

Tytuł zadania:

Doskonalenie mapy genetycznej łubinu wąskolistnego i poszukiwanie markerów sprzężonych z cechami użytkowymi ze szczególnym uwzględnieniem zawartości białka i alkaloidów

Okres realizacji:

Rok 2021

Kierownik zadania:

Instytut Genetyki Roślin Polskiej Akademii Nauk
dr hab. Michał Książkiewicz, mksi@igr.poznan.pl

Zespół wykonawców:

Instytut Genetyki Roślin Polskiej Akademii Nauk:
dr Sandra Rychel-Bielska
dr Wojciech Bielski
mgr inż. Jolanta Belter (pracownik techniczny)

Uniwersytet Przyrodniczy we Wrocławiu
dr hab. inż., prof. UP Renata Galek
dr Bartosz Kozak
Mariusz Skrzyński (pracownik techniczny)
mgr inż. Barbara Wiśniewska (pracownik techniczny)

Poznańska Hodowla Roślin Spółka z o.o.
mgr Paweł Barzyk (pracownik techniczny)

Cele projektu:

Temat badawczy nr 1

Uzyskanie liczby nasion linii kolekcyjnych łąbinu wąskolistnego oraz linii wsobnych populacji mapującej Emir × LAE-1 wystarczającej do przeprowadzenia doświadczeń polowych zaplanowanych w kolejnych latach projektu

Temat badawczy nr 2

Opracowanie zestawu loci polimorfizmu pojedynczego nukleotydu do genotypowania przez sekwencjonowanie dla linii rodzicielskich populacji mapujących Emir × LAE-1 i Graf × Mutant cienkościenny

Temat badawczy nr 3

Uzyskanie zestawu markerów polimorfizmu pojedynczego nukleotydu metodą genotypowania przez sekwencjonowanie linii wsobnych populacji mapujących Emir × LAE-1 i Graf × Mutant cienkościenny

Wszystkie cele zostały zrealizowane

Materiał:

- populacja mapująca łubinu wąskolistnego Emir × LAE-1 (133 linie wsobne), która segreguje w zakresie wysokości roślin, liczby pędów bocznych, typu wzrostu, pęknięcia strąków, plonu, kwitnienia i dojrzewania
- populacja mapująca łubinu wąskolistnego Graf × Mutant cienkościenny (97 linii wsobnych), która segreguje w zakresie grubości okrywy nasiennej
- 300 linii z kolekcji łubinu wąskolistnego: populacje dzikie (115), linie prymitywne (4), odmiany (88), linie hodowlane (87), mutanty (6)

Metody:

- wysiew nasion (siewnik), zbiór strąków (ręczny/automatyczny), ocena koloru nasion i kwiatów, obliczenie masy tysiąca nasion i współczynnika rozmnażania
- izolacja DNA przy użyciu Genomic Mini AX Plant Spin (AABiotechnology) oraz Maxwell® RSC PureFood GMO and Authentication Kit (Promega)
- ocena stężenia i czystości DNA metodą spektrofotometryczną - Nanodrop 2000 (Thermo Fisher Scientific) lub na drodze elektroforezy - BioAnalyze 2100 (Agilent)
- amplifikacja PCR, trawienie enzymami restrykcyjnymi, elektroforeza
- sekwencjonowanie genomowe NovaSeq6000 (Illumina): fragmenty o długości ok. 350 bp, sparowane końce, odczyty 2×150 bp, min. 10-krotne pokrycie genomu
- kontrola jakości surowych odczytów sekwencjonowania (Trimmomatic, fastQC)
- identyfikacja loci polimorficznych (potoki analityczne GATK v4.1 i freebayes)
- filtracja loci za pomocą algorytmu uczenia maszynowego z wykorzystaniem bibliotek języka Python: keras, Tensorflow2, pandas, numpy oraz scikit-learn
- celowane genotypowanie przez sekwencjonowanie (SeqSNP, Future Synthesis)

Rozmnożenie linii kolekcyjnych łubinu wąskolistnego (temat 1):

- W 2021 roku wysiano **300 linii** kolekcyjnych. Uzyskano średnio **254 nasiona** na linię (po odliczeniu nasion na kolejne rozmnożenie), plon był silnie skorelowany ze stopniem udomowienia: najwyższy w odmianach, najniższy w populacjach dzikich.
- Dla **150 linii** uzyskano plon nasion na doświadczenia polowe w pełnym zakresie (> 200), zaś dla kolejnych **72 linii** zwiększono pulę nasion o co najmniej 100 szt.

