

Sprawozdanie z realizacji zadania badawczego nr 11 za 2021 rok

Tytuł zadania **Identyfikacja mechanizmu molekularnego odporności żyta ozimego na rdzę brunatną.**

Kierownik zadania: **dr hab. inż. Kamila Nowosad , prof. Uczelni**

Celem prac realizowanych w 2021 roku było:

- 1) Wytypowanie genotypów żyta o największej tolerancji lub całkowitej odporności na porażenie rdzą brunatną (*Puccinia recondita*).
- 2) Wytworzenie populacji mapującej żyta ozimego.
- 3) Identyfikacja genów o zmienionej ekspresji pod wpływem infekcji patogena.

W roku 2021 wykonano następujące prace:

Ad 1)

Materiał badawczy stanowiły przekazane przez hodowców materiały wyjściowe do hodowli nowych odmian żyta ozimego, linie wsobne z kolekcji oraz zarejestrowane odmiany, które zostały użyte jako wzorce.

1. Formy żyta ozimego przekazane przez:

1.1. DANKO Hodowla Roślin Sp. z o. o.

LS 48N/21, LS 48N/21, LS 57N/21, LS 72N/21, LS 72N/21, LS 171N/21, LS 184N/21, LS 202N/21, LS 226N/21, LS 227N/21, LS 223N/21, LS 237N/21, LS 244N/21, LS 260N/21, LS 272N/21, LS 286N/21, LS 293N/21, LS 312N/21, LS 340N/21, LS 389N/21, S 82P/05, SE 8P/10, SE31P/11, SE47P/11, SE62P/12, SE67P/12, SE71P/12, SE96P/16, SE98P/16, SE101P/16, SE104P/16, SE105P/16, SE118P/17, SE122P/17, SE125P/17, SE128P/17, SE141P/17, SE144P/17, SE150P/17, SE174P/17, SO123R/19, SO125R/19, SO126R/19, SO143R/19, SO157R/19, SO161R/19, SO163R/19, SO182R/19, SO131R/19, SO133R/19, SO135R/19, SO141R/19, SO145R/19, SO150R/19, SO154R/19, SO168R/19, SO189R/19, SO191R/19, SO201R/19.

1.2. Poznańska Hodowla Roślin Sp. z o.o.

PHR_R1 - PHR_R60

1.3. Hodowla Roślin Smolice Sp. z o. o. Grupa IHAR

Bojko-18, Bojko-30, Rostockie-75, 45-4Rj-71, 46-4Rj-71, 97-2Rjz-1-02, 100-2Rj-1-72, 100-2Rj-1-74, 100-2Rj-26-04, 101-2Rj-11-73, 108-2Rj-71, 108-2Rj-72, 124-4R-71, 143-2Rj-572, 164-Rj-5-71, 164-Rj-6-71, 168-2R-71, 209-4R-71, 212-4R-71, 214-4R-71, 228-3R-71, 318-3R-81, 363-R-71, 364-2R-81, 427-2R-81, 479-R-81, 46-4Rjz-1-81, 98-2Rj-3-71, 98-2Rj-3-72, 100-2Rjz-2-02, 101-2Rj-71, 164-Rj-6-72, 189-4R-71, 198-R-71, 202-4R-71, 250-3R-72, 256-R-J-81, 257-R-2-02, 274-R-4-82, 281-3R-82, 298-3R-82, 323-3R-81, 323-3R-82, 347-R-81, 372-2R-83, 402-2R-82, 460-R-81, 469-R-81, 475-R-81.

W roku 2021 testowaniu w kierunku poszukiwania odporności na rdzę brunatną poddano łącznie 170 genotypów żyta ozimego w doświadczeniach infekcyjnych. Były to dostarczone przez hodowców linie wsobne, populacje oraz odmiany wzorcowe.

Ocenę podatności genotypów żyta na porażenie przez rdzę brunatną (*Puccinia recondita*) wykonano metodą laboratoryjną. Na szlakach Petriego przygotowano pożywkę składającą się z agaru z dodatkiem benzomidazolu. Liście dziesięciodniowych siewek układano na pożywce, a następnie infekowano przygotowanym wcześniej inokulatem rdzy brunatnej. Materiał do przygotowania zawiesiny inokulacyjnej zebrano w różnych miejscowościach (m.in: Wrocław, Smolice, Kondratowice, Radzików), aby otrzymać szeroką genetycznie populację zarodników patogena. Oceny porażenia wykonano po dziesięciu dniach od inokulacji, w skali czterostopniowej, gdzie 1 – brak porażenia, a 4 - silne objawy porażenia rdzą brunatną na liściu.

Wyniki ocen zostały poddane analizom statystycznym. W celu prawidłowego statystycznego opracowania otrzymanych danych eksperymentalnych przeprowadzono ich transformację, w celu spełnienia warunków koniecznych do analiz, czyli uzyskania ciągłości i normalności rozkładu empirycznego. Wyniki surowe przekształcono wykorzystując wzór podany przez Węgrzyna i in. (1996), a następnie przeprowadzono jednoczynnikową analizę wariancji zgodnie z modelem kompletnej randomizacji (ANOVA), co umożliwiło sprawdzenie istotności zróżnicowania badanych genotypów żyta ozimego pod względem odporności na mączniaka prawdziwego i rdzę brunatną. Do grupowania genotypów o tej samej wartości średnich, czyli do ich podziału na grupy jednorodne wykorzystano wielokrotny test Duncana.

