

Identyfikacja wybranych genów związanych z typem wzrostu roślin ogórka (*Cucumis sativus* L.)

Okres realizacji zadania: 2022 r.

Miejsce realizacji zadania:

Katedra Genetyki Hodowli i Biotechnologii Roślin
Szkola Główna Gospodarstwa Wiejskiego w Warszawie

Kierownik zadania:

prof. dr hab. Grzegorz Bartoszewski

e-mail: grzegorz_bartoszewski@sggw.edu.pl

Główni wykonawcy:

dr inż. Renata Słomnicka, dr inż. Karolina Kaźmińska, dr inż. Aleksandra Korzeniewska, dr inż. Helena Olczak-Woltman

Cel zadania

Identyfikacja wybranych genów odpowiedzialnych za typ wzrostu roślin ogórka oraz opracowanie markerów molekularnych dla zidentyfikowanych genów, potencjalnie przydatnych w hodowli twórczej tego gatunku

Cele zadania w 2022 roku

1. Ocena fenotypowa populacji F_2 ogórka i uzyskanie nasion F_3 dla wybranych populacji
2. Analiza danych resekwencjonowania i identyfikacja polimorfizmów w genomie ogórka oraz sekwencjonowanie transkryptomów
3. Izolacja DNA z roślin ogórka i testowanie markerów genetycznych na potrzeby mapowania

Cele zadania zaplanowane na 2022 rok zostały w pełni osiągnięte

Materiały i metody

Materiał roślinny

- Dziesięć linii ogórka o zmienionym typie wzrostu i linie kontrolne o normalnym typie wzrostu L500 (B10DH) i L501 (Gy14).

Metody

- **Doświadczenia z roślinami w szklarni, tunelu foliowym i w polu:** ocena fenotypowa typu wzrostu roślin (1 cecha) dla segregujących populacji (12 populacji) reprezentujących linie o zmienionym typie wzrostu, pobranie tkanek do izolacji DNA, uzyskanie nasion przez samozapylenie (minimum 60 roślin F_2 dla trzech populacji).
- **Analiza genetyczna:** w oparciu o dane fenotypowe dla P_1 , P_2 , F_1 , F_2 oraz BC_1 i BC_2 , analiza statystyczna - test χ^2 (12 populacji).
- **Analizy bioinformatyczne:** identyfikacja polimorfizmów strukturalnych SV i CNV w oparciu o dane resekwencjonowania, wstępna analiza danych transkryptomicznych RNA-seq (2 analizy).
- **Izolacja RNA z roślin ogórka:** zestaw Plant RNeasy, ocena jakości preparatów z wykorzystaniem elektroforezy i spektrofotometru (6 linii w 3 powtórzeniach biologicznych).
- **Sekwencjonowanie transkryptomów RNA-seq:** z wykorzystaniem technologii Illumina (18 transkryptomów).
- **Izolacja DNA z roślin ogórka:** zestaw NucleoSpin Plant II, ocena preparatów z wykorzystaniem elektroforezy i spektrofotometru (3 populacje po 80 roślin).
- **Testowanie starterów i opracowanie markerów DNA:** zaprojektowanie starterów PCR (30 par starterów), analiza produktów amplifikacji, wytypowanie markerów dla genów kandydackich, testowanie markerów genetycznych (3 markery) na segregujących populacjach F_2 .