Tab. 1. Plon nasion w liniach kolekcyjnych łubinu wąskolistnego uzyskany w 2021 r.

Stopień udomowienia	Liczba linii	Średnia liczba nasion	Odchylenie st.
Populacje dzikie	115	194	176
Linie hodowlane	87	266	242
Odmiany	88	323	232
Linie prymitywne	4	253	142
Mutanty	6	202	124

Rozmnożenie populacji mapującej Graf × Mutant cienkościenny (temat 1):

- W 2021 roku wysiano **129 linii** wsobnych populacji mapującej Graf × Mutant cienkościenny. Uzyskano średnio po **570 nasion** na linię (min. 52, maks. 1520).
- Dla **116 linii** uzyskano plon nasion umożliwiający prowadzenie doświadczeń polowych, zaś dla kolejnych **12 linii** uzyskano co najmniej 100 nasion. Nasiona będzie można też uzupełnić z posiadanych zasobów (pokolenie RILs₇).
- Współczynnik rozmnożenia nasion wyniósł od 1,0 do 57,2 (**średnio 14,4**), zaś masa tysiąca nasion mieściła się przedziale od 83,3 do 184,2 g (**średnio 131,3 g**).
- Zaobserwowano **osiem** wzorców ubarwienia nasion i **cztery** barwy kwiatów.

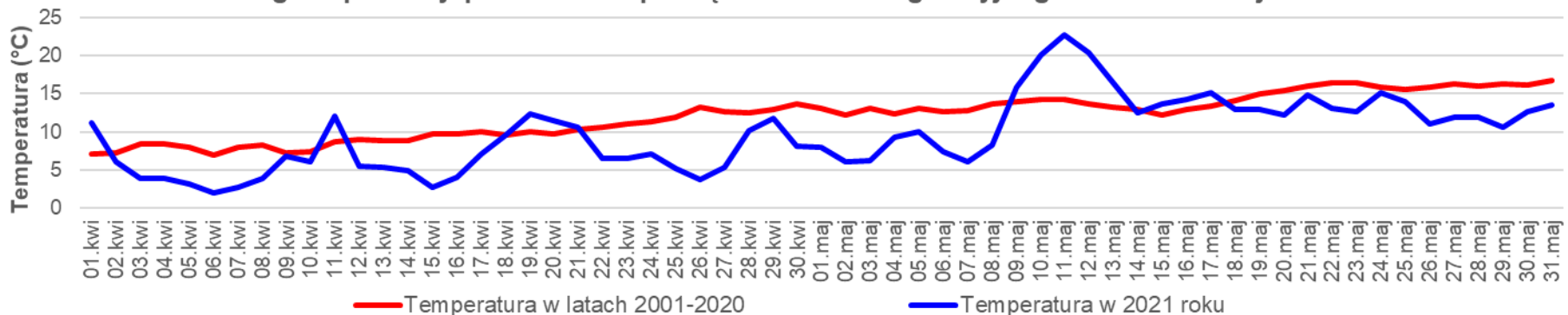
Rozmnożenie populacji mapującej Emir×LAE-1 (temat 1):

- W 2021 roku wysiano **133 linie** wsobne populacji mapującej Emir×LAE-1. Uzyskano średnio po **2585 nasion** na linię (min. 456, maks. 6008).
- Dla **wszystkich linii** uzyskano plon nasion do doświadczeń polowych.
- Współczynnik rozmnożenia nasion wyniósł od 1,9 do 28,7 (**średnio 11,5**), zaś masa tysiąca nasion mieściła się przedziale od 89,8 do 191,1 g (**średnio 140,1 g**).
- Zaobserwowano **trzy** wzorce ubarwienia nasion i **trzy** barwy kwiatów.