W doświadczeniach infekcyjnych wykonano oceny podatności posiadanych form żyta na porażenie przez *Puccinia recondita*. Oceną objęto łącznie 170 genotypów żyta pochodzących z firm hodowlanych.

Wyniki ocen dla badanych genotypów żyta zostały poddane analizie statystycznej. Do analizy wariancji oraz podziału genotypów na grupy jednorodne wykorzystano dane po transformacji, które po pogrupowaniu zostały z powrotem przeliczone na skalę oceny.

Analiza wariancji wykazała istotne zróżnicowanie badanych genotypów żyta pod względem ich odporności na rdzę brunatną, to znaczy że stopień porażenia liści przez rdzę brunatną w warunkach sztucznej inokulacji był zróżnicowany w zależności od genotypu. Wśród 170 badanych genotypów żyta, znaleziono 28 obiektów które charakteryzowały się pożądaną, wysoką odpornością na tego patogena. W zestawieniu wiele z badanych genotypów było podatnych na porażenie i zostały one ocenione na 3 lub 4 w stosowanej czterostopniowej skali oceny porażenia. Wystąpiło jednak 28 obiektów o znikomym porażeniu lub bez śladów zarodników na liściach (**164-Rj-6-71, 101-2Rj-71, SO182R/19, SO154R/19, PHR_R23, PHR_R39, 45-4Rj-71, 164-Rj-5-71, 214-4R-71, 164-Rj-6-72, 202-4R-71, 323-3R-82, SE122P/17, PHR_R21, PHR_R16, PHR_R21, PHR_R24, PHR_R52, 100-2Rj-1-72, 100-2Rj-1-74, 108-2Rj-72, 98-2Rj-3-72, 100-2Rjz-2-02, 189-4R-71, 256-R-J-81, 298-3R-82, 469-R-81, LS 56N/21**), które można uznać za odporne na rdzę brunatną. Genotypy te utworzyły pierwszą grupę jednorodną. Kolejne grupy jednorodne tworzyły genotypy, które w teście laboratoryjnym uzyskały oceny świadczące o większej wrażliwości na porażenie przez rdzę brunatną. Z informacji literaturowych wynika, że istnieje wysoka korelacja pomiędzy odpornością określoną metodą laboratoryjną, a polową odpornością na rdzę brunatną. Wymienione genotypy można będzie wykorzystać w programach hodowli nowych, tolerancyjnych na rdzę brunatną odmian żyta.

Uzyskane w bieżącym roku wyniki ocen porażenia genotypów żyta przez rdzę brunatną nie odbiegają znacząco do rezultatów z innych badań. Wśród analizowanych genotypów żyta ozimego stwierdzono nieliczne formy charakteryzujące się pełną odpornością rdzę brunatną. Zakres zmienności porażenia badanych linii przez patogena wynosił od 1,0 do 4,0. Analiza statystyczna otrzymanych wyników wykazała istotne zróżnicowanie podatności na porażenie rdzę brunatną badanych genotypów. Wśród badanych genotypów wystąpiły formy odporne na *Puccinia recondita*. Podobne wyniki uzyskano w badaniach linii wsobnych wyprowadzonych w Katedrze Genetyki, Hodowli Roślin i Nasiennictwa

Uniwersytetu Przyrodniczego we Wrocławiu. Linie wsobne żyta ozimego wchodzące w skład kolekcji charakteryzowały się również istotnym zróżnicowaniem podatności na rdzę brunatną.

1. Analizy wyników oceny porażenia genotypów żyta przez rdzę brunatną (*Puccinia recondita*) w wyniku sztucznej inokulacji wykazały, że istnieją genotypy odporne lub wysoce tolerancyjne na tego patogena, które można polecić do programów hodowli odpornościowej.

2. Udało się znaleźć stosunkowo dużą grupę genotypów bez śladów zarodników rdzy brunatnej na liściach (28 genotypów), które można uznać za odporne na porażenie przez tego patogena.

3. Genotypy odporne mogą zostać wykorzystane do założenia przez hodowców szkółek rozmnożeniowych i selekcyjnych w celu uzyskania z nich nowych cennych linii o kompleksowej

Ad 2)

Materiał badawczy stanowiły linie wsobne (homozygotyczne) o potwierdzonej odporności na rdzę brunatną, jak i linie silnie podatne na porażenie. Linie podatne o sprawdzonej odporności zostały wyprowadzone z różnych populacji wyjściowych i w wyniku wieloletniego chowu wsobnego stanowią homozygoty. Linie odporne zostały przetestowane w warunkach polowych w latach poprzedzających wykonanie doświadczenia. Spośród linii wytypowanych przez hodowców jako odporne, pełną odporność w warunkach polowych udało się potwierdzić tylko dla jednej formy RB1/ TN1874 pochodząca z kolekcji Instytutu Hodowli i Aklimatyzacji Roślin w Radzikowie. Linia ta została wybrana jako komponent do krzyżowania.

