

MIROSLAW TYRKA¹
GRZEGORZ FIC¹
MAGDALENA SZELIGA¹
MARCIN JAROMIN¹
PAWEŁ KRAJEWSKI²
PAWEŁ MILCZARSKI³
TADEUSZ DRZAZGA⁴
PRZEMYSŁAW MATYSIK⁵
RÓŻA MAZUR⁶
TERESA SIKORA⁷
EDWARD WITKOWSKI⁸
JUSTYNA BUCZKOWICZ¹
DOROTA TYRKA¹

¹ Politechnika Rzeszowska im. Ignacego Łukasiewicza

² Instytut Genetyki Roślin PAN w Poznaniu

³ Zachodniopomorski Uniwersytet Technologiczny w Szczecinie

⁴ Małopolska Hodowla Roślin

⁵ Hodowla Roślin Strzelce, Strzelce

⁶ Poznańska Hodowla Roślin, Tulce

⁷ DANKO Hodowla Roślin, Modzurów

⁸ Hodowla Roślin Smolice, Smolice

Kierownik Tematu: dr hab. Mirosław Tyrka Zakład Biotechnologii i Bioinformatyki, Politechnika Rzeszowska im. Ignacego Łukasiewicza, Al. Powstańców Warszawy 6, 35-959 Rzeszów, +48 17 8651927, e-mail: mtyrka@prz.edu.pl

Prace zostały wykonane w ramach badań podstawowych na rzecz postępu biologicznego w produkcji roślinnej na podstawie decyzji Ministra Rolnictwa i Rozwoju Wsi nr decyzji HOR.hn.802.12.2018, Zadanie 109.

Selekcja genomowa pszenicy ozimej

Genomic selection of winter wheat

Słowa kluczowe: genotypowanie przez sekwencjonowanie, pszenica zwyczajna, zmienność epigenetyczna

CEL BADAŃ

Celem badań było poszerzenie wiedzy na temat uwarunkowań genetycznych i epigenetycznych kształtujących zmiany w wysokości plonowania. Plon jest niezwykle złożoną cechą, a hipoteza badawcza zakłada, że część tej zmienności jest kształtowana

przez zmiany epigenetyczne. Realizacja przewidzianego cyklu badań powoli na określenie udziału i roli tych zmian w kształtowaniu plonu.

Celem zadania jest:

1. wdrożenie metody genotypowania przez sekwencjonowanie (GBS) wrażliwej na metylację,
2. wykorzystanie metody GBS do oznaczenia zmienności genetycznej w grupie 200 genotypów pszenicy zwyczajnej scharakteryzowanych pod względem plonu,
3. ocena zmienności fenotypowej w doświadczeniu porównawczym w wybranych lokalizacjach.

OPIS WYNIKÓW

Pierwszy cel badawczy, tj. wdrożenie i optymalizację metody GBS realizowano w etapach: a) wybór endonuleaz do fragmentacji DNA, b) wybór dodatkowego enzymu wrażliwego na zmiany metylacyjne w celu redukcji złożoności bibliotek, c) projektowanie adaptorów do konstrukcji bibliotek, d) konstrukcja i ocena ilościowa bibliotek, e) sekwencjonowanie bibliotek f) obróbka bioinformatyczna uzyskiwanych danych pozwalająca na ocenę pokrycia i powtarzalności wyników.

Dla 43 enzymów restrykcyjnych rozpoznających różne motywy 6-nuleotydowe określono liczbę fragmentów o wielkości w zakresie od 75 do 150 nukleotydów. Przewidziano % fragmentów unikalnych i ich rozkład w genomie podzielonym na partycje o wielkości 300 kpz. W efekcie przeprowadzonych symulacji komputerowych stwierdzono, że liczba fragmentów o wielkości w zakresie 75–150 pz generowana przez poszczególne enzymy wynosi średnio 147 tysięcy i w zależności od zastosowanego enzymu restrykcyjnego waha się w szerokim zakresie od 14 do 483 tys. fragmentów. Udział fragmentów unikalnych dla zestawu enzymów restrykcyjnych wynosił od 63,5% do 92,5%. Średnie pokrycie genomu przez uzyskiwane fragmenty wyniosło 73% i wahało się od 99,4% do 21,6%. (tab. 1). Na podstawie danych o liczbie fragmentów unikalnych i pokryciu do wstępnych testów wybrano 3 enzymy restrykcyjne.

W drugim etapie symulacji przewidziano ilości i rozkład fragmentów po dodatkowym trawieniu redukcyjnym enzymem rozpoznającym motywy 4pz. Na podstawie symulacji komputerowej wybrano 5 kombinacji enzymów (oznaczonych M1-M5) i zaprojektowano adaptory dla 12 losowych barkodów odpowiadających odmianom pszenicy ozimej wybranym losowo z listy odmian roślin uprawnych (KWS Ozon, Mewa, Muszelka, Natula, Nutka, Opcja, Ostka Strzelecka, Ostroga, Owacja, Pokusa, Sailor i Satyna).

