

Современные технологии в селекции растений

Продолжающий уровень
13 марта – 24 марта 2023 года

Презентация к Лабораторному практикуму «Создание собственных молекулярных маркеров на основе результатов GWAS и QTL картирования»

23 – 24 марта
2022 года

Мартынова Елена, Степан Болдырев, Баик Алина

Создание маркера. С чего начать?

Маркер - это не только **полиморфизм** в ДНК, ассоциированный с признаком интереса, но и **система**, позволяющая **визуализировать** искомый полиморфизм и **дифференцировать** между **различными** его **аллелями**

Перед тем как приступить к разработке маркера необходимо ответить на следующие вопросы.

Вопросы для размышления

- С каким типом полиморфизма Вы работаете - полиморфизм в размере локуса ДНК (например, SSR или InDel), полиморфизм единичного нуклеотида (SNP). Разные полиморфизмы «удобнее» детектировать разными методами
- Кодоминантный или доминантный маркер
- Количество растений для генотипирования - много или мало - Вам нужна высокопроизводительная система
- Доступный/удобный способ визуализации результатов
- Возможности автоматизации получения результата
- Доступность реагентов и оборудования

Ответив на данные вопросы, важно провести

- Оценка стоимости создания маркера
- Оценка стоимости генотипирования с использованием маркера

Создание маркера. Какие бывают системы генотипирования.



Низкопроизводительные/неавтоматизированные системы

- SNP, InDel - большое разнообразие аллель-специфичных системы на основе ПЦР и электрофоретической визуализации (чаще доминантный, реже кодоминантный)
- SNP CAPS маркеры - ПЦР+рестрикция (кодоминантный)
- SSR - системы на основе ПЦР + гель-электрофорез

ПЦР + секвенирование по Сэнгеру (кодоминантный)



Фрагментный анализ (кодоминантный) - низкопроизводительная автоматизированная система для SSR, InDEI и иногда SNP



Системы генотипирования с использованием визуализации в режиме реального времени - низко-и среднепроизводительные, частично автоматизированные системы

- Стандартный аллель-специфичный ПЦР (доминантный в принципе, может быть кодоминантным)
- TaqMan зонды
- KASPar системы
- HRM системы

И другие системы для детекции SNP, иногда SSR

Высокопроизводительное генотипирование

- Bead chip arrays
 - AmpliSeq assays
 - GBS (генотипирование с помощью секвенирования)
- И другие



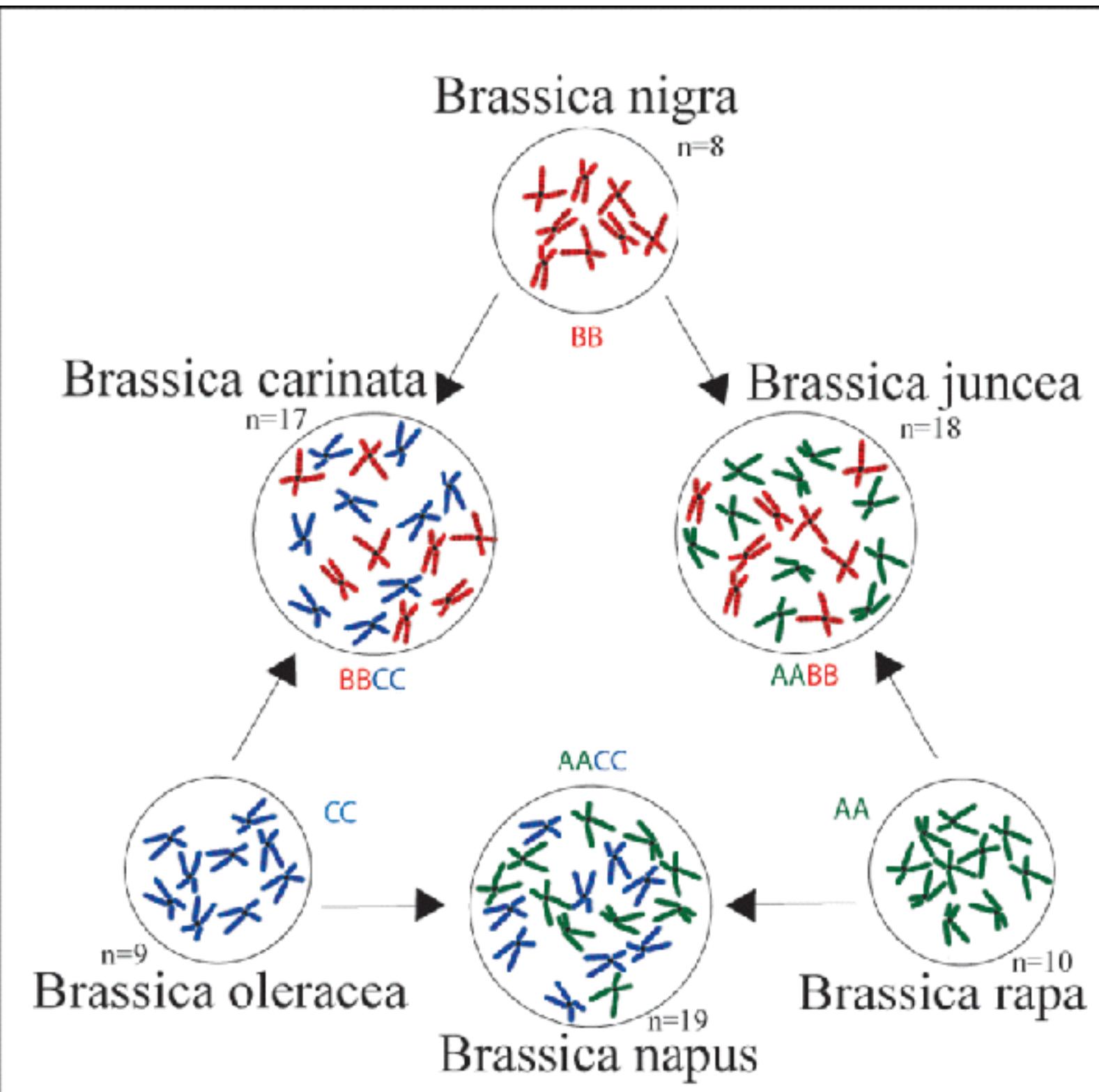
DArTs



Создание маркера. Этапы на примере QTL-картирования.



Геном рапса (*B. napus*)



- Рапс является аллополиплоидом (*Brassica napus* L.; AnAnCnCn, $2n = 38$)
- Произошел в результате гибридизации между двумя близкородственными видами семейства Brassicaceae, *Brassica rapa* (AA, $2n = 20$) and *Brassica oleracea* (CC, $2n = 18$) не ранее, чем 10,000 лет назад
- Наиболее близкий родственник *Arabidopsis thaliana*
- Размер генома - примерно 1,132 Гб
- Первое полногеномное секвенирование рапса и сборка генома de novo было проведено в 2014 году для европейского сорта озимого рапса 'Darmor-bzh' (Chalhoub et al., 2014) с использованием подходов секвенирования нового поколения технологии 454 Roche и Illumina и досеквенирования методом Сэнгера.
- Повторное секвенирование было проведено в 2017 году для сорта полу-озимого рапса (semi-winter) «ZS11» с использованием платформы Illumina HiSeq2000 (Sun et al., 2017).

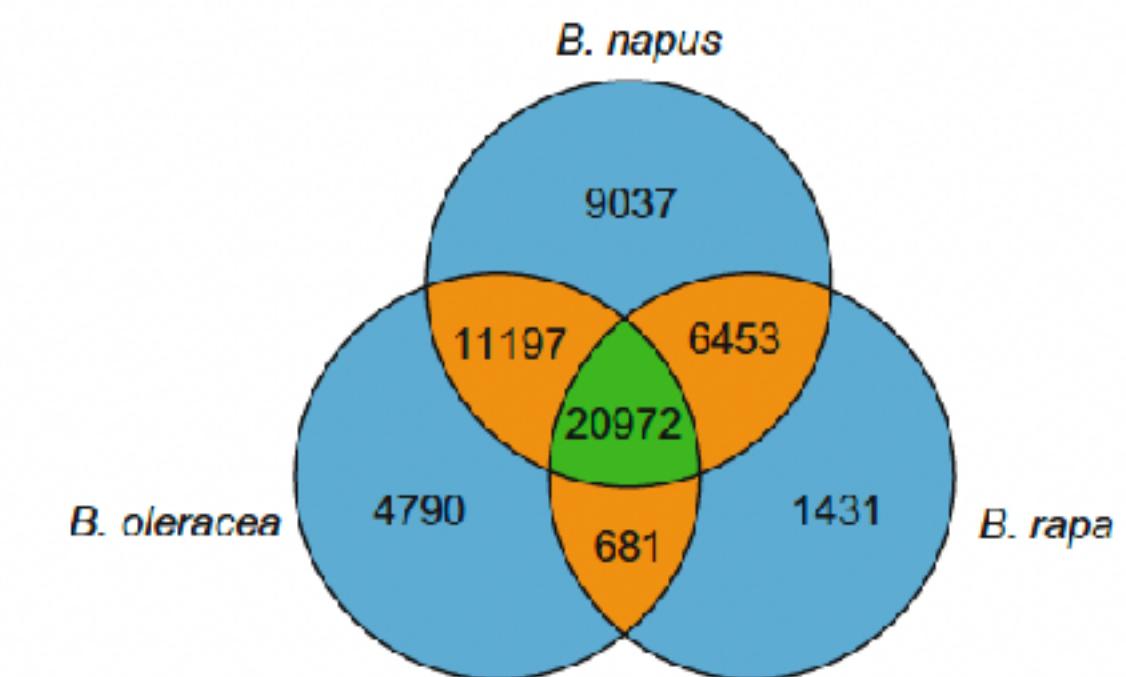


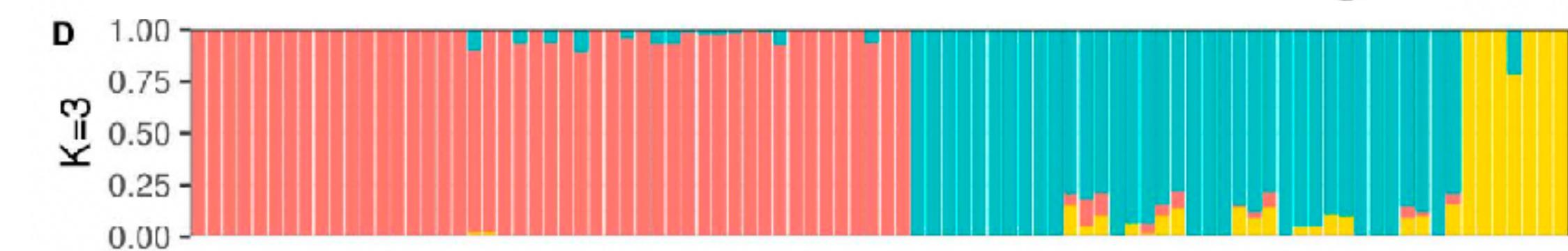
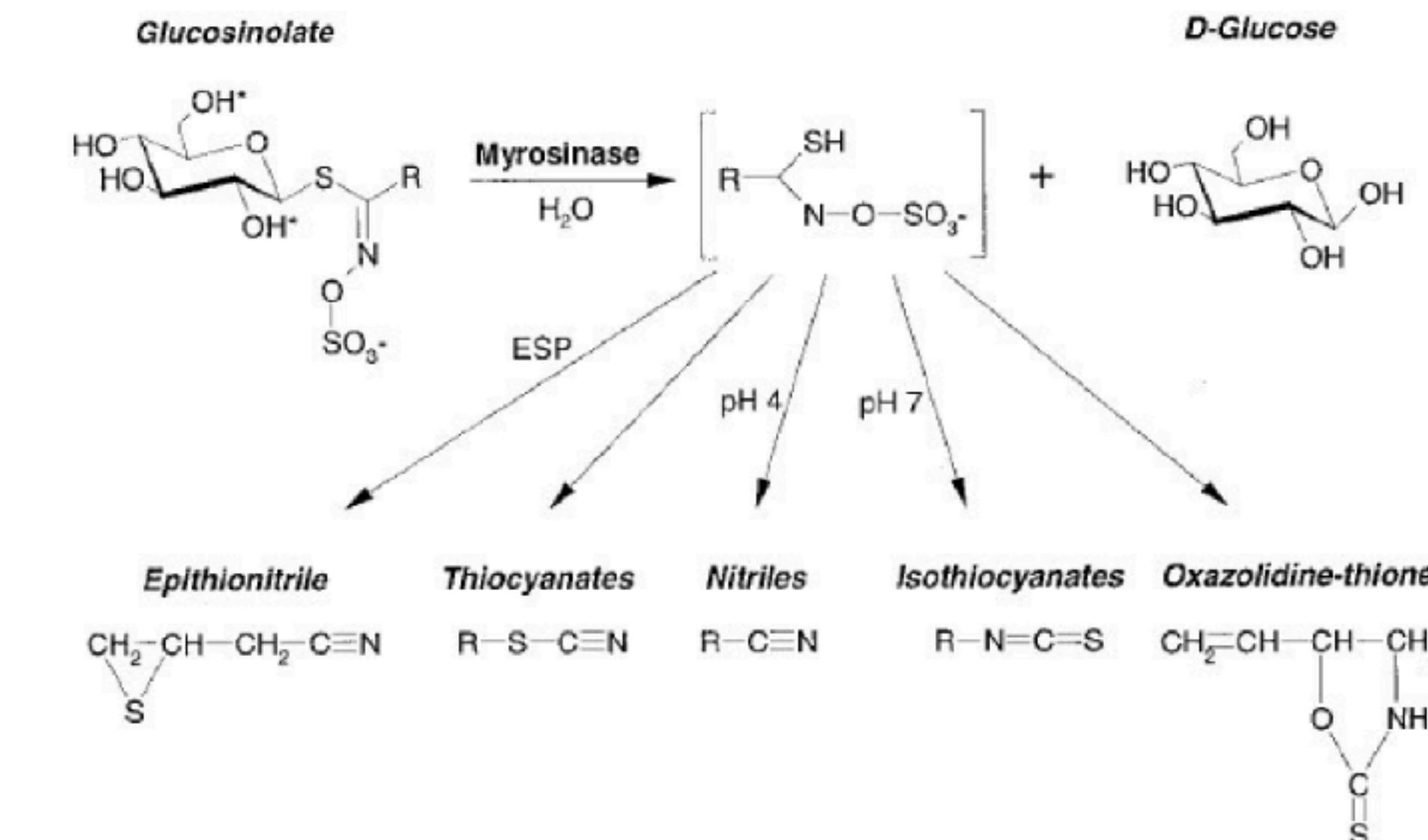
Fig. 4.4 Gene families clustering among three-related species

Первый вариант: количественный признак

Содержание глюкозинолатов в масле семян рапса - данные GWAS

Глюкозинолаты представляют собой группу вторичных метаболитов, синтезируемых представителями семейства Brassicaceae. Эти метаболиты, с одной стороны, способствуют защите растений от патогенов и, с другой стороны, обуславливают горький вкус растений семейства Brassicaceae. Было также показано, что глюкозинолаты негативно влияют на функцию почек и печени у высших животных.