Linie zostały wysiane na poletkach doświadczalnych Rolniczego Zakładu Doświadczalnego Swojec Uniwersytetu Przyrodniczego we Wrocławiu i przekrzyżowane, w celu uzyskania nasion pierwszego pokolenia mieszańcowego F₁. Z uwagi na to iż, linie mogą posiadać geny samoniezdgodności, co w przypadku żyta jest dość powszechne, które mogą w znaczny sposób utrudnić krzyżowania i uzyskanie ziarniaków mieszańców wykonano zostało krzyżowanie 10 linii. Układ krzyżowań obejmował 9 linii podatnych krzyżowanych w układach w obie strony z linią odporna RB1.

Efekty krzyżowania były zadowalające pod względem otrzymanej liczby ziarniaków mieszańcowych F₁ dla trzech kombinacji:

1. SE 128N/17 x RB1
2. RB1 x UP122
3. RB1 x SE 128N

Ziarniaki mieszańców zostały wysiane w celu wyprowadzenia populacji mapującej. Nasiona wysiewano punktowo w rozstawie 20 x 5 na poletkach dwurzędowych. Na każdym z poletek wysiano po 40 ziarniaków. Kombinacje mieszańcowe po przezimowaniu, zostaną ocenione i wybrana zostanie jedna kombinacja. Wybrana populacja F₁ zostanie w kolejnym roku rozmnożona celem wyprowadzenia z niej populacji mapującej o odpowiedniej liczebności do dalszych badań.

Zastosowanie różnych kombinacji linii stanowiło dobre rozwiązanie aby uniknąć efektu samoniezdgodności u roślin żyta ozimego. Wykonanie krzyżowań linii wsobnych w układzie A x B i BxA, z 10 liniami pozwoliło na uzyskanie odpowiedniej liczby ziarniaków do wysiewu i wytworzenia populacji mapującej.

1. Wybrane kombinacje linii wsobnych żyta ozimego pozwoliły na wytworzenie populacji F₁.
2. Wysiane populacje zostaną poddane ocenie po przezimowaniu. Wybrana zostanie jedna najliczniejsza populacja F₁, która posłuży do rozmnożenia i uzyskania populacji mapującej F₂.

Ad 3)

Materiał do analiz RNAseq stanowiło całkowite RNA wyizolowane z młodych liści genotypów żyta. Materiał izolowany był 24 godziny po inokulacji 14 dniowych roślin wodnym roztworem zarodników

rdzy brunatnej. Rośliny kontrolne traktowane były wodą destylowaną. W badaniach wykorzystano dwa genotypy, jeden wykazujący połowę odporność na porażenie rdzą i drugi podatny na porażenie. Materiał pobrany został w trzech powtórzeniach biologicznych.

Całkowite RNA do analiz wyizolowano przy użyciu kitu Total RNA Mini (A&A Biotechnology). Ilość oraz jakość (integralność) wyizolowanego materiału sprawdzono na urządzeniu Qiaxcel Advance wykorzystując zestawu RNA QC (Qiagen). Biblioteki do sekwencjonowania uzyskano korzystając z zestawu TruSeq Stranded mRNA LT Sample Prep Kit (Illumina) według protokołu TruSeq Stranded mRNA Sample Preparation Guide, Part # 15031047 Rev. E, co pozwoliło na uzyskanie bibliotek o znanej orientacji nici (strand sensitive). Sekwencjonowanie wykonano na urządzeniu HiSeq2500 w technologii 150 PE. Wyniki sekwencjonowania zapisane zostały jako pliki fastq (zawierające informację sekwencyjną oraz informację o jakości uzyskanych odczytów). Otrzymane surowe wyniki poddane zostały analizie jakości. W tym celu wykorzystany został program *fastqc*. Odczyty słabej jakości $Q < 30$ oraz adaptory usunięte zostały za pomocą programu Trimmomatic (Bolger, Lohse, and Usadel 2014).

Oczyszczone dane (clean data) zostały zmapowane do genomu referencyjnego (Rabanus-Wallace i in. 2021) pobranego ze strony [GenBanku](#) przy wykorzystaniu alignera wrażliwego na slipping HiSat2 (Kim et al. 2019). Utworzone pliki SAM z informacją o położeniu poszczególnych odczytów przekształcone zostały w posortowane pliki BAM za pomocą narzędzia samtools (Li 2011). Tak przygotowane pliki BAM posłużyły do przygotowania macierzy zaliczeń odczytów przypisanych do poszczególnych genów za pomocą programu featureCounts z pakietu [Subread](#). Uzyskane wyniki ekspresji zostały znormalizowane według metody TPM (Conesa et al. 2016) za pomocą własnego skryptu w programie R. Analiza porównawcza ekspresji genów (DE) wykonana została w pakiecie DESeq2 (Love, Anders, and Huber 2019) w programie R (Team and others 2013). Do identyfikacji ortologów genów zaangażowanych w proces odpowiedzi rośliny na infekcję patogenu grzybowego został wykorzystany program BLAST (Altschul et al. 1990) i algorytm tblastn. Transkryptom referencyjny żyta (Hane et al. 2017) przeszukiwano za pomocą sekwencji białkowych wybranych genów pobranych z bazy [UniProt](#). Analizy GO wykonane zostały przy użyciu pakietu GOSec (Young et al. 2012) w programie R. Pakiet GOSec uwzględnia w analizie wpływ długości genu i bierze poprawkę na tą wartość. W analizie GO wykorzystano test Walleniusa. Za istotnie nadreprezentowane uznano numery GO dla których wartość $p < 0.05$. Na podstawie znormalizowanych danych ekspresji metodą TPM wykonano mapę PCA dla 1 i 2 składowej głównej dla 6 badanych próbek (3 Genotyp I, 3 Genotyp II). Na podstawie znormalizowanych danych ekspresji genów skonstruowano także mapę cieplną obrazującą współczynniki korelacji pearsona pomiędzy poszczególnymi próbkami. PCA i mapę cieplną wykonano w programie R. Analizy bioinformatyczne i obliczenia statystyczne wykonane zostały na serwerze „bioscienta” należącym do KGHRiN.