Wpływ przebiegu pogody na uzyskane wyniki (temat 1):

- Przyczyn **niskiego plonu** linii kolekcyjnych należy upatrywać w niekorzystnych warunkach atmosferycznych w pierwszych tygodniach po wysianiu nasion, a mianowicie **niskiej temperaturze** przy jednocześnie **wysokiej sumie opadów**.
- Średnia temperatura powietrza w okresie pierwszych 38 dni od wysiania nasion była **o 3,7°C niższa** od średniej z ostatniego dwudziestolecia dla tego samego okresu, największe odchylenie dobowe od normy dziennej wyniosło -9,5°C, zaś łącznie **przez dwa tygodnie odchylenie ujemne przekraczało 5°C**:

Przebieg temperatury powietrza na początku sezonu wegetacyjnego na tle średniej z 2001-2020



Sekwencjonowanie linii rodzicielskich Graf i Mutant cienkościenny (temat 2):

- Uzyskano **70,2 mln** surowych odczytów o śr. długości 150 pz dla linii Graf oraz **69,2 mln** odczytów o śr. długości 150 pz dla linii Mutant cienkościenny
- Po usunięciu adaptorów oraz odczytów niskiej jakości, łączna długość sekwencji dla linii Graf wyniosła **10,6 mld pz** zaś dla linii Mutant cienkościenny **10,5 mld pz**, co przekłada się na ok **16,5 krotne pokrycie genomu**.

Identyfikacja loci polimorfizmu pojedynczego nukleotydu (SNP) (temat 2):

- Liczba markerów **SNP** uzyskanych dla linii rodzicielskich populacji **Graf × Mutant cienkościenny** za pomocą potoku GATK wyniosła **779 789**, zaś w przypadku potoku freebayes – **753 970**, z czego **422 508** było wspólnych dla obu metod.
- Dla form rodzicielskich populacji **Emir × LAE-1** uzyskano **574 152** loci **SNP** w potoku GATK oraz **780 082** w potoku freebayes, z czego **396 128** było wspólnych.
- Uzyskany zestaw loci SNP przyrównano do zestawu markerów **MACE** dla kolekcji nasiennej, opublikowanego w 2020 r. Wyodrębniono **1101** loci wspólnych z MACE dla populacji Emir×LAE-1 i **928** loci dla populacji Graf×Mutant cienkościenny.
- Uzyskany zestaw loci SNP przyrównano do zestawu markerów **DArT-seq** polimorficznego w populacji mapującej Emir × LAE-1, opublikowanego w 2019 r. Zidentyfikowano **434** loci SNP wspólne z DArT-seq dla populacji Emir × LAE-1, oraz **135** markerów DArT-seq dla populacji Graf × Mutant cienkościenny.
- Finalnie do dalszych analiz wybrano **94 036 markery SNP** akceptowalnie równo rozmieszczone w genomie, preferując loci w genach, MACE i DArT-seq.

Izolacja DNA z linii wsobnych obu populacji mapujących (temat 3):

- Wykonano izolację DNA **140 linii** wsobnych z populacji mapującej Emir × LAE-1 oraz **131 linii** wsobnych z populacji mapującej Graf × Mutant cienkościenny. Zestaw ten zawierał duplikaty kilku linii wsobnych wykazujących segregację koloru nasion, w celu uzyskania informacji o podobieństwie sekwencji obu wariantów.
- Dla populacji Emir × LAE-1 uzyskano **stężenie DNA** w zakresie od 204 do 859 ng/μl (**średnia 472 ng/μl**), zaś dla populacji Graf × Mutant cienkościenny uzyskano stężenie DNA w zakresie od 241 do 1237 ng/μl (**średnia 509 ng/μl**).
- **Wszystkie** wyniki były **powyżej** minimalnego progu wymaganego dla celowanego genotypowania przez sekwencjonowanie (SeqSNP), czyli **50 ng/μl**.
- **Wszystkie** próby **przeszły** test jakości i **zostały** poddane genotypowaniu SeqSNP.

Analiza polimorfizmu opracowanych wcześniej markerów zasocjowanych z cechami użytkowymi (temat 3):

- Badania wykazały, że zestaw loci polimorficznych w obrębie badanych populacji mapujących obejmuje **6 markerów MACE** zasocjowanych w kolekcji nasiennej z **plonem** oraz **3 markery MACE** zasocjowane z **terminem dojrzałości**.

Genotypowanie przez sekwencjonowanie SeqSNP (temat 3):

- W ostatecznie opracowanym zestawie sond SeqSNP jest **20 tys. loci SNP**, w tym **12 706** loci polimorficznych między Emir i LAE-1, **11 600** loci polimorficznych między Graf i Mutant cienkościenny oraz **4306** loci wspólnych dla obydwu populacji. Charakterystykę tych loci przedstawiono w **Tabeli 2**.