Современные программы селекции рапса направлены на снижение присутствия глюкозинолатных компонентов в рапсовом масле и жмыхе.

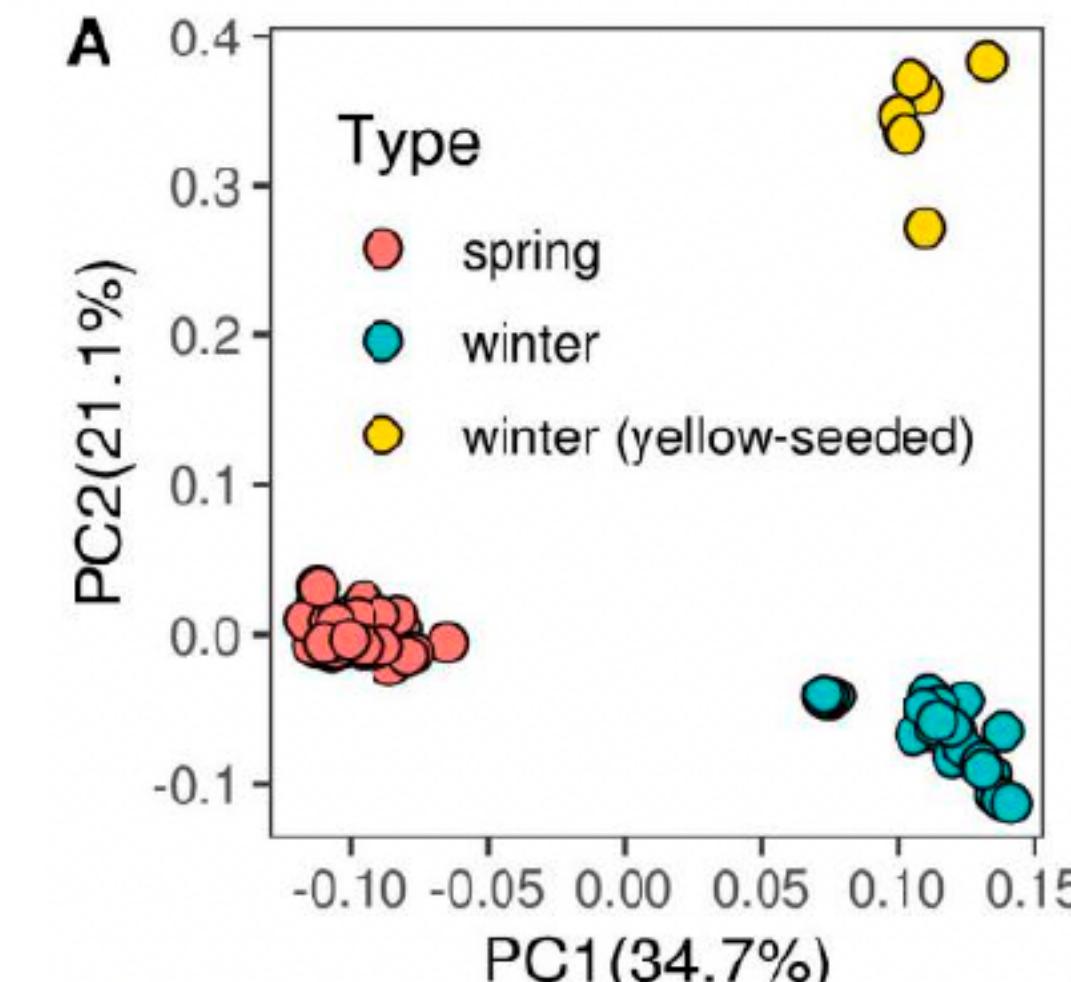


В работе использовали 47 инбредных линий ярового и 43 инбредные линии озимого рапса.

Библиотеки для NGS секвенирования получали методом GBS с использованием рестрикционных эндонуклеаз *PstI* and *MspI*.

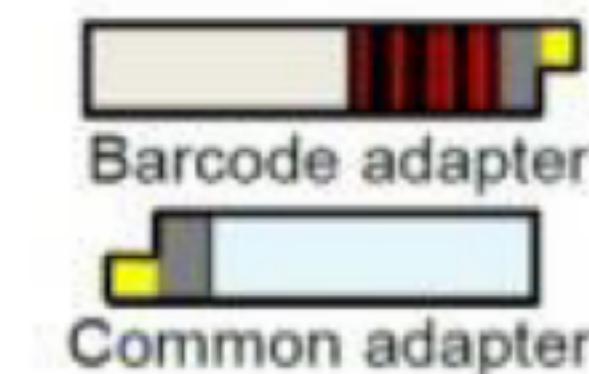
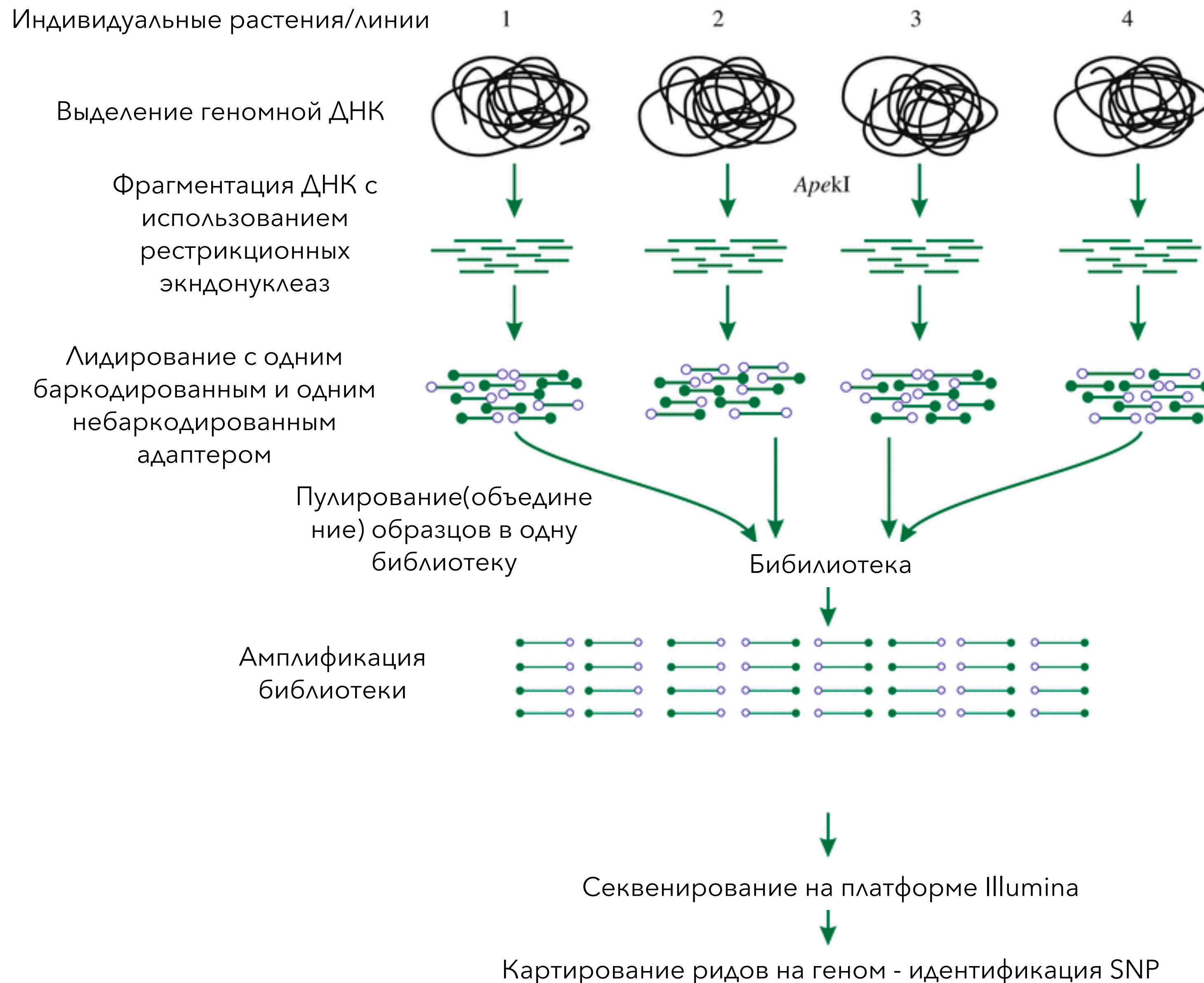
GWAS проводился с использованием программного обеспечения TASSEL software version 5.2 на основе смешанных линейных моделей. Предварительно проводился анализ популяционной структуры исследуемых линий методом PCA (PLINK software version 1.9) и программы ADMIXTURE.

Gubaev et al., 2020



GBS - генотипирование секвенированием

Индивидуальные растения/линии

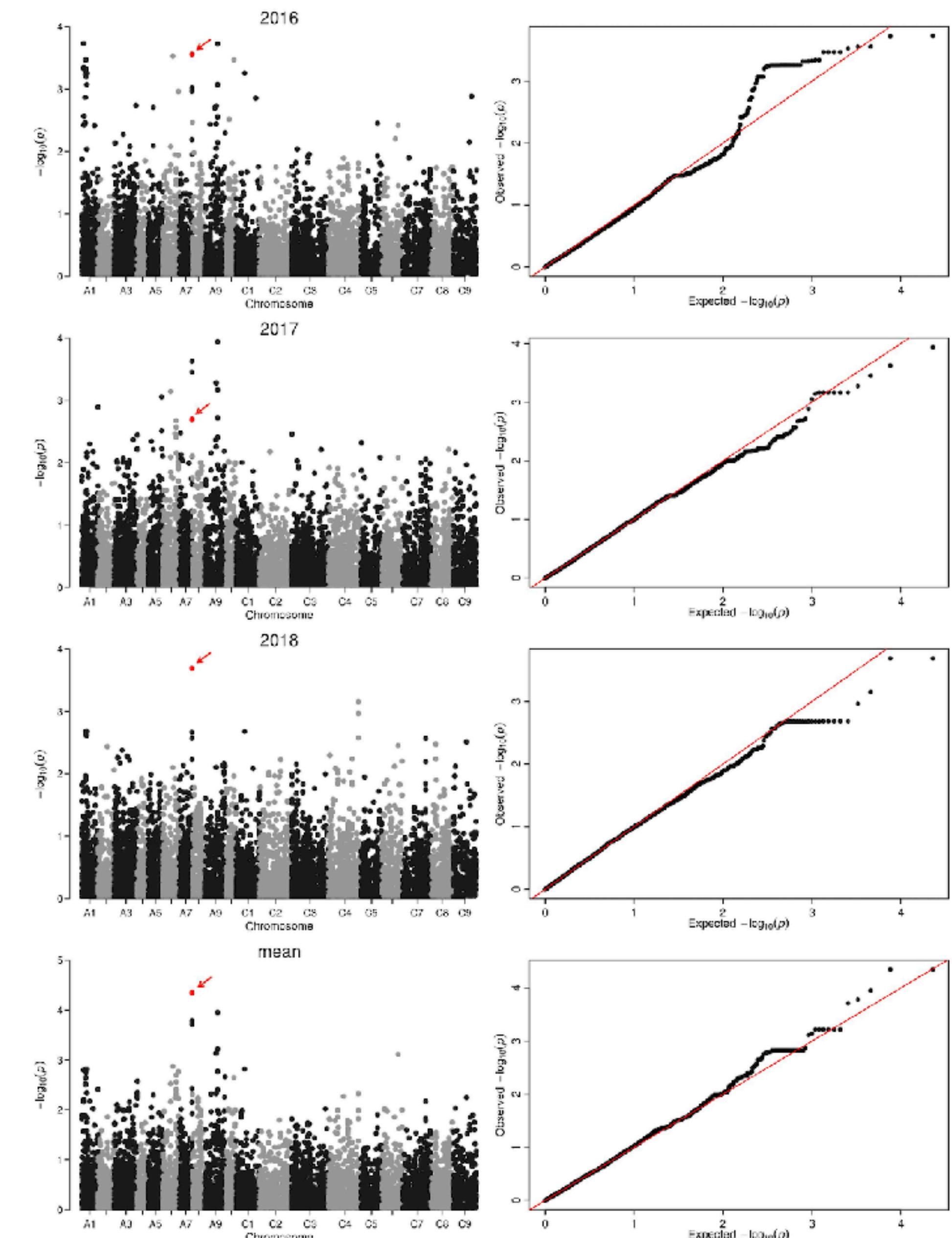
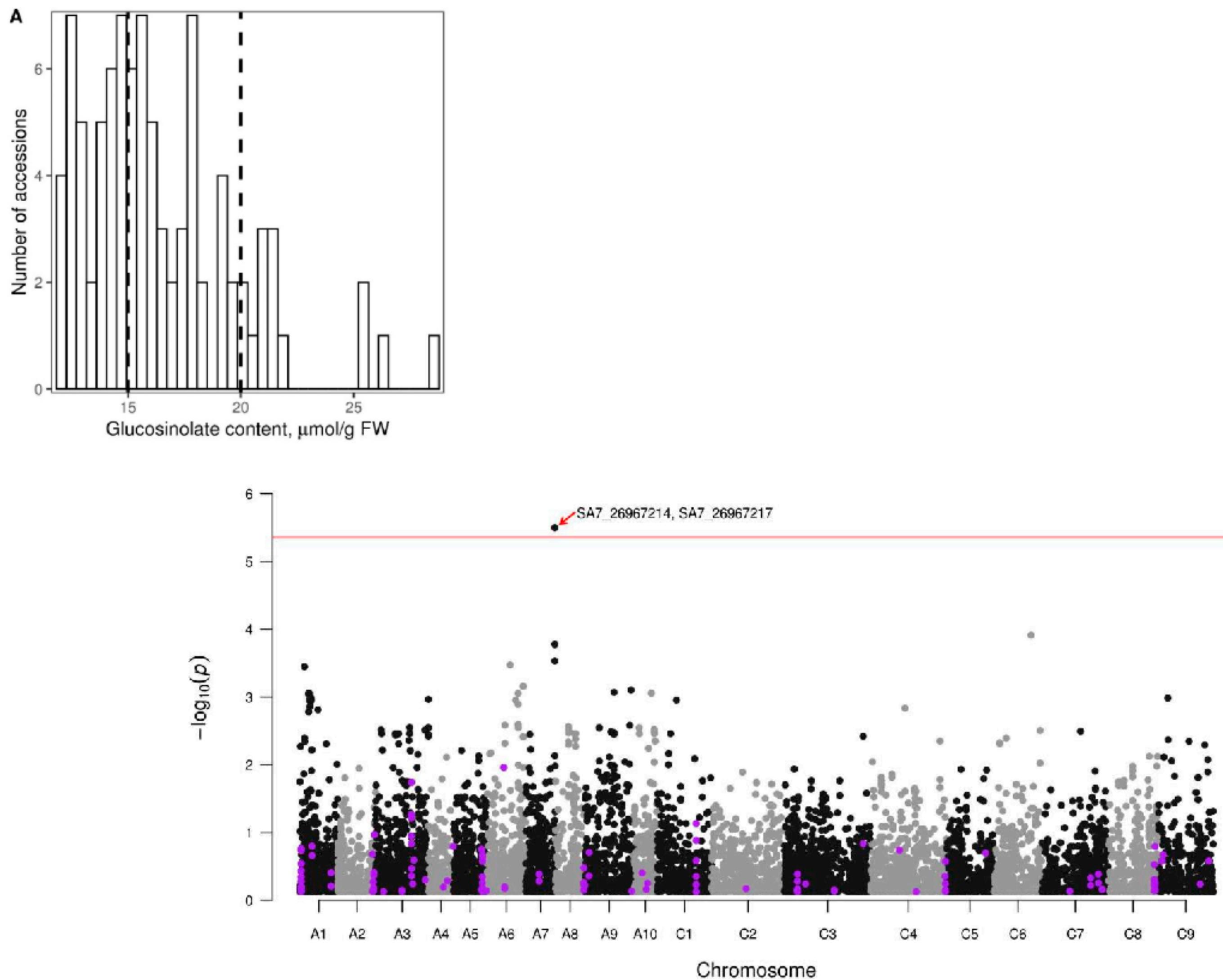