Dla wszystkich próbek udało się wyizolować RNA w wystarczającej ilości (6 – 32 g, Tab 1) do przeprowadzenia analizy RNAseq. Wyizolowane RNA było wysokiej jakości (wartości wsp. RIN > 8, tab 1). Uzyskane wyniki wskazują, że dobrana metoda izolacji była odpowiednia dla analizowanego materiału badawczego.

Tabela 1. Wyniki izolacji RNA - zestawienie

ID próbki	Ilość wyizolowanego RNA	RIN
UP1	19.539	7.9
UP2	14.019	8.1
UP3	31.131	8.2

UP4	14.019	8.2
UP5	17.397	7.3
UP6	6.477	8.3

Uzyskanie wysokiej jakości RNA było warunkiem koniecznym do przeprowadzenia sekwencjonowania transkryptomu. Ponieważ uzyskane RNA było wysokiej jakości, proces tworzenia bibliotek przebieg bez zakłuceń i udało się uzyskać wysokiej jakości biblioteki, które następnie z powodzeniem zostały zsekwencjonowane. Uzyskane wyniki sekwencjonowania były bardzo dobre. Dla każdej biblioteki uzyskano 31 - 32 mln odczytów i 6,2 - 5,4 Gpz odczytanych zasad. Większość (ok. 95%) uzyskanych odczytów miała wartości Phred (jakości odczytu) powyżej 30, co wskazuje na udane sekwencjonowania. Podsumowanie wyników sekwencjonowania zawiera tabela 2. Uzyskanie wysokiej jakości odczytów sekwencjonowania pozwoliło na wykonanie analiz porównania ekspresji (DE).

Tabela 2. Wyniki sekwencjonowania - zestawienie

ID biblioteki	Łączna liczba odczytów	Łączna liczba odczytanych zasad	CG	Q30..
UP1	41,195,974	6,220,592,074	58.27	95.21
UP2	41,557,290	6,275,150,790	58.21	95.11
UP3	41,316,498	6,238,791,198	57.99	95.16
UP4	41,991,232	6,340,676,032	58.54	95.31
UP5	40,943,890	6,182,527,390	58.8	95.21
UP6	42,144,012	6,363,745,812	57.2	95.19

Mapowanie do genomu referencyjnego

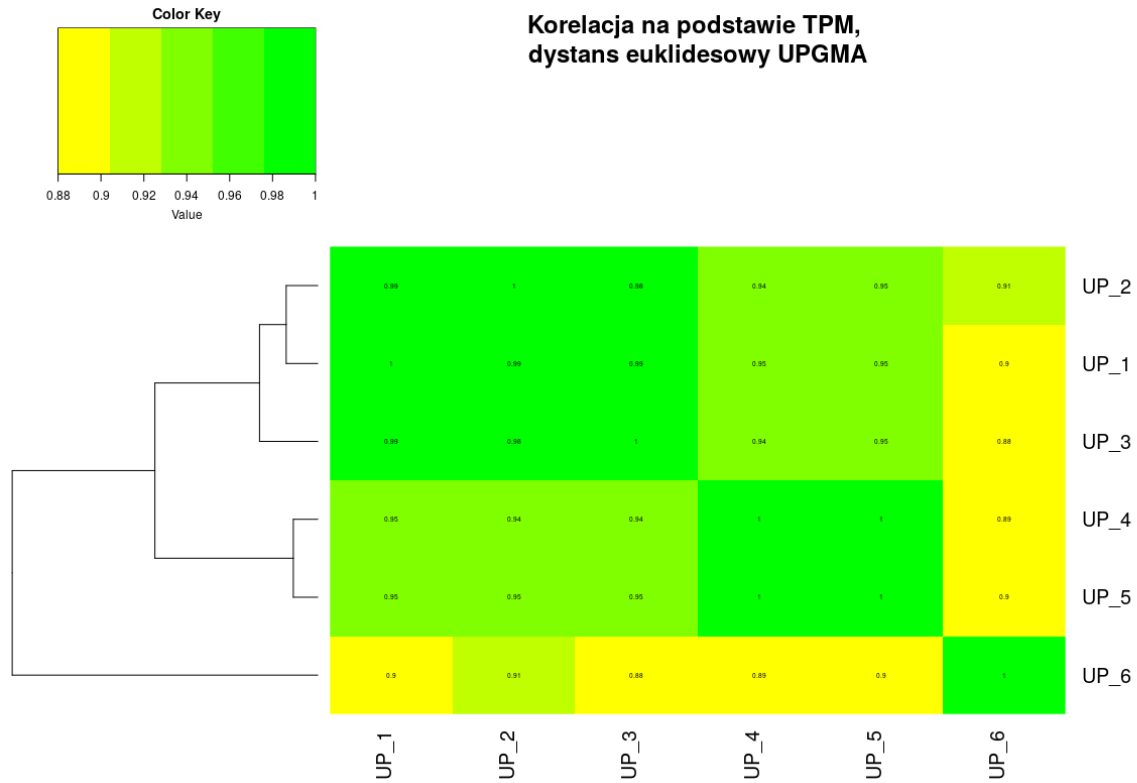
Przeprowadzone mapowanie oczyszczonych wyników pozwoliło na przypisanie większości odczytów do pojedynczego miejsca w genomie. Dla wszystkich analizowanych próbek co najmniej 80% z nich zostało zmapowanych do pojedynczej lokalizacji w genomie referencyjnym. Tak przygotowane wyniki mogły zostać wykorzystane w dalszych analizach ekspresji porównawczej (Tabela 3).