Tab. 2. Loci SNP analizowane w ramach celowanego genotypowania przez sekwencjonowanie (**SeqSNP**) **140 linii** z populacji mapującej **Emir × LAE-1 (E×L)** oraz **131 linii** z populacji mapującej **Graf × Mutant cienkościenny (G×AU)**.

Chromosom	Markery w genach	Markery MACE	Markery DArT-seq	Markery E×L	Markery G×AU	Markery wspólne
NLL-01	612	102	19	737	855	222
NLL-02	712	139	37	939	1126	472
NLL-03	686	61	52	999	724	308
NLL-04	403	85	26	631	623	270
NLL-05	610	37	23	877	851	417
NLL-06	808	82	34	1015	1205	426
NLL-07	347	30	9	471	543	201
NLL-08	338	13	16	391	474	53
NLL-09	451	81	35	549	685	174
NLL-10	286	74	11	534	363	144
NLL-11	602	74	17	783	834	225
NLL-12	298	30	17	463	351	154
NLL-13	405	78	35	718	545	282
NLL-14	222	26	9	425	345	137
NLL-15	348	27	30	602	298	117
NLL-16	344	35	20	601	464	230
NLL-17	208	1	4	389	204	63
NLL-18	301	26	22	548	175	64
NLL-19	292	35	21	523	298	126
NLL-20	373	37	6	511	637	221
Suma	8646	1073	443	12706	11600	4306

Wnioski:

1. W kolekcji nasiennej łąbinu wąskolistnego występuje znaczne zróżnicowanie potencjału plonowania, ujawniające się w szczególności w niekorzystnych warunkach atmosferycznych na początku sezonu wegetacyjnego.
2. Liczba nasion oraz masa tysiąca nasion w populacjach mapujących łąbinu wąskolistnego Graf × Mutant cienkościenny i Emir×LAE-1 są cechami ilościowymi, warunkowanymi przez co najmniej kilka genów (świadczy o tym rozkład ciągły).
3. Kolor kwiatów w populacjach mapujących Graf × Mutant cienkościenny i Emir ×LAE-1 warunkuje najprawdopodobniej jeden gen główny (barwa biała/niebieska).
4. Zastosowana metodyka umożliwiła opracowanie zestawu markerów SNP do genotypowania SeqSNP obu populacji mapujących łąbinu wąskolistnego.
5. Zidentyfikowanie zestawu loci SNP współdzielonych z opublikowanymi markerami MACE i DArT-seq umożliwi utworzenie zintegrowanej mapy genetycznej.
6. Dla populacji Emir × LAE-1 uzyskano zestaw loci równomiernie rozmieszczony w genomie, umożliwiający konstrukcję mapy genetycznej o wysokiej rozdzielczości.
7. Dla populacji Graf × Mutant cienkościenny uzyskano zestaw loci dla każdego chromosomu, ale jednocześnie zaobserwowano występowanie kilku obszarów o niskiej gęstości markerów, których uzupełnienie zaplanowano w roku 2022 (zgodnie z kontynuacją tematu nr 3 ujętą w harmonogramie wieloletnim).
8. Występowanie polimorfizmu dla markerów MACE zasocjowanych z plonem i terminem dojrzałości pozwoli na porównanie zidentyfikowanych nowych loci QTL z opublikowanymi wynikami dla kolekcji nasiennej łąbinu wąskolistnego.

Doniesienia konferencyjne:

1. Kozak B., Książkiewicz M., Rychel-Bielska S., Bielski W., Galek R. „*Identyfikacja loci polimorficznych i przygotowanie zestawu markerów do wysokoprzepustowego genotypowania populacji mapującej łubinu wąskolistnego Graf x Mutant cienkościenny*”. Konferencja „Genetyka Aplikacyjna Roślin – wyzwania XXI wieku, 22-24 września 2021, Warszawa. Streszczenia, str. 78.
2. Strony w tegorocznym sprawozdaniu, które zawierają informacje prezentowane na tej konferencji: 12-15, 21.

Realizacja mierników:

Temat	Miernik – l.p.	Wartość podana w opisie zadania	Wartość zrealizowana	Stopień realizacji
1	1.1	300	300	1,00
	1.2	133	133	1,00
	1.3	129	129	1,00
2	2.1	2	2	1,00
	2.2	4	4	1,00
3	3.1	2	2	1,00
	3.2	8-20 tys.	20 tys.	1,00