Оригинальная статья:

Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, et al. (2011) A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species. PLoS ONE 6(5): e19379. <https://doi.org/10.1371/journal.pone.0019379>

Первый вариант: количественный признак

Содержание глюкозинолатов в масле семян
рапса - данные GWAS



Первый вариант: количественный признак

Содержание глюкозинолатов в масле семян рапса - данные GWAS

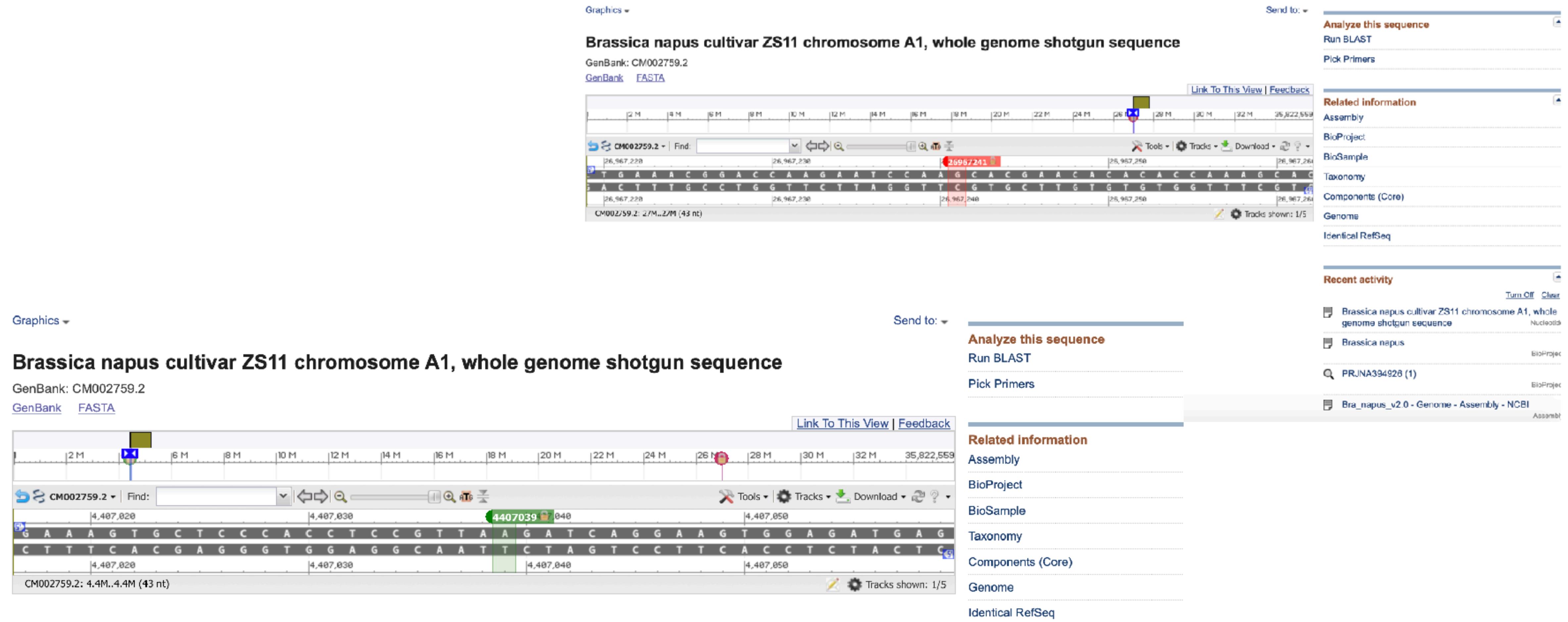
Таблица SNP с наиболее высоким достоверным уровнем
ассоциации с признаком интереса - содержания глюкозинолатов

SNP	Chromosome	Position	MarkerR2 (2016)	P-value (2016)	MarkerR2 (2017)	P-value (2017)	MarkerR2 (2018)	P-value (2018)	P-value (Irwin-Hall)
SA1_4407039	1	4407039	0,11347	0,01364	0,1176	0,01279	0,23472	0,00018605	0,0003542070588012
SA6_21541176	6	21541176	0,1273	0,00796	0,11252	0,01762	0,20244	0,00029697	0,0003348087881904
SA7_26967211	7	26967211	0,09292	0,01688	0,18138	0,00035337	0,14105	0,00108	0,0001676897603784
SA7_26967214	7	26967214	0,20484	0,0002031	0,13835	0,00203	0,17205	0,00027665	3,15E-06
SA7_26967217	7	26967217	0,20484	0,0002031	0,13835	0,00203	0,17205	0,00027665	3,15E-06
SA7_26967241	7	26967241	0,08556	0,02298	0,1914	0,0002377	0,1435	0,00096431	0,0002923848038200
SC6_35450938	16	35450938	0,10741	0,00354	0,08359	0,00832	0,08979	0,00382	0,0001229312

Эффект маркера на признак - не более 20%.

Genome Browsers and Genome Assemblies

[NCBI.NIH.GOV](https://www.ncbi.nlm.nih.gov), Ensemble Plants, Phytozome, <https://sunflowergenome.org/>, etc



Rapeseed Genome assembly

Bra_napus_v2.0 Cultivar SZ11 «SZ11»_assembly

[HTTPS://WWW.NCBI.NLM.NIH.GOV/ASSEMBLY/GCA_000686985.2](https://www.ncbi.nlm.nih.gov/assembly/GCA_000686985.2)

Bra_napus_v2.0

Organism name: [Brassica napus \(rape\)](#)

Infraspecific name: Cultivar: ZS11

BioSample: [SAMN02742820](#)

BioProject: [PRJNA237736](#)

Submitter: BGI-Shenzhen

Date: 2017/09/15

Assembly level: Chromosome

Genome representation: full

RefSeq category: representative genome

GenBank assembly accession: GCA_000686985.2 (latest)

RefSeq assembly accession: GCF_000686985.2 (latest)

RefSeq assembly and GenBank assembly identical: no ([hide details](#))

- Only in RefSeq: chromosomes MT, Pltd, linear plasmid (in non-nuclear assembly-unit)
- Data displayed for RefSeq version

WGS Project: [JMKK02](#)

Assembly method: SOAPdenovo v. 2; Celera v. 8.3rc2

Expected final version: yes

Genome coverage: 200x

Sequencing technology: Illumina HiSeq 2000

IDs: 1200221 [UID] [GenBank](#) ▾

Send to: ▾

Due to the large size of this record, sequence and annotated features are not shown. Use the "Customize view" panel to change the display.

Brassica napus cultivar ZS11 chromosome C6, whole genome shotgun sequence

GenBank: CM002774.2

[FASTA](#) [Graphics](#)

Go to: ▾

LOCUS CM002774 45505715 bp DNA linear CON 15-SEP-2017
DEFINITION Brassica napus cultivar ZS11 chromosome C6, whole genome shotgun sequence.

Assembly Unit: Primary Assembly (GCF_000686985.2)

Molecule name	GenBank sequence		RefSeq sequence	Unlocalized sequences count
Chromosome A1	CM002759.2	=	NC_027757.2	0
Chromosome A2	CM002760.2	=	NC_027758.2	0
Chromosome A3	CM002761.2	=	NC_027759.2	0
Chromosome A4	CM002762.2	=	NC_027760.2	0
Chromosome A5	CM002763.2	=	NC_027761.2	0
Chromosome A6	CM002764.2	=	NC_027762.2	0
Chromosome A7	CM002765.2	=	NC_027763.2	0
Chromosome A8	CM002766.2	=	NC_027764.2	0
Chromosome A9	CM002767.2	=	NC_027765.2	0
Chromosome A10	CM002768.2	=	NC_027766.2	0
Chromosome C1	CM002769.2	=	NC_027767.2	0
Chromosome C2	CM002770.2	=	NC_027768.2	0
Chromosome C3	CM002771.2	=	NC_027769.2	0
Chromosome C4	CM002772.2	=	NC_027770.2	0
Chromosome C5	CM002773.2	=	NC_027771.2	0
Chromosome C6	CM002774.2	=	NC_027772.2	0
Chromosome C7	CM002775.2	=	NC_027773.2	0
Chromosome C8	CM002776.2	=	NC_027774.2	0
Chromosome C9	CM002777.2	=	NC_027775.2	0
unplaced	n/a	n/a	n/a	1,449

Rapeseed Genome assembly

Новая сборка - Da-Ae

Объект секвенирования - проростки рапса сорта Da-Ae (США)

Метод секвенирования - PacBio Sequel; Illumina HiSeq;

Глубина секвенирования - 100x;

Авторы сборки - ученые из университета UC Davis, John Davis (South Carolina)

Адрес проекта в базе данных NCBI

https://www.ncbi.nlm.nih.gov/assembly/GCF_020379485.1/

BRAD V3.0; <http://brassicadb.cn/>

Global statistics

Total sequence length	1,001,886,053
Total ungapped length	1,001,800,256
Gaps between scaffolds	0
Number of scaffolds	3,167
Scaffold N50	48,209,797
Scaffold L50	9
Number of contigs	4,007
Contig N50	1,580,367
Contig L50	177
Total number of chromosomes and plasmids	22
Number of component sequences (WGS or clone)	3,167

Первый вариант: количественный признак

Задание 1.

Основываясь на данных Таблицы SNP с наиболее высоким достоверным уровнем ассоциации с признаком найти в геноме рапса (сборка Bra_napus_v2.0 ZS11) все указанные в таблице SNP.

Для каждого SNP выделить из генома участок длиной 251 нт, содержащий SNP ровно посередине (250 нт слева и 250 нт справа).

Для каждого фрагмента создать отдельный файл в простом текстовом редакторе типа TextPad или WordPad и сохранить в формате .fasta. SNP выделить с помощью квадратных скобок (например, [A]).

Создание маркера - Дизайн праймеров

Праймеры-синтетические олигонуклеотиды, состоящие из 16-30 оснований. Они комплементарны участкам ДНК, между которыми находится последовательность-мишень.

Праймер(англ. primer) является обязательным компонентом, необходимым для работы ДНК-полимеразы: к его **3'-ОН концу** фермент присоединяет нуклеотиды, комплементарные матрице.

Праймер к 5'-концу гена называют **прямым праймером (forward, For)**, к 3'-концу гена -**обратным праймером(reverse, Rev)**.