Tabela 3. Wyniki mapowania - zestawienie

ID biblioteki	Liczba odczytów	% nie zmapowanych	% jedno mapowanie	% więcej niż jedno mapowanie
UP1	20,004,438	7.78%	87.03%	5.19%
UP2	20,145,786	8.17%	86.71%	5.12%
UP3	20,046,240	7.77%	86.22%	6.01%
UP4	20,426,039	7.02%	85.88%	7.1%
UP5	19,894,533	7.06%	86.61%	6.33%
UP6	20,503,083	7.25%	80.68%	12.07%

Analiza ekspresji różnicowej (DE)

Po wczytaniu danych do R w pakiecie DESeq2 obliczono współczynniki normalizacji biblioteki (Tab. 4). Wszystkie analizowane biblioteki miały podobną głębokość sekwencjonowania co ilustruje współczynnik normalizacji bliski 1. Wyniki uzyskane w obrębie powtórzeń biologicznych było do siebie zbliżone co ilustruje rysunek 1. Potwierdza to poprawność wykonanych prac laboratoryjnych i podnosi wiarygodność uzyskanych wyników.

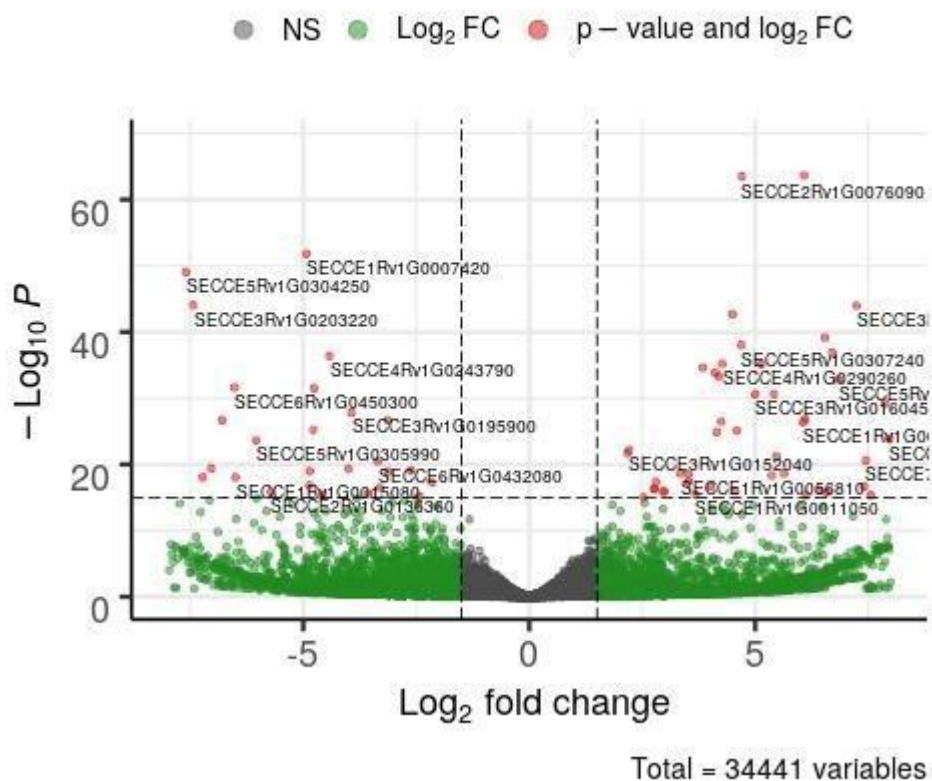


Rysunek 1. Wykres wartości korelacji między analizowanymi próbkami

Wykonana analiza porównawcza ekspresji pozwoliła na wskazanie 2296 genów z istotną statystycznie ($FDR < 0,01$) zmianą ekspresji (2x lub więcej wyższa i niższa) w roślinach uzyskanej metodami klasycznymi. Uzyskane wyniki ilustruje wykres typu „Volcano plot” (rys. 2) Zestawienie genów z zmienioną ekspresją zawiera suplement (Stab1).