Wyniki: ocena fenotypowa segregujących populacji ogórka

- Typ wzrostu roślin dla populacji uzyskanych w wyniku krzyżowania linii L504, L505 i L511 o zmienionym typie wzrostu z liniami o niezmienionym typie wzrostu L500 i L501 wykonano **w szklarni**. Ocena tych materiałów była możliwa dla roślin w fazie 2-4 liści.
- Ocenę typu wzrostu dla populacji uzyskanych z krzyżowania linii L506, L507 i L512 z linią L500 wykonano **w tunelach foliowych**, zaś dla populacji uzyskanych z krzyżowania tych linii z linią L501 **w polu**. Ocena tych materiałów była możliwa w fazie 15-25 międzywęźli.
- Dla populacji F_2 oceniono od 99 do 365 roślin w zależności od dostępności nasion, dla segregujących krzyżowań wstecznych BC_2 oceniono minimum 96 roślin, dla P_1 i P_2 oraz F_1 i BC_1 minimum 10 roślin.
- Obserwowano większe zróżnicowanie pojedynków F_2 o zmienionym typie wzrostu w porównaniu z roślinami linii rodzicielskiej.
- Łącznie wykonano ocenę typu wzrostu roślin dla 12 populacji F_2 wraz z towarzyszącymi im odpowiednimi liniami rodzicielskimi, F_1 i BC_1 oraz BC_2 .

Doświadczenie w szklarni – populacja F_2 L500 x L511



Doświadczenie w tunelu – populacja F_2 L500 x L506



Wyniki: analiza genetyczna

- W oparciu o uzyskane dane fenotypowe dla pokolenia F_2 , krzyżowań wstecznych BC_1 i BC_2 oraz P_1 , P_2 i F_1 i wykonano analizę genetyczną dla sześciu linii o zmienionym typie wzrostu.
- Dla wszystkich badanych linii potwierdzono segregację 3:1 w F_2 i 1:1 w BC_2 co oznacza, że za zmieniony typ wzrostu tych linii odpowiadają pojedyncze geny recesywne.
- Dla populacji F_2 L501 x L511 uzyskano stosunek segregacji inny niż 3:1 co może wynikać z tła genetycznego linii matecznej L501 (linia żeńska Gy14).

Wyniki: uzyskanie nasion F_3

- Dla wytypowanych roślin F_2 o niezmienionym typie wzrostu dla populacji uzyskanych w wyniku krzyżowania L500 z liniami L504, L505 i L511 wykonano samozapylenia i otrzymano nasiona.
- Uzyskano wystarczające ilości nasion F_3 na potrzeby dalszych badań.

Podsumowanie wyników analizy genetycznej

Krzyżowanie	Segregacja typu wzrostu F_2	Wartość $Chi^2_{emp.}$
L500 x L504	3:1	0.06
L500 x L505	3:1	1.47
L500 x L506	3:1	0.75
L500 x L507	3:1	2.19
L500 x L511	3:1	0.60
L500 x L512	3:1	1.77
L501 x L504	3:1	1.54
L501 x L505	3:1	0.45
L501 x L506	3:1	2.20
L501 x L507	3:1	0.20
L501 x L511	inny	6.39*
L501 x L512	3:1	2.66

* odrzucono hipotezę o segregacji 3:1

Zestawienie liczebności pojedynków F_2 i liczby nasion

Populacja	Liczba pojedynków F_2	Liczba nasion F_3
L500 x L504	60	85-1175
L500 x L505	66	30-621
L500 x L511	76	87-988

Wyniki: analiza bioinformatyczna polimorfizmów strukturalnych

- Analiza danych resekwencjonowania genomów linii o zmienionym typie wzrostu pozwoliła na identyfikację wariantów strukturalnych w odniesieniu do genomu referencyjnego B10 v3.
- Zidentyfikowano znacznie więcej wariantów strukturalnych zidentyfikowano z wykorzystaniem programu CNVnator (średnio 13 336 CNV dla danej linii) w porównaniu z wariantami SV identyfikowanymi za pomocą programu BreakDancer (średnio 3 653 SV dla danej linii).
- Tylko niewielki udział wariantów strukturalnych był związany z genami i ich rejonami regulatorowymi (6% CNV i 24% SV).
- Dla linii L508 zidentyfikowano około 20% więcej wariantów strukturalnych, co potwierdza, że linia ta na poziomie genetycznym znacząco różni się od pozostałych.
- Zidentyfikowane polimorfizmy strukturalne okazały się mało przydatne w projektowaniu markerów molekularnych dla wybranych genów kandydackich – zaprojektowano dwa markery na bazie polimorfizmu SV jednak nie potwierdzono tych polimorfizmów eksperymentalnie.