W celu weryfikacji symulacji, zaadaptowano metodykę Poland i in. (2012) do wytypowanych kombinacji enzymów restrykcyjnych. Stosowano adaptory do pierwszego enzymu. Fragmenty o docelowej wielkości 75–150 nukleotydów, po uwzględnieniu ligowanych adaptorów, wycinano z żelu, oczyszczano i charakteryzowano na fluorymetrze. Próbki łączono w ilościach równomolowych i sekwencjonowano na aparacie Miseq zgodnie z instrukcją.

Tabela 1

Charakterystyka polimorfizmu w bibliotekach M1-5. W analizie rozkładu podano liczbę genotypów u których wykryto zmienność w tym samym locus

Biblioteka	M1		M2		M3		M4		M5	
Liczba wszystkich sekwencji	89396		97884		3046		406		6106	
Liczba sekwencji unikalnych	55550		56926		977		252		2083	
Średnie pokrycie	1,39	1,27	1,06	1,08	2,29	2,38	1,38	1,34	2,21	2,16
Rozkład										
12	2	0	0	0	5	3	0	0	3	3
11	0	0	4	3	5	5	0	0	12	10
10	3	0	0	0	16	15	0	0	30	13
9	1	1	3	3	31	14	0	0	45	29
8	13	1	3	2	41	15	0	0	66	29
7	40	9	3	3	47	14	3	2	113	42
6	147	21	13	11	75	28	2	1	179	52
5	489	111	40	33	137	43	1	1	261	66
4	1758	526	167	122	194	51	6	4	391	124
3	5240	2249	627	434	392	102	28	13	650	150
2	16316	8294	4325	2877	571	158	49	28	1065	278
1	65387	44338	92699	53438	1532	529	317	203	3291	1287

Uzyskiwane dane z sekwencjonowania przypisywano do poszczególnych próbek na podstawie adresów, następnie czyszczono z sekwencji o błędach przekraczających 1 nukleotyd na 1000. Oczyszczone dane posłużyły do wydzielenia dla każdej z bibliotek oddzielnie puli fragmentów unikalnych (unigenów). Wprowadzenie tego etapu do analizy miało na celu przyspieszenie docelowych obliczeń dla dużej liczby próbek. Najwyższa liczba sekwencji unikalnych była dla bibliotek M1 i M2. Sekwencje unikalne w bibliotekach M3-M5 stanowiły od 324 tysięcy do 575 tysięcy. Następnie wykonano dopasowanie unikalnych sekwencji do genomu referencyjnego pszenicy i uzyskano pliki (SAM, Sequence Alignment Map) zawierające informacje o pozycji fizycznej markera na chromosomach pszenicy. W kolejnym etapie w sekwencjach uzyskanych z poszczególnych bibliotek i genotypów identyfikowano mutacje punktowe (SNP). Dane o występowaniu SNP w poszczególnych badanych odmianach w bibliotekach M1-5 zebrano i przeanalizowano pod kątem zmienności (tab. 1). W bibliotekach M1 i M2 stwierdzono występowanie odpowiednio 89 tys. i 98 tys. fragmentów co przekracza możliwości rzetelnej oceny na aparacie Miseq. W przypadku bibliotek M3-M5 uzyskiwane liczby markerów były akceptowalne i pokrycie wystarczające do identyfikacji SNP można uzyskać na posiadanym aparacie. Analizując rozkład fragmentów w badanych odmianach do dalszych analiz na 200 genotypach wybrano kombinację M5.

W ramach realizacji drugiego etapu prac kombinację enzymów M5 wykorzystano do uzyskania i sekwencjonowania 3 bibliotek (po 67 genotypów każda). Uzyskane dane są w trakcie obróbki bioinformatycznej i będą wykorzystane do identyfikacji markerów epigenetycznych po skompletowaniu wyników oceny fenotypowej.

W ramach realizacji trzeciego celu badawczego założono doświadczenia w 5 lokalizacjach (Strzelce, Modzurów, Nagradowice, Smolice, Kobierzyce). Określono wysokość plonu ziarna o wilgotności 15% dla 200 genotypów badanych w 4 seriach w odniesieniu do wzorców w układzie bloków niekompletnych. We wszystkich seriach stwierdzono

istotne efekty związane z genotypem badanych linii. Analiza wariancji z uwzględnieniem lat i miejscowości wskazuje na celowość uzupełnienia analiz o zmienne towarzyszące uwzględniające warunki pogodowe panujące w trakcie wegetacji.

WNIOSKI

W efekcie przeprowadzonych badań wdrożono skalowalną metodykę skanowania całego genomu pszenicy. Wydajność opracowanego systemu docelowo powinna pozwalać na uzyskiwanie do 6000 markerów na genotyp. Proponowany system ma wydajność pośrednią pomiędzy markerami DArT bazującymi na hybrydyzacji (około 1000 markerów) a systemem DArTseq (około 20000 markerów). Uzyskano dane o plonowaniu 200 linii pszenicy ozimej po uzupełnieniu o dane z doświadczenia założonego w 2018 roku będą wartościowe bez względu na rozwój technologii markerowych i mogą służyć obliczaniu efektów selekcyjnych oraz identyfikacji markerów epigenetycznych związanych z plonowaniem.

LITERATURA

Poland J. A., Brown P. J., Sorrells M. E., Jannink J.-L. 2012. Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. *PLoS ONE* 7 (2): e32253.