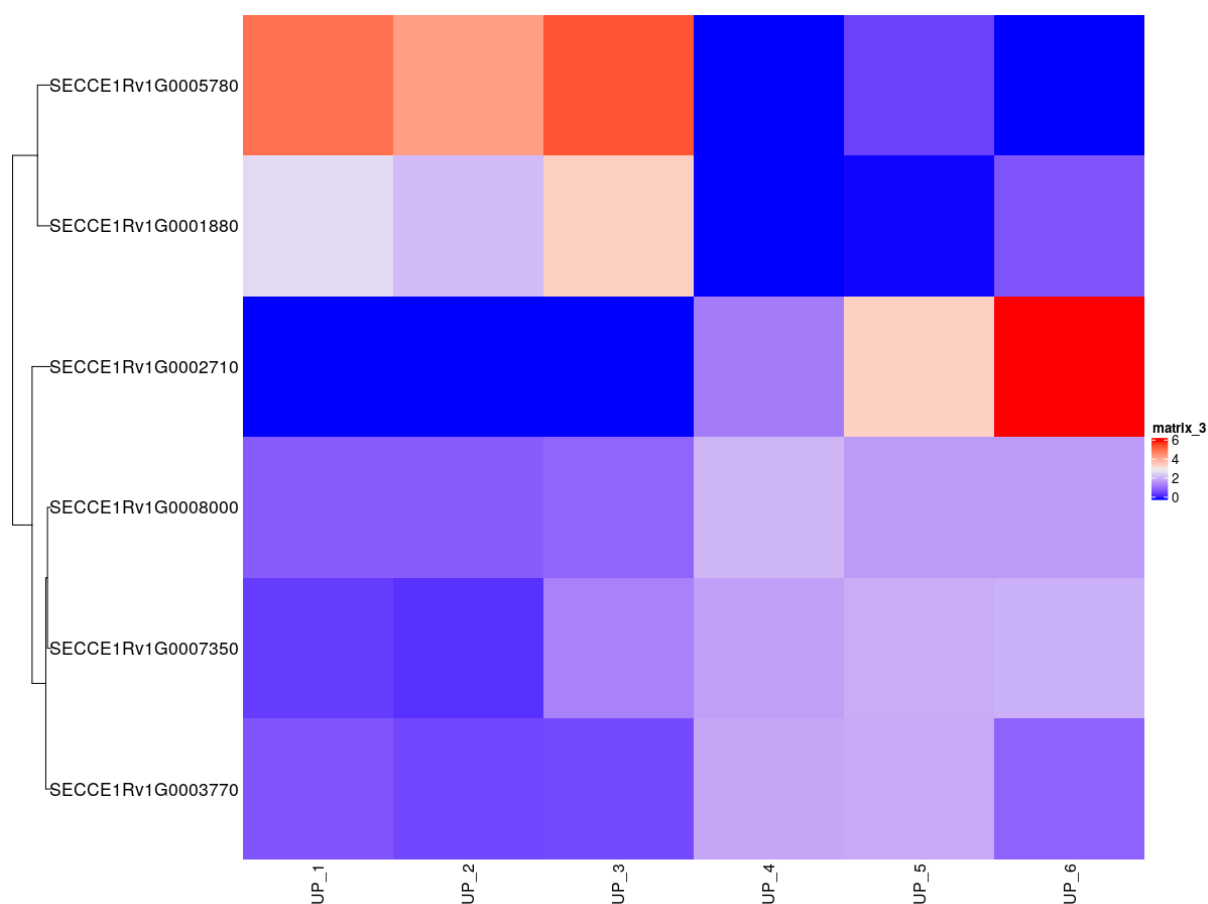
Tabela 4. Wyniki normalizacji bibliotek

ID biblioteki	wsp. normalizacji
UP1	1,05
UP2	0,93
UP3	1,07
UP4	0,93
UP5	0,94
UP6	1,16



Rysunek 2. Wykres typu "Volcano plot ilustrujący wyniki ekspresji różnicowej DE.

Wśród tych transkryptów znalazły się między innymi geny związane z udokumentowaną odpowiedzią roślin na infekcję patogenu takie jak np.: SECCE1Rv1G0001880; Pm3-like disease resistance protein, SECCE1Rv1G0002710; Cyclotide - peptyd który pełni rolę obronną przed grzybami u roślin, SECCE1Rv1G0003770; Disease resistance protein (TIR-NBS-LRR class) family, SECCE1Rv1G0005780; RNA-binding (RRM/RBD/RNP motifs) family protein - białko opisane jako związane z odpornością na grzyby, SECCE1Rv1G0007350; CsAtPR5 - gen opisany jako jeden z genów PM u pszenicy, SECCE1Rv1G0008000; GDSL esterase/lipase 1 - gen opisany jako gen odporności na grzyby. Na rysunku 3 przedstawiono poziom ekspresji wybranych genów z udokumentowanym udziałem w odpowiedzi roślin na infekcję patogenu.



Rysunek 4. Wykres - mapa cieplna ilustrujący poziom ekspresji wybranych genów w badanych próbkach.

W kolejnym kroku odczytu uzyskane w analizach RNAseq zostały wykorzystane do opracowania transkryptomu *de novo*. Analizy przeprowadzone zostały w programie Trinity (<https://github.com/trinityrnaseq/trinityrnaseq>) z przyjętymi parametrami domyślnymi. W wyniku działania programu został odtworzony transkryptom. Otrzymany transkryptom zawierał 133176 transkryptów oraz 83198 genów. Średnia długość otrzymanych transkryptów wynosiła 902,4 pz, a średnia długość genów 700,6 pz. Następnie za pomocą narzędzi dostępnych w programie Trinity została przeprowadzona kontrola jakości otrzymanego transkryptomu. W tym celu obliczony został parament E90N50 - Contig N. Wyniki tych analiz zestawiono w tabeli 5.