В базах данных нуклеотидных последовательностей обычно приводится только одна цепь ДНК -**значающая**, та, которая транскрибируется в виде мРНК. По ней подбирают **прямой праймер**, т.е. тот праймер, с которого будет начинаться амплификация искомого фрагмента ДНК.

Обратный праймер подбирают **для комплементарной цепи**, но также в направлении **5'→3'**, данный праймер будет фланкировать искомый ДНК-фрагмент.

Правила подбора праймеров - очень важный момент! Плохо подобранные праймеры могут непоправимо «испортить» вашу ПЦР реакцию

- Размер праймера должен находиться в пределах 16-25(30) нуклеотидов.
- CG-состав праймера должен находиться в пределах 40-60%. Нуклеотиды G и C должны быть равномерно распределены по всей длине праймера;
- Желательно чтобы на 3'-конце праймера находился нуклеотид G или C, так как это увеличивает эффективность ПЦР;
- При этом следует избегать наличия 3 и более оснований G или C на 3'-конце праймера, поскольку это может приводить к ошибочному отжигу праймера на ГЦ-богатых участках ДНК;
- Существует мнение о том, что на 5'-конце праймера желательно преимущественное присутствие А или Т нуклеотидов;
- Температура плавления (Tm) данной пары праймеров должна быть сходной и, желательно, находиться в пределах 55-65C;
- Разница в температуре отжига двух праймеров в паре должна составлять не более 5C;
- Праймеры не должны быть само-и взаимно-комplementарными, то есть формировать стабильных дуплексов сами на себя или друг с другом. В противном случае, эффективность ПЦР может значительно снижаться из-за истощения запаса праймеров в реакции;
- Нуклеотиды 3'-конца праймера в случае стандартной ПЦР должны быть строго комплементарны матрице(замены возможны на 5' конце длинных праймеров).
- Участки отжига должны быть уникальными, т.е. праймеры не должны иметь длинных комплементарных участков вне специфических мест отжига;

Не забывайте об антипараллельности цепей ДНК при подборе обратного праймера. Если он будет иметь нужный состав, но другую полярность, ПЦР-продукт получен не будет. Такой праймер (параллельный) не будет отжигаться на матрицу.

Создание маркера - Дизайн праймеров

**Основной путь подбора праймеров:
поиск оптимальной пары Primer3plus**

+ проверка физических свойств праймера (температуры плавления/отжига) OligoCalc

+ Проверка возможной самокомплементарности и формирования димеров

+ Проверка специфичности PrimerBlast

+ GenomeCheck (дополнительная проверка уникальности и специфичности на референсном геноме)

<http://biotools.nubic.northwestern.edu/OligoCalc.html>

<https://primer3plus.com/cgi-bin/dev/primer3plus.cgi>

<https://www.genscript.com/tools/pcr-primers-designer>

<https://eu.idtdna.com/pages/tools/oligoanalyzer?returnurl=%2Fcalc%2Fanalyzer> (OligoAnalyzer Tool)

Создание маркера - Дизайн праймеров

Расчёт температуры отжига праймера можно осуществить самостоятельно

Для точного расчёта оптимальной температуры существует множество программ и алгоритмов.

Упрощенный расчёт можно провести по формулам:

$T_m = [(A+T)2^{\circ}\text{C}] + [(G+C)4^{\circ}\text{C}]$ (если длина праймера ≤ 20 оснований)

$T_m = 22 + 1,46([2(G+C)] + (A+T))$ (если длина праймера составляет 20–30 оснований)

Более точные результаты можно получить при использовании формулы:

$$T_m = 81,5 + 16,6(LGM) + 0,41(\%G+C) - (600/L),$$

где M – ионная сила раствора (в молях/л), L – длина олигонуклеотида.

Эта формула пригодна для олигонуклеотидов длиной 14-70 оснований (Сулимова и др., 2006).

Создание маркера - Дизайн праймеров

News: Try out our new tool: [Wiley-DNA-Editor](#) - A DNA/Plasmid editor running in your browser!

Primer3Plus

pick primers from a DNA sequence

[More...](#)

[Source Code](#)

[Help](#)

[About](#)

< Back

Pair 1:

Left Primer 1:

Start: 10 Length: 20 bp Tm: 60.0 °C GC: 55.0 % Any: 0.0 End: 0.0 TB: 6.0 HP: 0.0 3' Stab: 4.8 Penalty: 0.036

Right Primer 1:

Start: 488 Length: 22 bp Tm: 60.0 °C GC: 45.5 % Any: 0.0 End: 0.0 TB: 8.0 HP: 45.9 3' Stab: 2.9 Penalty: 2.041

Pair: Product Size: 479 bp Any: 0.0 End: 0.0 TB: 13.0 Penalty: 2.078

[Send to Primer3Manager](#) [Reset Form](#)

1	ATTATTTCTA	GCTCTCCACA	AACTCCAGCA	ATTGATCCGG	ATTATGCCTT
51	GAATCACATC	CTTCTTTTTC	TCCGAAACCT	TAAGGACCTT	ATGGGTATCA
101	AGAAGATCCT	TAATCGAGAA	AGCAAAAATA	TTTGGGATCT	TACACCACGA
151	ACTAATTCCG	TTCCAAACAT	TCGATGCAAC	GAAGCAAGCA	ATGAAAAGAT
201	GATCAGGAGA	TTCATCATCC	GAACCACAAA	AAGGGCACGA	AGAACATCTCCA
251	GACAGGACGT	TTCTTTTCCT	TAAAGCTTCT	CTGGTCGGAA	TTCTATCCTT
301	AAACATTGAG	TGATATATTT	TACAATATTT	GTATGTATCT	TCATGAGGGA
351	AAATGAGAAG	TTGAAGACAA	CCCCTTTAG	GAGTTTGACC	CTTTTGACAA
401	ATTGACTTTT	ATTAAACAT	GGCATTGTTG	TTTCATCTTC	GGTTTAGGTG
451	AGATTTAAGC	ATGATTGAT	CAAGGATTGG	CTGAACCAAC	AAAGACTAGT
501	T				

Select all Primers

Primer3Plus
pick primers from a DNA sequence

Load server settings Default Activate Settings

Select primer pairs to detect the given template sequence. Optionally targets and included/excluded regions can be specified.

Task: generic

Pick Primers Reset Form

Main General Settings Advanced Settings Internal Oligo Penalty Weights Advanced Sequence

Sequence id: Paste template sequence below: Or upload sequence file: Выбрать файл... или не выбран Upload File

Mark selected region: < > Clear Save Sequence

Excluded Regions: < >

Targets: [345.90]

Included Region: { }

Primer overlap positions: -

Pair OK Region List:

Pick left primer Pick hybridization probe Pick right primer
or use left primer below. (internal oligo) or use oligo below.
or use right primer below ($\beta \rightarrow 3'$ on opposite strand).

Oligo Calc: Oligonucleotide Properties Calculator

Enter Oligonucleotide Sequence Below

OD calculations are for single-stranded DNA or RNA

Nucleotide base codes

TGG TTC AGC CAA TCC TTG ATC A

Reverse Complement Strand(5' to 3') is:

TGA TCA AGG ATT GGC TGA ACC A

5' modification (if any)

3' modification (if any)

Select molecule

50 nM Primer

1 Measured Absorbance at 260 nanometers

50 mM Salt (Na⁺)

Calculate

Swap Strands

BLAST

mfold

Physical Constants

Melting Temperature (T_M) Calculations

Length: 22 Molecular Weight: 6685.4 4 GC content: 45 % 1 53 °C (Basic)

1 ml of a sol'n with an Absorbance of 1 at 260 nm 2 60.1 °C (Salt Adjusted)

is 4.442 microMolar 5 and contains 29.7 micrograms. 3 54.21 °C (Nearest Neighbor)

Thermodynamic Constants Conditions: 1 M NaCl at 25°C at pH 7.

RinK 33.404 cal/(°K²mol)

deltaH 177.2 Kcal/mol

deltaG 28 Kcal/mol

deltaS 465 cal/(°K²mol)

Deprecated Hairpin/self dimerization calculations

4 (Minimum base pairs required for single primer self-dimerization)

4 (Minimum base pairs required for a hairpin)

Check Self-Complementarity

Первый вариант: количественный признак

Задание 2.

Подберите праймеры для использования в ПЦР с целью амплификации фрагментов, содержащих SNP интереса. Данные фрагменты в дальнейшем будут подвергнуты секвенированию по методу Сэнгера для верификации нуклеотида в полиморфной позиции у каждого растения в тестовой выборке с известными фенотипами (и известными генотипами). Используйте программу Primer3Plus + OligoCalc.

Рекомендации по подбору праймеров для амплификации:

- 1). Размер фрагмента должен составлять примерно 200 - 300 п.н.
- 2). SNP интереса должен располагаться не ближе 50 п.н. к каждому из концов фрагмента.
- 3). Праймеры должны иметь T_m около 60С.
- 4). Праймеры должны быть максимально уникальными для анализируемой части генома (мы увидим дальше как проверить уникальность праймеров).

Второй вариант: качественный признак

Содержание токоферолов в масле подсолнечника

Plant material and phenotype data

QTL-mapping of tocopherol content

Phenotypes of interest: the relative content of each tocopherol type (alpha, beta, gamma, delta) i.e ratios of tocopherols in plants.

