą brać udział w odpowiedzi na porażenie roślin patogenem wywołującym rdzę brunatną. Spośród wytypowanych genów sekwencja genu nazwanego *Sck33* wykazała najwyższe podobieństwo z sekwencjami genów *Lrk 33*, *Lrk19* oraz *Tak33* pszenicy. Podobieństwo to wyniosło ponad 93%, ze średnią długością dopasowania porównywanych fragmentów przewyższającą 80%. Gen ten wykazał także wysokie podobieństwo do sekwencji genu *Lr10-like* zwłaszcza w rejonie kodującym, a reakcja roślin żyta z tym genem na porażenie przez zarodniki rdzy brunatnej wykazywały podobieństwo do reakcji roślin z genem *Lr 1*. Dla tego genu zaprojektowano pary starterów, które poddano weryfikacji i jedna z tych par może być wykorzystywana do identyfikacji tego genu w materiałach genetycznych żyta. Poziom ekspresji tego genu u badanych roślin żyta poddanych inokulacji przez zarodniki rdzy brunatnej wykazywał istotną różnicę w stosunku do roślin kontrolnych, co może sugerować, że gen ten może brać udział w odpowiedzi rośliny na infekcję przez ten patogen.

Ponadto Autorka rozprawy zidentyfikowała 11 genów, które przypuszczalnie związane są z udomowieniem, dla których także zaprojektowała startery do amplifikacji fragmentów tych genów z genomem żyta. Na uwagę zasługuje także opracowanie struktury kilku genów potencjalnie zaangażowanych w proces udomawiania u żyta wykazujących homologię do genów innych gatunków zbóż, a także przypuszczalne struktury genów związanych z reakcją odporności zidentyfikowanych w klonach BAC.

W opisanych badaniach Doktorantka wykorzystwała bibliotekę BAC linii wsobnej żyta L318, a podobieństwo sekwencji i pokrycie porównywalnych fragmentów z genomem referencyjnym żyta nie zawsze było wysokie i dotyczyło przede wszystkim krótkich fragmentów genów. Proszę o wyjaśnienie skąd mogą być niskie dopasowania sekwencji lub mniejsze bądź większe niezgodności.

Mam także pytanie dotyczące czynników, które mogą być przyczyną błędnej predykcji bioinformatycznej przypuszczalnych genów, co ma miejsce np. w przypadku homologów u innych pokrewnych gatunków.

Rozprawę kończy obszerna i dobrze przeprowadzona dyskusja, w której Doktorantka zestawiała uzyskane wyniki z dostępnymi w literaturze badaniami innych Autorów, odnosząc się kolejno do wszystkich etapów prowadzonych badań własnych. Ten rozdział został opracowany przez Doktorantkę bardzo starannie, zawiera dogłębną analizę wyników własnych,

co znacząco podnosi walory naukowe dysertacji i w całej rozciągłości zasługuje na wyróżnienie. Podsumowaniem jest sześć wniosków wyciągniętych na podstawie przeprowadzonych badań i analiz, które mają potwierdzenie w przedstawionej dokumentacji wyników.

W trakcie opracowania rozprawy Doktorantka nie ustrzegła się drobnych błędów językowych, literowych, skrótów myślowych (np. gen produkcji chleba pszennego) czy przejęzyczeń, które zaznaczyłem w tekście przesłanego do oceny maszynopisu, a które mogą być z łatwością usunięte w trakcie przygotowania Autoreferatu. Zauważone w pracy drobne niedociągnięcia nie umniejszają jej bardzo wysokiej wartości merytorycznej, poznawczej i naukowej.

Wniosek końcowy

Przedstawiona do oceny rozprawa doktorska wnosi nowe elementy wiedzy w dziedzinie genetyki i genomiki żyta oraz znacząco poszerza dotychczasową wiedzę na temat przydatności technologii DArTseq do efektywnego zakotwiczenia klonów BAC na mapie genetycznej żyta, wykorzystania biblioteki BAC do identyfikacji genów związanych z odpornością na choroby oraz udomowieniem u żyta. Zaprezentowane wyniki dają możliwość analizy rejonów genomu, na które wywierana była presja selekcyjna podczas udomowienia i hodowli żyta. Ponadto w pracy przedstawiono implementację i optymalizację rozwiązań metodycznych, takich jak sekwencjonowanie pojedynczych klonów BAC metodą Nanopore - ułatwiających wykorzystanie bibliotek BAC w badaniach z zakresu genetyki i genomiki. Uzyskane wyniki posiadają wysokie walory poznawcze i naukowe i stanowią punkt wyjścia do dalszych badań nad funkcją zidentyfikowanych genów, co przyczyni się do lepszego poznania genetycznej kontroli istotnych funkcji i cech użytkowych u żyta.

Uważam, że temat podjęty przez Doktorantkę jest bardzo aktualny, a pytania postawione w celach badawczych sformułowano w sposób jasny i prawidłowy. Dzięki zastosowaniu właściwych metod wytyczone cele zostały w pełni zrealizowane, postawione hipotezy badawcze zweryfikowane, a uzyskane wyniki i wnioski znajdują pełne potwierdzenie w materiale dokumentacyjnym przedstawionym w tabelach, na rysunkach, oraz mają bardzo wysoką wartość naukową.

Zgodnie z ustawą Prawo o szkolnictwie wyższym i nauce (Dz. U. z 2018 r. poz. 1668, z późn. zm.) praca mgr Ewy Borzęckiej pt. „Wykorzystanie biblioteki BAC do integracji map fizycznych i genetycznych oraz identyfikacji genów związanych z udomowieniem i

odpornością na choroby żyta” spełnia wszystkie wymogi stawiane rozprawom doktorskim przez, dlatego stawiam wniosek do Rady Dyscypliny Nauki Biologiczne Szkoły Głównej Gospodarstwa Wiejskiego w Warszawie o dopuszczenie Jej Autorki do dalszych etapów przewodu doktorskiego.

Ze względu na wagę podjętego tematu, zakres przeprowadzonych badań, zastosowanie nowoczesnych metod badawczych, efektywne wykorzystanie istniejących w Katedrze zasobów genomicznych, kompleksowość w podejściu do rozwiązywanego zagadnienia, a także opanowanie i zastosowanie wielu narzędzi bioinformatycznych oraz wdrożenie sekwencjonowania nowej generacji Nanopore, a przede wszystkim za wysoką wartość naukową uzyskanych wyników, wnioskuję o wyróżnienie rozprawy doktorskiej nagrodą.



Prof. dr hab. Henryk Bujak

Wrocław, dnia 5.11.2022 r.