Tabela 5: Statystyki Contig N dla opracowanego transkryptomu

Statystyka bazująca na kontigach całego transkryptu:		Statystyka bazująca jedynie na najdłuższych izoformach 'na gen'	
Contig N10:	3628	Contig N10:	3380
Contig N20:	2787	Contig N20:	2522
Contig N30:	2262	Contig N30:	2032

Contig N40:	1887	Contig N40:	1655
Contig N50:	1568	Contig N50:	1296

Dodatkowo odczyty uzyskane w RNAseq zostały zmapowane do otrzymanego transkryptomu za pomocą programu bowtie2. Dla każdej z analizowanych próbek ponad 95% odczytów mapowało się do opracowanego transkryptomu. Na podstawie uzyskanych wyników przyjęto, iż opracowany transkryptom jest dobrej jakości i może z powodzeniem zostać wykorzystany w dalszych badaniach (Conesa i in 2016). Otrzymane sekwencje dla opracowanego transkryptomu zapisano w pliku Żyto_transkryptom.fasta W kolejnym kroku otrzymany transkryptom został zanotowany na podstawie baz sekwencji GenBank oraz SweetProt. Do tego celu został wykorzystany program InterProScan. Przed uruchomieniem programu przygotowany został plik wsadowy poprzez identyfikację sekwencji kodujących (CDS) w opracowanym transkryptomu, a następnie przetłumaczenie tych sekwencji na sekwencje aminokwasowe. Anotacja została przypisana do 64244 transkryptów. Tak przygotowany transkryptom został wykorzystany w identyfikacji sekwencji mikrosatelitarnych. Do tego celu użyty został program MISA (Baldwin i in. 2012). W wyniku tych analiz zidentyfikowanych zostało 34119 loci SSR w obrębie 19776 genów, 7562 genów zawierało więcej niż jedno locus SSR. Z analizowanych motywów SRR najczęściej zidentyfikowano powtórzeń trzynukleotydowych - 16645, a najmniej monomorficznych 1662. Szczegółowe zestawienie przedstawia tabela 6. Pełna lista zidentyfikowanych motywów SSR wraz z ich lokalizacją zestawiona została w pliku Żyto_SSR.gff.

Tabela 6: Zestawienie liczby zidentyfikowanych sekwencji SSR.

Długość motywu	Liczba zidentyfikowanych sekwencji
1	1662
2	3204
3	16645
4	7806
5	2580
6	22222

Analiza markerów SNP polimorficznych pomiędzy formami rodzicielskimi populacji mapującej wykonana została według wytycznych najlepszych praktyk opracowanych przez Instytut Broad MIT i Harvardu. Zmapowane odczyty dla każdej próbki zostały posortowane po pozycji w transkryptomie. Tą modyfikację wykonano za pomocą programu 'samtools sort'. Następnie do otrzymanych próbek dodane zostały grupy odczytów, tak aby możliwe było połączenie wyników z powtórzeń biologicznych

z tego samego genotypu na dalszych etapach analizach bioinformatycznych. W kolejnym kroku za pomocą narzędzia Picard zaznaczone zostały odczyty zdublikowane. Tak przygotowane pliki zostały następnie zindeksowane za pomocą polecenia 'samtools index'. W ten sposób uzyskano pliki, które mogły zostać wykorzystane w analizach wywoływania wariantu za pomocą pakietu GATK. W pierwszym etapie identyfikacji polimorfizmów SNP analizy przeprowadzone zostały dla każdego z plików osobno za pomocą algorytmu HaplotypeCaller, a następnie tak uzyskane wyniki połączone zostały w jeden plik vcf za pomocą narzędzia CombineGVCFs. Ostatnim krokiem było wyznaczenie polimorfizmów pomiędzy analizowanymi genotypami za pomocą algorytmu GenotypeGVCFs. W wyniku tych prac uzyskano surowy plik vcf zawierający potencjalne polimorfizmy pojedynczego nukleotydu. Łączna liczba uzyskanych wariantów wyniosła 44052. W celu eliminacji mało wiarygodnych wyników, otrzymane markery SNP poddane zostały filtracji. Zastosowano następujące parametry: $QD < 2.0$, $QUAL < 30.0$, $SOR > 3.0$, $FS > 60.0$, $MQ < 20.0$, $MQRankSum < -12.5$, $ReadPosRankSum < -8.0$, $DP > 500$, $DP < 4$. Z uzyskanych wariantów pozytywnie filtrację przeszło 39413. Pełną listę uzyskanych wiarygodnych (po filtracji) wariantów SNP zawiera plik Żyto.vcf.