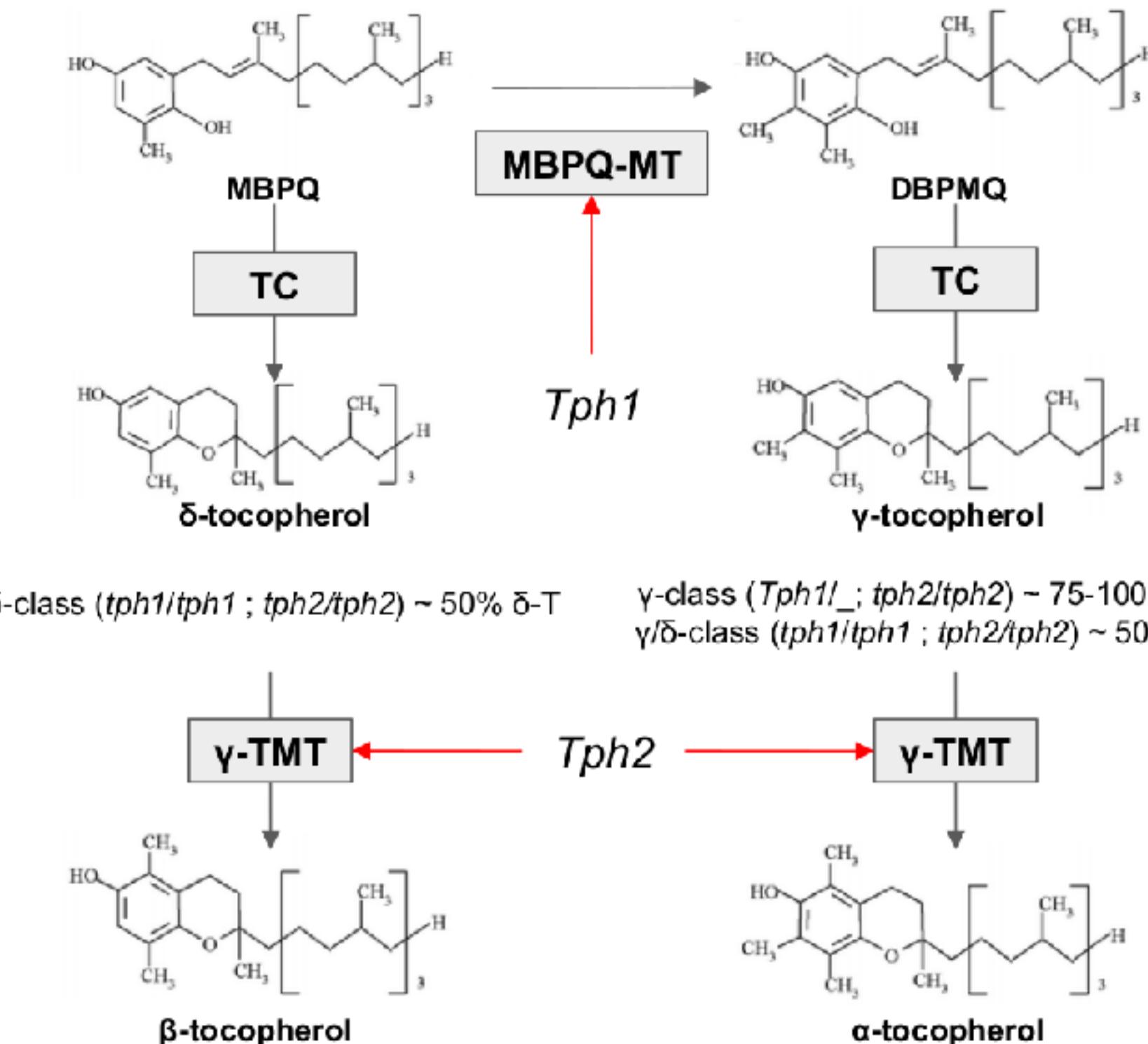
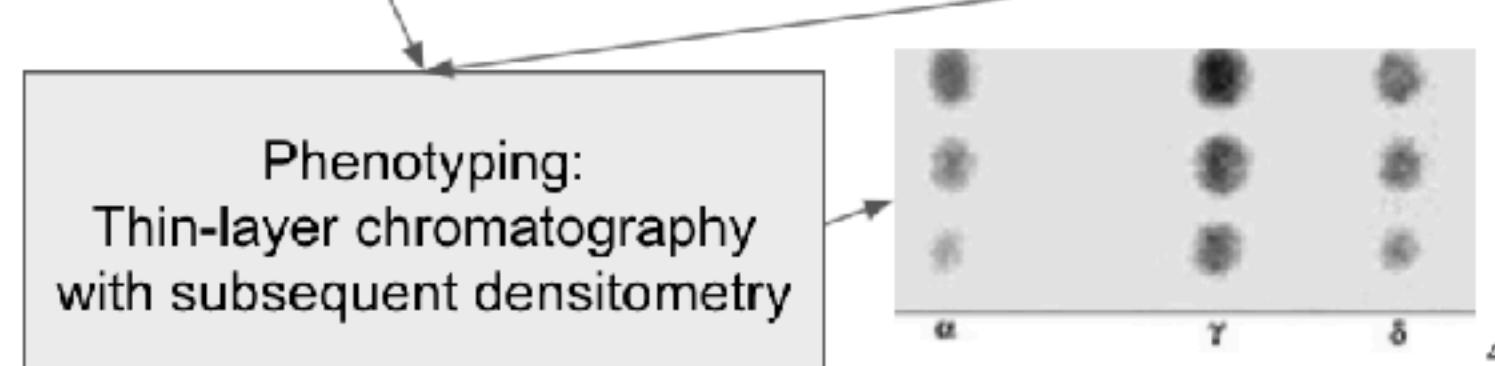
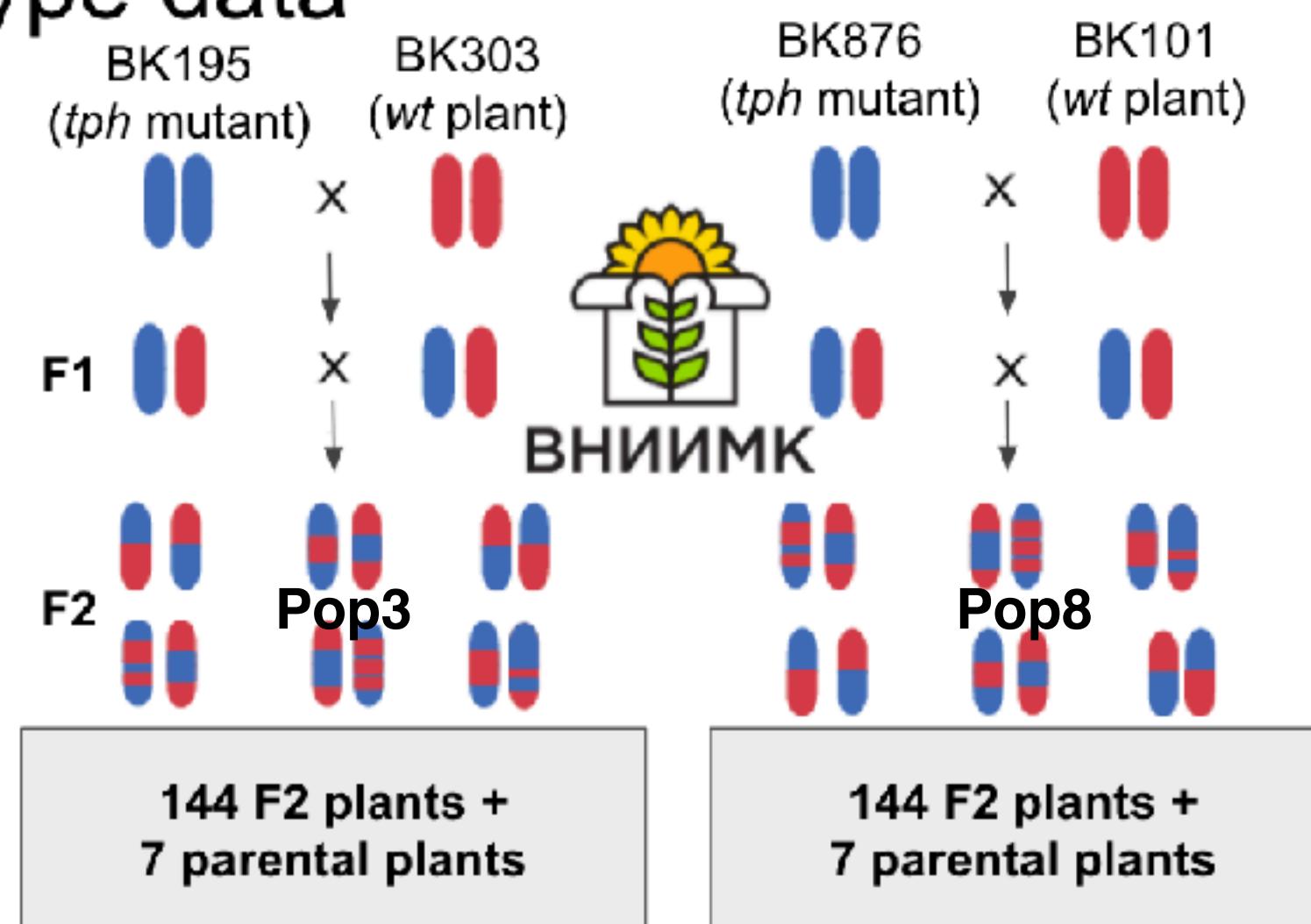
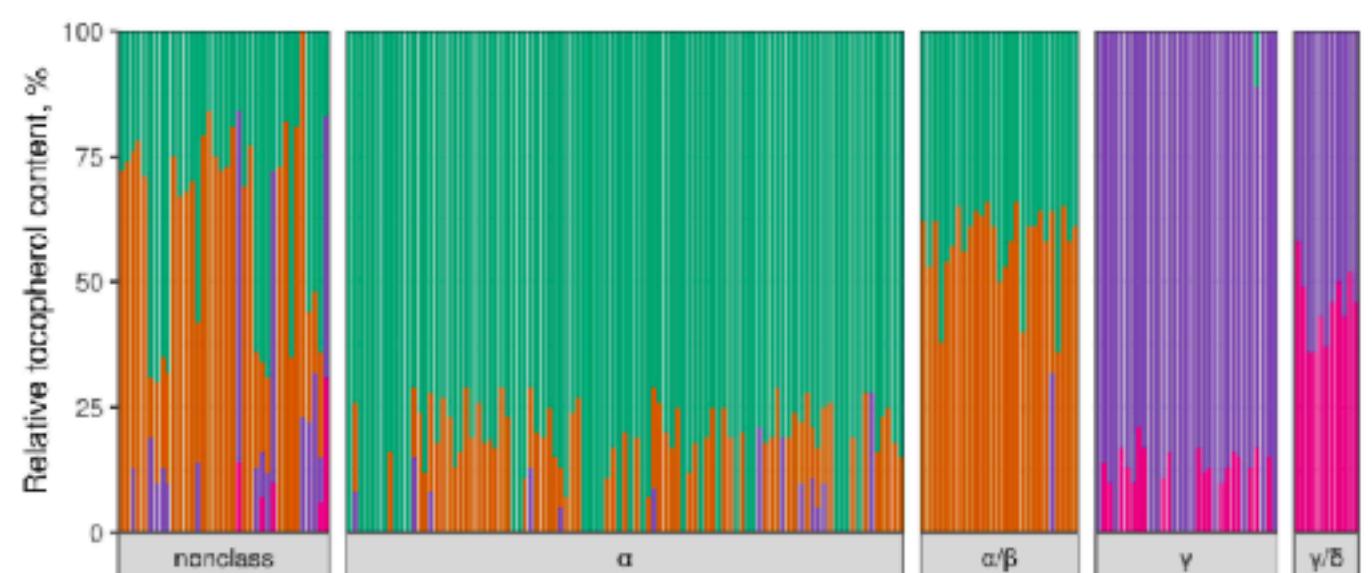
Parents:

- wild type plants (*Tph1**Tph1**Tph2**Tph2*):
 - BK101 ~80-99% of alpha tocopherol
 - BK303 ~80-99% of alpha tocopherol

tocopherol mutants (*tph1tph1tph2tph2*):

- BK195 ~40-60% of gamma and ~40-60% delta tocopherols
- BK876 ~40-60% of gamma and ~40-60% delta tocopherols

Токоферол type: alpha (green), beta (orange), gamma (purple), delta (pink)



Ключевые ферменты, участвующие в биосинтезе токоферолов (финальные стадии синтеза):
MBPQ-MT - 2-methyl-6-phytylbenzoquinol methyltransferase MT (*Tph1*)
gamma-TMT - gamma-tocopherol methyltransferase gamma-TMT (*Tph2*)

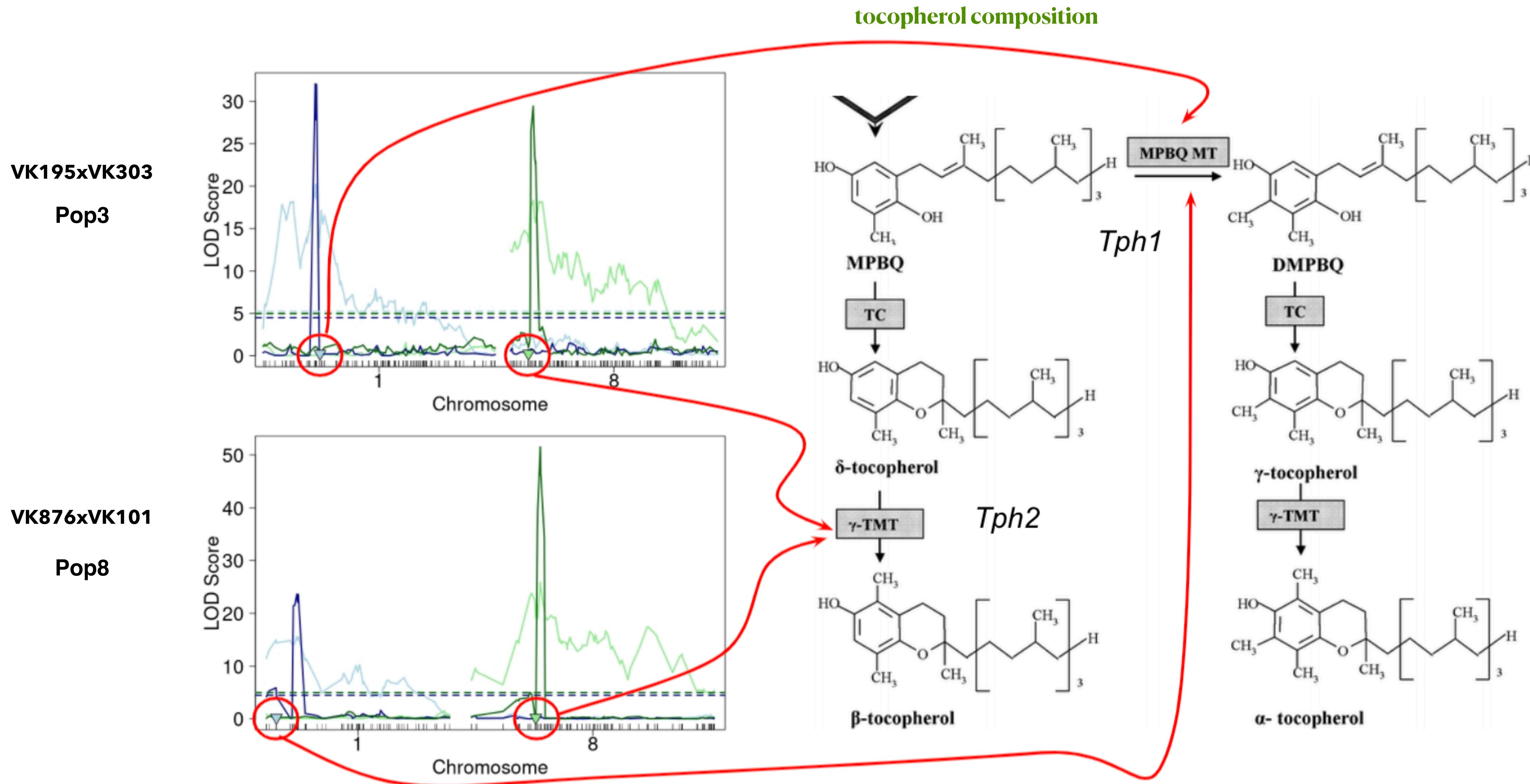
У подсолнечника **НЕТ** природного разнообразия соотношений токоферолов, семена подсолнечника дикого типа содержат почти 100% альфа-токоферол. Мутантные формы были получены с помощью спонтанного мутагенеза и имеют соотношения токоферолов, отличные от альфа-токоферола (дикий тип).