Podsumowanie wyników identyfikacji wariantów strukturalnych dla danych resekwencjonowania genomów 10 linii ogórka o zmienionym typie wzrostu w odniesieniu do genomu referencyjnego B10 v.3

Polimorfizm strukturalny SV (program BreakDancer)	Średnia dla linii
Łączna liczba zmian	3653
Ogółem w genach	876
Rejony promotorowe (1 kb)	78
Eksony	407
Introny	327
Rejony terminatorowe (1 kb)	64
Polimorfizm strukturalny CNV (program CNVnator)	Średnia
Łączna liczba zmian	13336
Ogółem w genach	825
Rejony promotorowe (1 kb)	168
Eksony	449
Introny	36
Rejony terminatorowe (1 kb)	172

Wyniki: izolacja RNA i analiza RNA-seq

- Badania wykonano dla pięciu linii o zmienionym typie wzrostu i linii kontrolnej L500 (B10DH) w trzech powtórzeniach biologicznych.
- Uzyskano dobrej jakości RNA wolne od zanieczyszczeń DNA, które wykorzystano do sekwencjonowania transkryptomów RNA-seq.
- Uzyskano odczyty sekwencyjne dobrej jakości (Q20>97%), które zmapowano na genomie referencyjnym 9930 v3.
- Średni odsetek zmapowanych odczytów wyniósł 93,9% co świadczy o wysokiej jakości odczytów sekwencyjnych.

Wyniki: wstępna analiza bioinformatyczna danych RNA-seq

- Potwierdzono zgodność wyników RNA-seq dla trzech powtórzeń biologicznych dla każdej z badanych linii.
- Wstępnie zidentyfikowano geny ulegające zmienionej ekspresji w odniesieniu do linii kontrolnej L500. W zależności od linii było to od 483 do 1293 genów.
- Najwięcej genów ulegających zmienionej ekspresji zidentyfikowano dla linii L505 zaś najmniej dla L502 i L511.

Podsumowanie wyników oceny jakości wyników sekwencjonowania transkryptomów RNA-seq.

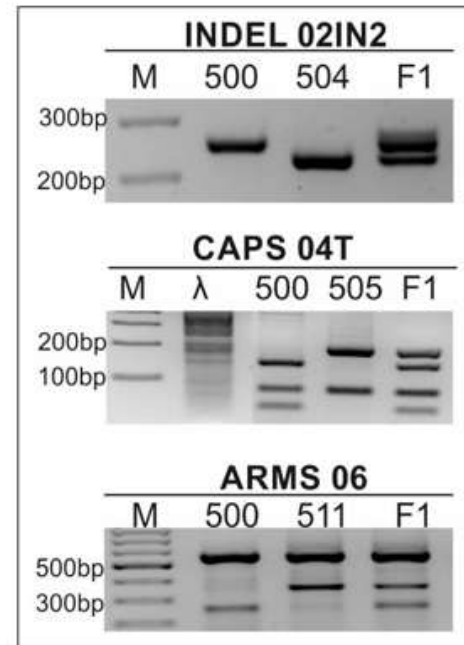
Odczyty mapowano na genomie referencyjnym 9930v3.

Linia	Powtórzenie biologiczne	Wartość Q20	Odsetek unikalnie zmapowanych odczytów
L500	1	97,3	93,8%
	2	97,3	93,8%
	3	97,5	93,9%
L502	1	97,3	93,9%
	2	97,6	93,8%
	3	97,4	93,7%
L505	1	97,5	93,7%
	2	97,3	93,7%
	3	97,5	93,7%
L507	1	97,5	94,1%
	2	97,5	94,0%
	3	97,4	93,9%
L511	1	97,5	94,1%
	2	97,4	93,8%
	3	97,4	94,0%
L512	1	97,5	93,9%
	2	97,4	93,7%
	3	97,4	93,8%
Średnia		97,4	93,9%