Analiza transkryptomu dostarczyła ponad 2000 genów różnicujących (DEG), których ekspresja zmieniła się (zwiększyła się lub zmniejszyła) ponad dwukrotnie. Zakażenie patogenne prowadzi do aktywacji tzw. genów związanych z patogenezą (PR). Można je podzielić na co najmniej 17 rodzin. Bardziej znanymi białkami PR są białka PR1 (przeciwgrzybicze i przeciwgrzybicze), PR2 (β -1,3-glukanazy), PR3 (chitynazy), białka PR4 (przeciwgrzybicze), PR6 (inhibitory proteiny). Do związków sygnałowych odpowiedzialnych za indukcję białek PR należą kwas salicylowy, etylen, ksylanaza, systemina polipeptydowa, kwas jasmonowy i prawdopodobnie inne. Po przeanalizowaniu poziomów transkrypcji genów zaangażowanych w syntezę białek PR zauważono, że większość z nich była obniżona w porównaniu z nie zakażoną kontrolą. Chociaż białka PR są uważane za markery infekcji, zmniejszona ekspresja odpowiednich genów może być spowodowana wyborem punktu czasowego po infekcji, w którym wykonano transkryptom. Być może wybrany punkt czasowy był zbyt wczesny, aby zaobserwować wzrost ekspresji genów PR (Laluk i in. 2011).

Jedną z pierwszych interakcji molekularnych roślin i patogenów najprawdopodobniej obejmuje fosforylację białek, przepływu jonów, reaktywne formy tlenu (ROS) i inne formy przekazywania sygnałów (Shen i in 2017). Dlatego też transkryptom został przeszukany pod kątem genów związanych z powyższymi procesami. Okazało się, że znaleziono szereg takich genów, dla których odnotowano zarówno aktywację, jak i represję. Nie powinno to dziwić, a ponieważ do profilowania transkryptomu wykorzystano całe rośliny, a regulacja percepcji, transdukcji i realizacji sygnału patogenu jest złożoną siecią interakcji, specyficznych dla poszczególnych narządów, a nawet tkanek, aby dokładnie określić zmiany w ekspresji poszczególnych genów, w przyszłych badaniach należy zastanowić się nad wykonaniem analizy transkryptomu poszczególnych części zakażonej rośliny. Obserwowane zmiany w poziomach transkryptów kinaz receptorowych wskazują na dynamiczne procesy obejmujące szlaki sygnałowe w komórkach roślinnych. Wśród kinaz receptorowych, gen kinazy receptorowej związanej ze ścianą komórkową (WAK) (którego ekspresja była silnie (> 2000 razy) zwiększona po infekcji) jest szczególnie interesujący. WAK-i są białkami podobnymi do receptorów, zawierającymi cytoplazmatyczną kinazę serynowo-treoninową, domenę transbłonową i mniej konserwowany region, który jest związany ze ścianą komórkową. Wykazano, że WAK-i mogą odbierać fragmenty pektyn powstające podczas ekspozycji na patogen (Kohorn i Kohorn 2012) i przesyłać sygnał do komórki w celu zaindukowania odpowiedzi.

Profil transkrypcyjny rośliny zmienia się radykalnie wskutek zakażenia, dlatego muszą mu towarzyszyć zmiany poziomu ekspresji genów czynników transkrypcyjnych (TF). Rzeczywiście, zaobserwowaliśmy duże zmiany, zarówno wzrost, jak i spadek ekspresji genów TF. Ponownie zatem, możliwe jest, że wskazanie czynników transkrypcyjnych, które są konkretnie odpowiedzialne za daną reakcję rośliny

(np. produkcję ROS lub ich neutralizację) wymaga analizy różnych tkanek w różnych punktach czasowych.

Ponieważ jedną z pierwszych reakcji rośliny na infekcję jest produkcja ROS w tzw. „wybuchu oksydacyjnym”, roślina musi natychmiast zastosować środki zaradcze w postaci wygaszaczy wolnych rodników, aby chronić własne komórki przed uszkodzeniem. Oprócz maszynarii enzymatycznej roślina wykorzystuje drobnocząsteczkowe związki zdolne do inaktywacji wolnych rodników, takie jak polifenole czy terpenoidy. Niestety nie stwierdzono zmian w ekspresji genów kodujących enzymy wygaszające ROS, jak dysmutaza ponadtlenkowa czy katalaza, jednak zauważono znaczny wzrost poziomu ekspresji genów kodujących białko zawierające powtórzenia pentatricopeptydowe, które bierze udział w obronie roślin i obniża akumulację ROS (Laluk i in. 2011). Zaobserwowaliśmy również zwiększoną ekspresję genów zaangażowanych w syntezę polifenoli, takich jak reduktaza cynamoilo-CoA, 4-reduktaza dihydroflawonolu czy 3'-hydroksylaza flawonoidowa oraz syntezę terpenoidów, takich jak izomeraza beta-karotenu czy cyklaza tokoferolu.

Odpowiedź nadwrażliwa (HR) jest częścią odporności i obejmuje szybką śmierć komórek i związane z nią pogrubienie ściany komórkowej w miejscu próby zakażenia, co wydaje się ograniczać rozprzestrzenianie się patogenów i dostęp do składników odżywczych. Analizując transkryptom, znaleziono szereg genów zaangażowanych w regulację śmierci komórkowej, takich jak białka z rodziny F-box, których ekspresja była głównie podwyższona w zakażonych roślinach. Zwiększona była również ekspresja genów kodujących kaspazę. Sugeruje to, że wybrany punkt czasowy 24 godzin po zakażeniu można scharakteryzować reakcją PCD rośliny w miejscu zakażenia oraz przygotowaniem do produkcji białek PR.