Библиотеки для NGS секвенирования получали методом GBS с использованием рестрикционных эндонуклеаз *PstI* and *MspI*. QTL картирование проводилось с помощью программного обеспечения rQTL методом интервального картирования на основе регрессии Haley-Knott и composite interval mapping

Второй вариант: качественный признак

Содержание токоферолов в масле подсолнечника

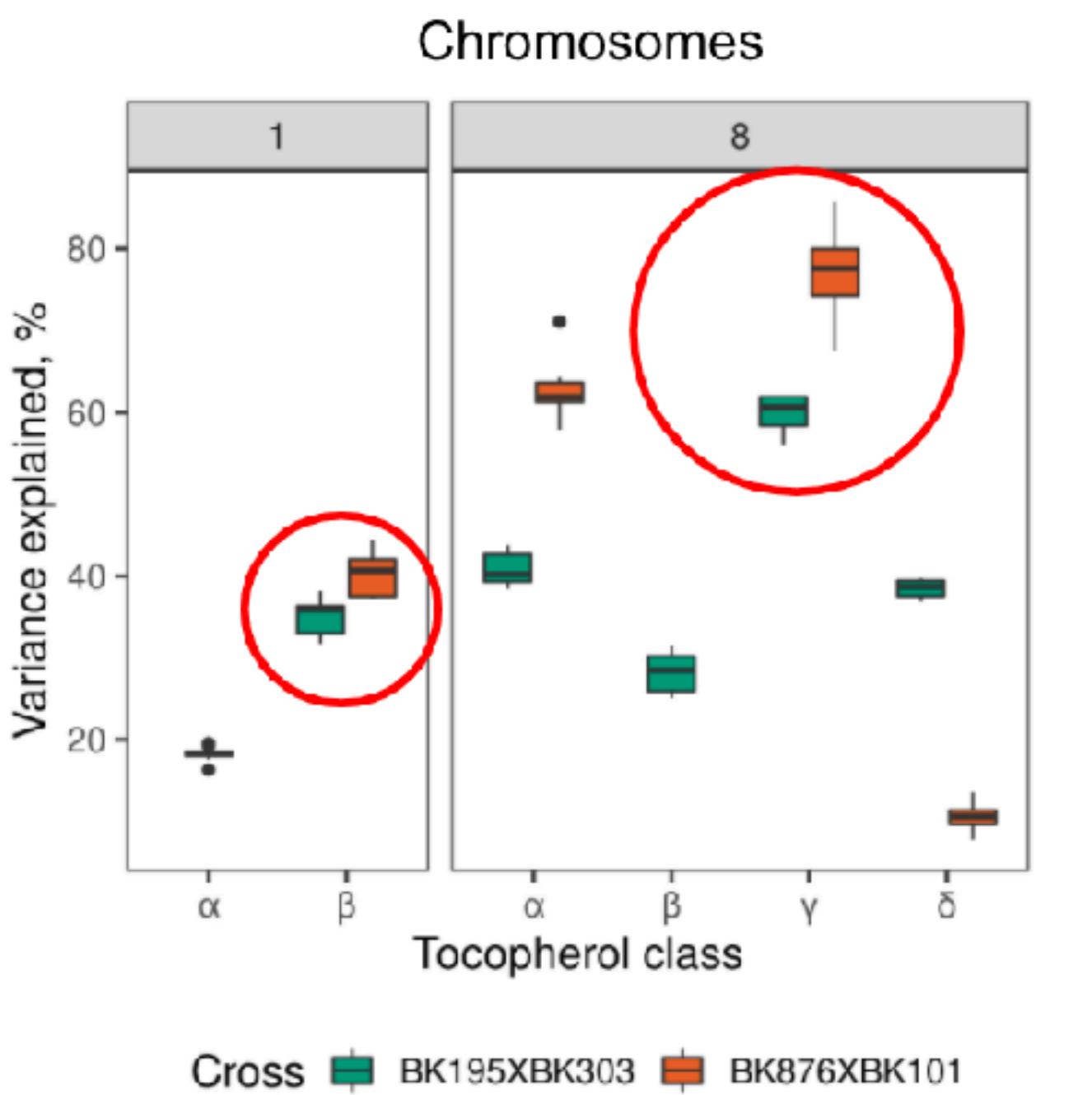
Two loci of major effects were identified on chromosomes 1 and 8 controlling



Второй вариант: качественный признак

Содержание токоферолов в масле подсолнечника

Идентифицированные SNPs



Marker	Gene	Population	Number.of.correctly.predicted.phenotypes...proportion.among.known.genotypes.	Number.of.unknown.genotypes	LOD.Score	mutant_allele (AA) - 0	wt_allele (BB) - 1
S1_68674384	Tph1	pop3	108 (88.52%)	14 (10.29%)	15.3237781627599	A	T
S1_56824042	Tph1	pop3	117 (91.41%)	8 (5.88%)	19.4637701179648	A	G
S1_55474204	Tph1	pop3	119 (92.25%)	7 (5.15%)	19.4638015101479	T	C
S1_55157943	Tph1	pop3	119 (92.25%)	7 (5.15%)	20.197719209787	A	G
S1_55221598	Tph1	pop3	119 (92.97%)	8 (5.88%)	20.1977192104096	T	C
S1_55153369	Tph1	pop3	118 (91.47%)	7 (5.15%)	19.3526455495233	C	T
S1_54474863	Tph1	pop3	112 (91.08%)	13 (9.56%)	18.5754085867686	T	A
S1_54419348	Tph1	pop3	108 (90.76%)	17 (12.5%)	18.5754085831441	G	A
S1_55074343	Tph1	pop3	115 (90.55%)	9 (6.62%)	18.5754085807879	G	T
S1_54598692	Tph1	pop3	115 (91.27%)	10 (7.35%)	18.5754085862152	A	T
S1_55196434	Tph1	pop3	119 (92.25%)	7 (5.15%)	20.1789815085749	T	C
S1_55496596	Tph1	pop3	119 (92.97%)	8 (5.88%)	20.1789813052948	G	A
S1_1055391	Tph1	pop3	92 (82.88%)	25 (18.38%)	12.7743637904406	C	G
S8_23248519	Tph2	pop3	111 (90.24%)	13 (9.56%)	16.5195219645418	A	G
S8_24158872	Tph2	pop3	117 (92.13%)	9 (6.62%)	17.7622764956217	G	A
S8_24015037	Tph2	pop3	117 (92.13%)	9 (6.62%)	17.7622796672159	C	T
S8_24695806	Tph2	pop3	115 (92.74%)	12 (8.82%)	17.8534509520552	A	G
S8_25947498	Tph2	pop3	115 (92%)	11 (8.09%)	18.3371934980591	C	T
S8_26605036	Tph2	pop3	112 (89.6%)	11 (8.09%)	16.1138508608025	C	T
S8_27014900	Tph2	pop3	108 (90%)	16 (11.76%)	16.1138494209159	C	T
S8_27950609	Tph2	pop3	105 (88.24%)	17 (12.5%)	15.7018278511307	A	G
S8_27815509	Tph2	pop3	103 (89.57%)	21 (15.44%)	15.7018451547262	A	G
S8_27576453	Tph2	pop3	105 (89.74%)	19 (13.97%)	15.8268119749985	G	A
S8_27469254	Tph2	pop3	101 (89.38%)	23 (16.91%)	15.826815925037	C	T
S8_30735413	Tph2	pop3	114 (93.44%)	14 (10.29%)	18.26466011484648	C	T
S8_30578572	Tph2	pop3	115 (93.5%)	13 (9.56%)	18.2646601627863	C	T
S8_29865494	Tph2	pop3	109 (93.16%)	19 (13.97%)	18.2646591760182	A	G
S8_29968177	Tph2	pop3	112 (93.33%)	16 (11.76%)	18.2646582072689	T	A
S8_30547454	Tph2	pop3	113 (93.39%)	15 (11.03%)	18.2646572470172	A	G
S8_31293423	Tph2	pop3	114 (92.68%)	13 (9.56%)	18.0347839371589	G	T
S8_31527479	Tph2	pop3	114 (91.94%)	12 (8.82%)	18.0820538957073	T	G
S8_33366420	Tph2	pop3	116 (92.06%)	10 (7.35%)	18.1605890000485	C	T
S8_33006425	Tph2	pop3	116 (92.06%)	10 (7.35%)	18.1605891532897	G	T
S8_32855147	Tph2	pop3	117 (92.13%)	9 (6.62%)	18.1605886597497	T	G
S8_32289442	Tph2	pop3	115 (92%)	11 (8.09%)	18.1618986789839	A	C
S8_44578937	Tph2	pop3	100 (85.47%)	19 (13.97%)	11.3199025395607	C	T
S1_2642582	Tph1	pop8	106 (86.89%)	21 (14.69%)	12.944970775749	G	A
S1_4459825	Tph1	pop8	113 (88.98%)	16 (11.19%)	14.9945921858843	A	G
S1_4445571	Tph1	pop8	114 (89.08%)	15 (10.49%)	14.9945924063595	C	G
S1_4480460	Tph1	pop8	110 (88%)	18 (12.59%)	14.3788731892557	C	T
S1_73083687	Tph1	pop8	113 (91.13%)	19 (13.29%)	14.8622321215383	T	C
S1_73256273	Tph1	pop8	104 (90.43%)	28 (19.58%)	14.8622311170446	G	T
S1_72757269	Tph1	pop8	118 (90.08%)	12 (8.39%)	14.1615913934855	A	G
S1_72750028	Tph1	pop8	117 (90%)	13 (9.09%)	14.1615913802784	G	T
S1_72558643	Tph1	pop8	114 (89.76%)	16 (11.19%)	14.1366590359312	A	G
S1_72373816	Tph1	pop8	118 (90.08%)	12 (8.39%)	14.2905104054621	C	G
S1_72423427	Tph1	pop8	119 (89.47%)	10 (6.99%)	14.2905104138193	C	T
S1_70724774	Tph1	pop8	111 (88.8%)	18 (12.59%)	15.6033922457912	T	A
S1_71141056	Tph1	pop8	112 (89.6%)	18 (12.59%)	15.6033920227286	A	T
S1_71031923	Tph1	pop8	111 (89.52%)	19 (13.29%)	15.6033917989473	T	C
S1_70999280	Tph1	pop8	112 (89.6%)	18 (12.59%)	15.6033915750782	T	A
S1_71599015	Tph1	pop8	121 (89.63%)	8 (5.59%)	14.9290747980466	A	C
S1_71228195	Tph1	pop8	115 (89.15%)	14 (9.79%)	14.9290747809422	T	C
S1_71251574	Tph1	pop8	116 (89.23%)	13 (9.09%)	14.9290747811839	C	A
S1_71435702	Tph1	pop8	120 (89.55%)	9 (6.29%)	14.9290747814228	A	G
S1_71566195	Tph1	pop8	120 (89.5%)	9 (6.29%)	14.9290746250135	T	A
S1_71748138	Tph1	pop8	120 (90.23%)	10 (6.99%)	15.566415291556	T	C
S1_71778451	Tph1	pop8	119 (90.15%)	11 (7.69%)	15.5664150703785	C	T
S1_71701057	Tph1	pop8	122 (90.37%)	8 (5.59%)	15.5664148478213	A	G
S1_80144596	Tph1	pop8	115 (83.94%)	6 (4.2%)	11.020331684211	A	T
S8_19958120	Tph2	pop8	126 (94.74%)	10 (6.99%)	19.0158224431079	A	T
S8_23941299	Tph2	pop8	124 (96.12%)	14 (9.79%)	24.6612633583864	T	C
S8_23732662	Tph2	pop8	128 (96.24%)	10 (6.99%)	24.6612638078175	A	G
S8_23902485	Tph2	pop8	125 (96.15%)	13 (9.09%)	24.6594772790411	T	C
S8_24016228	Tph2	pop8	124 (96.12%)	14 (9.79%)	24.6594772591678	C	T
S8_42026530	Tph2	pop8	117 (93.6%)	18 (12.59%)	20.595039671587	G	C

Второй вариант: качественный признак

Задание 3.

Учитывая комбинированные статистические показатели для каждого идентифицированного SNP - LOD-score и количество правильно предсказанных фенотипов, выберите два SNP для каждой популяции (рор3 и рор8) для каждого признака (Trh1 - увеличенное содержание бета-токоферола и Trh2 - увеличенное содержание гамма/дельта-токоферолов), которые на Ваш взгляд являются оптимальными кандидатами для разработки на их основе молекулярного маркера.

Показатель LOD или логарифм отношения шансов (logarythm of odds). Показатель LOD = \log_{10} (отношение шансов).

Отношение шансов - это отношение между вероятностью того, что SNP связан (ассоциирован) с интересующим признаком, и вероятностью того, что он не ассоциирован с признаком. Чем выше логарифм отношения шансов, тем более высока вероятность связи между SNP и признаком. Оценка LOD, равная 3, указывает на вероятность 1 к 10^3 того, что наблюдаемая ассоциация обусловлена случайностью (связь отсутствует), что часто принимается за пороговое значение.

Число правильно предсказанных фенотипов для идентифицированных генотипов - фактическое число (и процент) отдельных растений в выборке, для которых генотип был идентифицирован с помощью GBS и фенотип, предсказанный на основе генотипа, совпал с фактическим фенотипом для этого растения (содержание токоферола).

EnsemblPlants genome database

Первая версия сборки - Sunflower genome *Helianthus annuus*(HanXRQr1.0)

https://plants.ensembl.org/Helianthus_annuus/Info/Index

Существует вторая версия сборки - *Helianthus annuus* XRQ2.0

<https://www.heliagene.org/HanXRQr2.0-SUNRISE/>

Объект секвенирования - сорт XRQ (INRA (National Institute of Agricultural Research) genebank)

Метод секвенирования - Illumina Hiseq data + Pacbio

Существуют две другие версии генома, полученные другими авторами.

Используются более редко.

V2.0 - Chromosomes + unplaced contigs (2020)

V1.0 - Chromosomes + unplaced contigs (2017).

Второй вариант: качественный признак

Задание 4.

Используя сборку генома подсолнечника *Helianthus annuus* версии 1 (HanXRQr1.0) выделите нуклеотидную последовательность длиной 501 нт (или 1001 нт), которая будет содержать один из выбранных в ходе выполнения Задания 2 SNP в середине последовательности так, чтобы по обе стороны от SNP было по 250 нт или по 500 нт. Пометьте положение SNP в последовательности квадратными скобками. Через знак «/» добавьте альтернативный вариант SNP, указанный в Таблице (например, [T/C]).
Соберите все полученные последовательности в один файл в простом текстовом редакторе типа TextPad или WordPad Сохраните файл в формате .fa, .fas. или .fasta.

Primer specificity - an important thing to check. May be tricky^)

PrimerBlast + Ensemble
GenomeBrowser

https://www.ncbi.nlm.nih.gov/tools/primer-blast/index.cgi?LINK_LOC=reset

Primer-BLAST A tool for finding specific primers

Finding primers specific to your PCR template (using Primer3 and BLAST).