Wyniki: izolacja DNA, testowanie starterów i opracowanie markerów genetycznych

- Pobrano młode liście z roślin F_2 segregujących pod względem typu wzrostu populacji $L500 \times L504$, $L500 \times L505$ i $L500 \times L511$ i wyizolowano z nich DNA.
- Zaprojektowano 30 par starterów w oparciu o zidentyfikowane bioinformatycznie polimorfizmy SNP, INDEL i SV, które testowano na liniach rodzicielskich i F_1 trzech wytypowanych populacji.
- Opracowano dziewięć markerów dla wybranych genów kandydackich były to dwa markery typu CAPS, jeden typu ARMS i sześć typu INDEL.
- Trzy wybrane markery przetestowano na populacjach mapujących.
- Dla dwóch markerów uzyskano wysoką zgodność wyników fenotypowania typu wzrostu roślin z genotypowaniem, w przypadku trzeciego markera takiej zgodności nie było.

Profile amplifikacyjne markerów INDEL 02IN2, CAPS 04T i ARMS 06 uzyskane dla linii rodzicielskich i F_1 . Markery te wykorzystano do genotypowania populacji mapujących. M – marker wielkości DNA, linie rodzicielskie 500, 504, 505, 511



Wnioski

Ocena fenotypowa populacji mapujących F_2 ogórka i uzyskanie nasion F_3 dla wybranych populacji

1. Ocena typu wzrostu roślin ogórka w segregujących populacjach jest możliwa w fazie 2-4 liści dla linii L504, L505 i L511. Dla linii L506, L507 i L512 ocenę tę można wykonać na roślinach w fazie 15-25 międzywęźli.
2. Fenotypowanie segregujących populacji i analiza genetyczna potwierdziły, że zmieniony typ wzrostu roślin w badanych liniach ogórka jest warunkowany przez pojedyncze geny recesywne.
3. Dla trzech populacji uzyskano wystarczające ilości nasion na potrzeby oceny fenotypowej rodzin F_3 , która jest niezbędna do mapowania genów warunkujących zmieniony typ wzrostu roślin.

Analiza danych resekwencjonowania i identyfikacja polimorfizmów w genomie ogórka oraz sekwencjonowanie transkryptomów

1. Analizy bioinformatyczne danych resekwencjonowania pozwoliły na wskazanie rearanżacji strukturalnych. Dla linii L508 wykryto około 20% różnic więcej niż dla pozostałych linii, co potwierdza, że linia ta znacznie różni się od pozostałych.
2. Dzięki uzyskaniu dobrej jakości danych transkryptomicznych RNA-seq możliwe było zidentyfikowanie genów ulegających zróżnicowanej ekspresji dla pięciu linii o zmienionym typie wzrostu w odniesieniu do linii kontrolnej.
3. Analizy bioinformatyczne danych transkryptomicznych pokazały, że badane linie różnią się pod względem liczby genów ulegających zmienionej ekspresji - najwięcej takich genów odnotowano dla linii L505.

Izolacja DNA z roślin ogórka i testowanie markerów genetycznych na potrzeby mapowania

1. Na bazie zidentyfikowanych bioinformatycznie polimorfizmów zaprojektowano startery i opracowano markery do testowania na populacjach mapujących.
2. Spośród opracowanych markerów, trzy okazały się przydatne do genotypowania segregujących populacji F_2 . Wstępne wyniki pokazały, że wskazano obiecujące geny kandydackie dla genów zmienionego typu wzrostu w liniach L505 i L511.

Prezentacja wyników badań

Postery

Każmińska K, Słomnicka R, Mużacz S, Korzeniewska A, Bartoszewski G (2022) Phenotypic evaluation of cucumber lines characterized by dwarf plant architecture. 6th Polish Congress of Genetics, Kraków 2022, Book of abstracts, PR-4.11

Słomnicka R, Każmińska K, Bartoszewski G (2022) Assessment of genetic variation in cucumber lines characterized by dwarf phenotype. 6th Polish Congress of Genetics, Kraków 2022, Book of abstracts, PR-1.6