Primers for target on one template Primers common for a group of sequences

PCR Template Enter accession, gi, or FASTA sequence (A refseq record is preferred). Range From To Save search parameters Reset parameters

Forward primer Reverse primer

Or, upload FASTA file Выбрать файл файл не выбран

Primer Parameters Use my own forward primer (5'->3' on plus strand) Use my own reverse primer (5'->3' on minus strand) Min Max PCR product size 70 1000 # of primers to return 10 Min Opt Max Max T_m difference 57.0 60.0 63.0 3 Note: Parameter values that differ from the default are highlighted in yellow

Primer Pair Specificity Checking Parameters Specificity check: Enable search for primer pairs specific to the intended PCR template Search mode: Automatic Database: RefSeq representative genomes Exclusion: Exclude predicted RefSeq transcripts (accession with XM, XR prefix) Exclude uncultured/environmental sample sequences Organism: Helianthus annuus (taxid:4232) Add organism Entrez query (optional) Primer specificity stringency: Primer must have at least 2 total mismatches to unintended targets, including at least 2 mismatches within the last 5 bps at the 3' end. Ignore targets that have 6 or more mismatches to the primer. Max target amplicon size: 4000 Allow splice variants: Allow primer to amplify mRNA splice variants (requires refseq mRNA sequence as PCR template input) Get Primers Show results in a new window Use new graphic view Note: Parameter values that differ from the default are highlighted in yellow

Primer-BLAST A tool for finding specific primers

Finding primers specific to your PCR template (using Primer3 and BLAST).

Input PCR template: S0_19958120_Tph2_pop6 Range: 1 - 501

Your PCR template is highly similar to the following sequence(s) from the search database. To increase the chance of finding specific primers, please review the list below and select all sequences (within the given sequence ranges) that are intended or allowed targets.

Select: All None Selected:1

Accession	Title	Identity	Alignment length	Seq. start	Seq. stop	Gene
<input checked="" type="checkbox"/> NC_035440.2	Helianthus annuus cult var XRO/B chromosome 8, HanXRO/2.0-SUNRISE, whole genome shotgun sequence	100%	501	1995870	1995870	LOC110872310

Submit Show results in a new window

Input PCR template: S0_24016228_Tph2_pop6 Range: 1 - 501

Specificity of primers: Primer pairs are specific to Input template as no other targets were found in selected database: RefSeq Representative Genome Database (Organism limited to Helianthus annuus)

Other reports: Search Summary

An (almost) ideal situation

Graphical view of primer pairs

Template: 20 40 60 80 100 120 140 160 180 200 220 240 260 280 300 320 340 360 380 400 420 440 460 480 500

(0) Primer pairs for job KCL31X-0c1bV8nDentNTMC84OgMta1xLa

Query_1 ... 501 (501 nt)

Primer 1, Primer 2, Primer 3, Primer 4, Primer 5

Tracks shown: 2/3

Detailed primer reports

Primer pair 1

Forward primer	Sequence (5'-3')	Template strand	Length	Start	Stop	Tm	GC%	Self complementarity	Self 3' complementarity
GCAGGCTAGCATCACTCACT	Plus	20	299	318	59.82	55.00	6.00	1.00	2.00
ACACTCTCCCCGAGTTGATG	Minus	20	448	429	59.10	55.00	4.00		
Product length = 150									
Features associated with this product:									
nac domain-containing protein_78									
Forward primer	1	GCAGGCTAGCATCACTCACT	20						
Template	24016276		24016295					
Reverse primer	1	ACACTCTCCCCGAGTTGATG	20						
Template	24016425		24016406					

Primer specificity - an important thing to check. May be tricky^)

PrimerBlast + Ensemble
GenomeBrowser

[https://plants.ensembl.org/
Helianthus_annuus/Info/Index](https://plants.ensembl.org/Helianthus_annuus/Info/Index)

452 BLAST/BLAT ▾

Results for Job 1: BLASTN against Helianthus annuus HanXRQr1.0 (Genomic sequence) ⓘ

Job details ⓘ

Job name: Job 1: BLASTN against Helianthus annuus HanXRQr1.0 (Genomic sequence)
Species: Helianthus annuus
Assembly: HanXRQr1.0
Search type: BLASTN (NCBI BLAST)

A disastrous situation

Download results file New job

Results table ⓘ

Show All entries	Show/hide columns (4 hidden)	Filter			
Genomic Location	Overlapping Gene(s)	Orientation	Length	Score	E-val
13:35789628-35789647 [Sequence]		Forward	20 [Sequence]	20	0.0075
13:55897708-55897727 [Sequence]		Forward	20 [Sequence]	20	0.0075
13:68212470-68212489 [Sequence]		Reverse	20 [Sequence]	20	0.0075
17:82930489-82930508 [Sequence]		Forward	20 [Sequence]	20	0.0075
17:47405476-47405495 [Sequence]		Reverse	20 [Sequence]	20	0.0075
17:94068238-94068257 [Sequence]		Reverse	20 [Sequence]	20	0.0075
3:33091583-33091602 [Sequence]		Reverse	20 [Sequence]	20	0.0075
5:119350559-119350578 [Sequence]		Forward	20 [Sequence]	20	0.0075
15:164284506-164284525 [Sequence]		Forward	20 [Sequence]	20	0.0075
15:29683117-29683136 [Sequence]		Reverse	20 [Sequence]	20	0.0075
15:51901608-51901627 [Sequence]		Reverse	20 [Sequence]	20	0.0075
10:63183604-63183623 [Sequence]		Forward	20 [Sequence]	20	0.0075
10:111093818-111093837 [Sequence]		Forward	20 [Sequence]	20	0.0075

Check the gene

NCBI again

Gene ID 110872346

Search Results:

<https://www.ncbi.nlm.nih.gov/gene/?term=110872346>

Gene symbol

LOC110872346

Gene description

probable tocopherol O-methyltransferase, chloroplastic

Locus tag

HannXRQ_Chr08g0217461

Gene type

protein coding

Also known as

G-TMT; gamma-TMT

Full Report Send to:

Showing Current items.

LOC110872346 probable tocopherol O-methyltransferase, chloroplastic [*Helianthus annuus* (common sunflower)]

Gene ID: 110872346, updated on 22-Mar-2021

[Download Datasets](#)

Summary

Gene symbol: LOC110872346
Gene description: probable tocopherol O-methyltransferase, chloroplastic
Locus tag: HannXRQ_Chr08g0217461
Gene type: protein coding
RefSeq status: MODEL
Organism: *Helianthus annuus*
Lineage: Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliopsida; eudicotyledons; Gunneridae; Pentapetales; asterids; campanulids; Asterales; Asteraceae; Asteroideae; Heliantheae alliance; Heliantheae; Helianthus
Also known as: G-TMT; gamma-TMT

Genomic context

Location: chromosome 8
Exon count: 6

Annotation release: 101 (current), 100 (previous)

Genomic regions, transcripts, and products

Chromosome 8 - NC_035440.2

Genomic Sequence: NC_035440.2 Chromosome 8 Reference HanXRQr2.0-SUNRISE Primary Assembly

Go to reference sequence details

Go to nucleotide: Graphics, FASTA, GenBank

Tools, Tracks, Download, ?

NC_035440.2: 22M..22M (5,445 nt)

1,954 K | 21,954,500 | 21,955 K | 21,955,500 | 21,956 K | 21,956,500 | 21,957 K | 21,957,500 | 21,958 K | 21,958,500 | 21,959 K

Genes, NCBT Helianthus annuus Annotation Release 101, 2020-09-02

XM_022121127.2 LOC110872346 XP_021976819.1

Genes, TNSDC annotation provided by Sunflower Genomics Consortium

mRNA-putative toco... HanXRQr2_Chr08g0328591

RNA-seq exon coverage, aggregate (filtered), NCBT Helianthus annuus Annotation Release 101 - log base 2 scaled

RNA-seq intron-spanning reads, aggregate (filtered), NCBT Helianthus annuus Annotation Release 101 - log base 2 scaled

RNA-seq intron features, aggregate (filtered), NCBT Helianthus annuus Annotation Release 101

Tracks shown: 6/295

Bibliography

GeneRIFs: Gene References Into Functions

What's a GeneRIF?

Submit: New GeneRIF Correction

Gene ID 110865372

<https://www.ncbi.nlm.nih.gov/gene/110865372>

Be careful: Paralog MT-2

Gene symbol

LOC110865372

Gene description

2-methyl-6-phytyl-1,4-hydroquinone methyltransferase, chloroplastic

Locus tag

HannXRQ_Chr01g0006941

Gene type

protein coding

Pentapetales; asterids; campanulids; Asterales; Asteraceae; Asteroideae; Heliantheae alliance; Heliantheae; Helianthus

Also known as

IN37; MT-1

Второй вариант: качественный признак

Задание 5.

Подберите праймеры для использования в ПЦР с целью амплификации фрагментов, содержащих SNP интереса. Данные фрагменты в дальнейшем будут подвергнуты секвенированию по методу Сэнгера для верификации нуклеотида в полиморфной позиции у каждого растения в тестовой выборке с известными фенотипами (и известными генотипами). Используйте программу PrimerBlast (<https://www.ncbi.nlm.nih.gov/tools/primer-blast/>).

Рекомендации по подбору праймеров для амплификации:

- 1). Размер фрагмента должен составлять не более 200-300 пн для фрагментов размером 501 нт и не более 400 п.н. Для фрагментов размером 1001 пн.
- 2). SNP интереса должен располагаться не ближе 50 п.н. к каждому из концов фрагмента.
- 3). Праймеры должны иметь Т_m около 60С.
- 4). Праймеры должны быть максимально уникальными для анализируемой части генома (мы увидим дальше как проверить уникальность праймеров).

Задание 6.

Убедитесь в уникальности подобранных Вами праймеров для генома подсолнечника с помощью программы BLAST и/или Primer-BLAST.

Верификация SNP - Секвенирование по Сэнгеру

Популяция для верификации маркеров

Должна включать в себя минимум обоих родителей, а также выборку растений из картирующей популяции (mapping population) с известными фенотипами/SNP генотипами для сравнения. В идеале нужно использовать растения, которые не участвовали в проведении QTL картирования, для которых известны фенотипы.

QTL validation (recommendation from a Practical guide)

Прежде чем использовать QTL для МАС, его необходимо валидировать (подтвердить), чтобы исключить возможность статистических аномалий или ошибок. В процессе валидации проверяется, проявляется ли тот же QTL при выращивании материала в других местах и/или в другие годы, и можно ли обнаружить его эффект на отличном генетическим фоне, то есть в другой выборке/популяции растений. (<https://www.integratedbreeding.net/courses/marker-assisted-breeding/index2ea.html?id=139>)

Наша выборка

Родители - 5 растений для каждого из родителей исходного кросса (ВК195 и ВК 303)

Потомство F₂ - по 10 растений гомозигот каждого типа и 10 гетерозигот

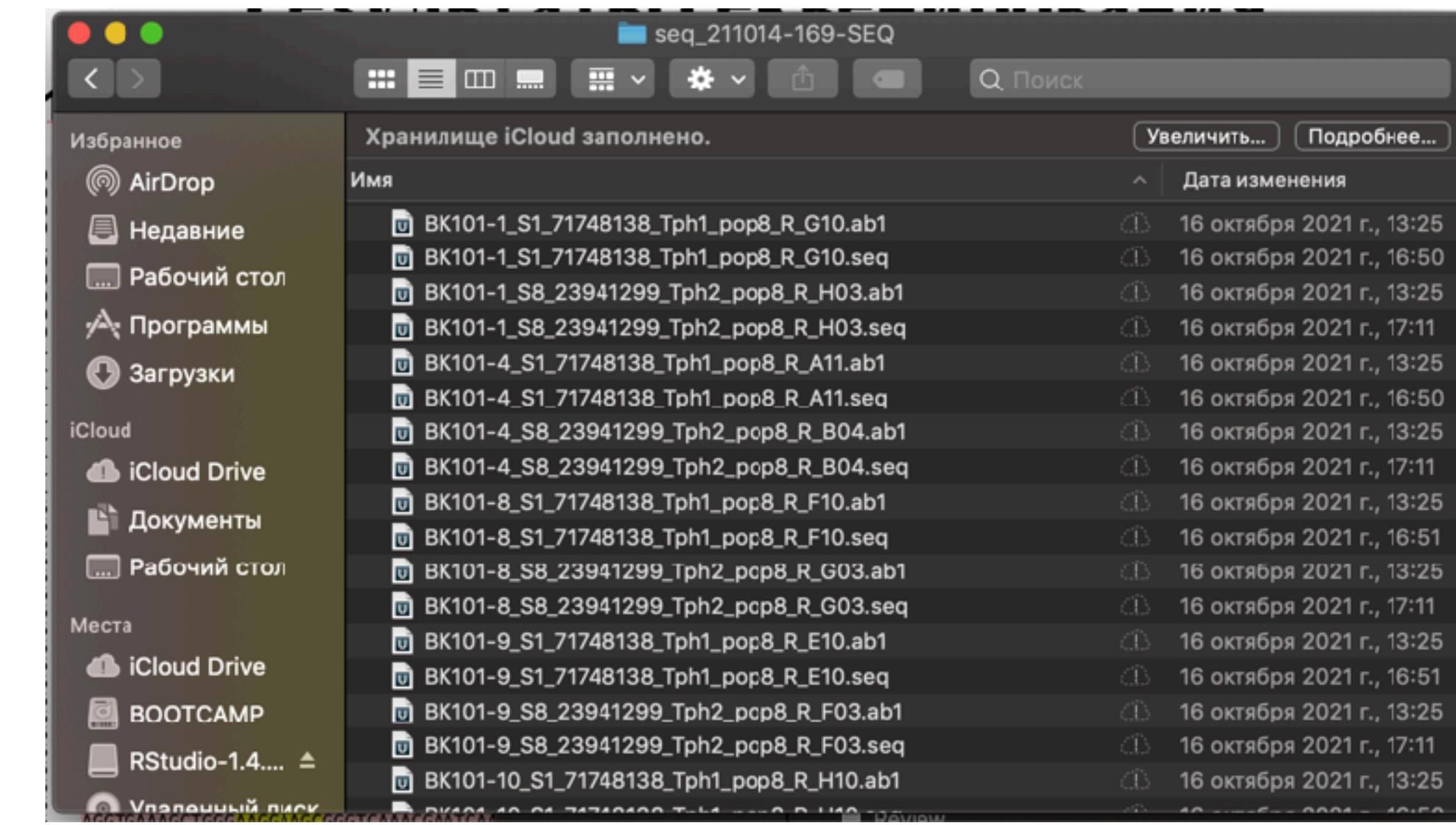
Результаты секвенирования

Множественное выравнивание последовательностей
<http://tcoffee.crg.cat/>

S1_55196434_Tph1_pop3 C(wt)/T(mut)

BK195-4 S1 5519 AGGTGAAAGCTGGGAAGGAAGCAGGTCAAACGAATCAAGTATGAAGGAAAAACATAATTGGT
BK195-5 S1 5519 AGGTGAAAGCTGGGAAGGAAGCGGGTCAAACGAATCAAGTATGAAGGAAAAACATAATTGGT
BK195-6 S1 5519 AGGTGAAAGCTGGGAAGGAAGCAGGTCAAACGAATCAAGTATGAAGGAAAAACATAATTGGT
BK195-8 S1 5519 AGGTGAAAGCTGGGAAGGAAGCAGGTCAAACGAATCAAGTATGAAGGAAAAACATAATTGGT
BK195-10 S1 551 AGGTGAAAGCTGGGAAGGAAGCAGGTCAAACGAATCAAGTATGAAGGAAAAACATAATTGGT
BK303-3 S1 5519 AGGTGAAAGCTGGGAAGGAAGCGGGTCAAACGAATCAAGTATGAAGGAAAAACATAATTGGT
BK303-4 S1 5519 AGGTGAAAGCTGGGAAGGAAGCAGGTCAAACGAATCAAGTATGAAGGAAAAACATAATTGGT
BK303-5 S1 5519 AGGTGAAAGCTGGGAAGGAAGCGGGTCAAACGAATCAAGTATGAAGG - AAAACATAATTGGT
BK303-7 S1 5519 AGGTGAAAGCTGGGAAGGAAGCAGGTCAAACGAATCAAGTATGAAGGAAAAACATAATTGGT
BK303-9 S1 5519 AGGTGAAAGCTGGGAAGGAAGCAGGTCAAACGAATCAAGTATGAAGGAAAAACATAATTGGT
sts10 S1 551964 AGGTGAAAGCTGGGAAGGAAGCAGGTCAAACGAATCAAGTATGAAGGAAAAACATAATTGGT
sts12 S1 551964 AGGTGAAAGCTGGGAAGGAAGCAGGTCAAACGAATCAAGTATGAAGGAAAAACATAATTGGT
sts14 S1 551964 AGGTGAAAGCTGGGAAGGAAGCAGGTCAAACGAATCAAGTATGAAGGAAAAACATAATTGGT
sts17 S1 551964 AGGTGAAAGCTGGGAAGGAAGCAGGTCAAACGAATCAAGTATGAAGGAAAAACATAATTGGT
sts19 S1 551964 AGGTGAAAGCTGGGAAGGAAGCAGGTCAAACGAATCAAGTATGAAGGAAAAACATAATTGGT
sts21 S1 551964 AGGTGAAAGCTGGGAAGGAAGCAGGTCAAACGAATCAAGTATGAAGGAAAAACATAATTGGT
sts22 S1 551964 AGGTGAAAGCTGGGAAGGAAGCAGGTCAAACGAATCAAGTATGAAGGAAAAACATAATTGGT
sts23 S1 551964 AGGTGAAAGCTGGGAAGGAAGCAGGTCAAACGAATCAAGTATGAAGGAAAAACATAATTGGT
sts26 S1 551964 AGGTGAAAGCTGGGAAGGAAGCAGGTCAAACGAATCAAGTATGAAGGAAAAACATAATTGGT
sts100 S1 55196 AGGTGAAAGCTGGGAAGGAAGCAGGTCAAACGAATCAAGTATGAAGGAAAAACATAATTGGT
sts107 S1 55196 AGGTGAAAGCTGGGAAGGAAGCAGGTCAAACGAATCAAGTATGAAGGAAAAACATAATTGGT
sts109 S1 55196 AGGTGAAAGCTGGGAAGGAAGCAGGTCAAACGAATCAAGTATGAAGGAAAAACATAATTGGT
sts112 S1 55196 AGGTGAAAGCTGGGAAGGAAGCAGGTCAAACGAATCAAGTATGAAGGAAAAACATAATTGGT
sts113 S1 55196 AGGTGAAAGCTGGGAAGGAAGCAGGTCAAACGAATCAAGTATGAAGGAAAAACATAATTGGT
sts114 S1 55196 AGGTGAAAGCTGGGAAGGAAGCAGGTCAAACGAATCAAGTATGAAGGAAAAACATAATTGGT
sts115 S1 55196 AGGTGAAAGCTGGGAAGGAAGCAGGTCAAACGAATCAAGTATGAAGGAAAAACATAATTGGT
sts116 S1 55196 AGGTGAAAGCTGGGAAGGAAGCAGGTCAAACGAATCAAGTATGAAGGAAAAACATAATTGGT
sts117 S1 55196 AGGTGAAAGCTGGGAAGGAAGCAGGTCAAACGAATCAAGTATGAAGGAAAAACATAATTGGT
sts118 S1 55196 AGGTGAAAGCTGGGAAGGAAGCAGGTCAAACGAATCAAGTATGAAGGAAAAACATAATTGGT
sts122 S1 55196 AGGTGAAAGCTGGGAAGGAAGCAGGTCAAACGAATCAAGTATGAAGGAAAAACATAATTGGT
sts124 S1 55196 AGGTGAAAGCTGGGAAGGAAGCAGGTCAAACGAATCAAGTATGAAGGAAAAACATAATTGGT
sts125 S1 55196 AGGTGAAAGCTGGGAAGGAAGCAGGTCAAACGAATCAAGTATGAAGGAAAAACATAATTGGT
sts126 S1 55196 AGGTGAAAGCTGGGAAGGAAGCAGGTCAAACGAATCAAGTATGAAGGAAAAACATAATTGGT
sts127 S1 55196 AGGTGAAAGCTGGGAAGGAAGCAGGTCAAACGAATCAAGTATGAAGGAAAAACATAATTGGT
sts128 S1 55196 AGGTGAAAGCTGGGAAGGAAGCAGGTCAAACGAATCAAGTATGAAGGAAAAACATAATTGGT
sts139 S1 55196 AGGTGAAAGCTGGGAAGGAAGCAGGTCAAACGAATCAAGTATGAAGGAAAAACATAATTGGT
sts147 S1 55196 AGGTGAAAGCTGGGAAGGAAGCAGGTCAAACGAATCAAGTATGAAGGAAAAACATAATTGGT
sts161 S1 55196 AGGTGAAAGCTGGGAAGGAAGCAGGTCAAACGAATCAAGTATGAAGGAAAAACATAATTGGT
sts176 S1 55196 AGGTGAAAGCTGGGAAGGAAGCAGGTCAAACGAATCAAGTATGAAGGAAAAACATAATTGGT
sts178 S1 55196 AGGTGAAAGCTGGGAAGGAAGCAGGTCAAACGAATCAAGTATGAAGGAAAAACATAATTGGT
sts180 S1 55196 AGGTGAAAGCTGGGAAGGAAGCAGGTCAAACGAATCAAGTATGAAGGAAAAACATAATTGGT
sts188 S1 55196 AGGTGAAAGCTGGGAAGGAAGCAGGTCAAACGAATCAAGTATGAAGGAAAAACATAATTGGT
sts189 S1 55196 AGGTGAAAGCTGGGAAGGAAGCAGGTCAAACGAATCAAGTATGAAGGAAAAACATAATTGGT
sts196 S1 55196 AGGTGAAAGCTGGGAAGGAAGCAGGTCAAACGAATCAAGTATGAAGGAAAAACATAATTGGT

cons *****

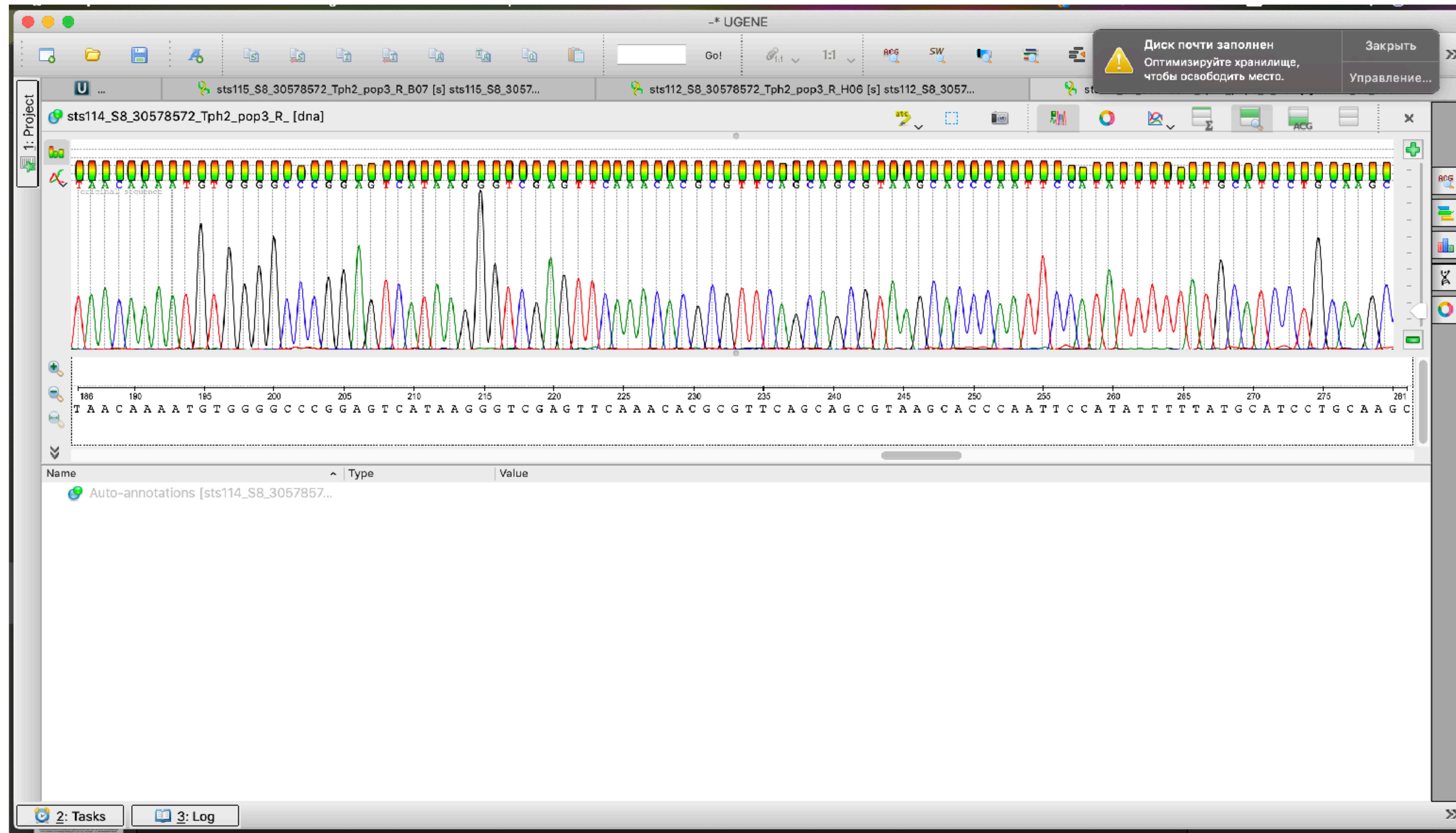


Результаты секвенирования в Fasta формате

```
>BK195-4_S1_55196434_Tph1_pop3_R_E12
ATTTGACTCTAGTGTATTACTTACGCTAATAGGTAGTTGACTTTGAAAAGTCAAACTAACCAAAATGAGATCAAATGGTAATTGATATAAGTCCTGAAGATGAG
ATAGTCTTAATAATGGTACTAATCTAACCGTCTTACGAATGATAAAAAGGTAGTTGACACTTGACAAGCTAGGTGAAAGCTGGGAAGGAAGCAGGTCAAACGAA
TCAAGTATGAAGGAAAAACATAATTGGTCAAGGTAAAGCTTACACTCTTTTAATTGTAGAATATTCCCTATTGTATTGATTACAATAAGATAAAAGTA
TTTGAATGTATGTTCCGACATGTAGTTGACAGGAATGAAATAAGTGCACATGATAAAAAACCATGTGTA
>BK195-5_S1_55196434_Tph1_pop3_R
ATTTGACTCATAGTGTATTACTTACGCTAATAGGTAGTTGACTTTGAAAAGTCAAACTAACCAAAATGAGATCAAATGGTAATTGATATAAGTCCTGAAGATGAG
ATAGTCTTAATAATGGTACTAATCTAACCGTCTTACGAATGATAAAAAGGTAGTTGACACTTGACAAGCTAGGTGAAAGCTGGGAAGGAAGCAGGTCAAACGAA
TCAAGTATGAAGGAAAAACATAATTGGTCAAGGTAAAGCTTACACTCTTTTAATTGTAGAATATTCCCTATTGTATTGATTACAATAAGATAAAAGTA
TTTGAATGTATGTTCCGACATGTAGTTGACAGGAATGAAATAAGTGCACATGATAAAAAACCATGTGAA
>BK195-6_S1_55196434_Tph1_pop3_R
ATTTGACTCTAGTGTATTACTTACGCTAATAGGTAGTTGACTTTGAAAAGTCAAACTAACCAAAATGAGATCAAATGGTAATTGATATAAGTCCTGAAGATGAG
ATAGTCTTAATAATGGTACTAATCTAACCGTCTTACGAATGATAAAAAGGTAGTTGACACTTGACAAGCTAGGTGAAAGCTGGGAAGGAAGCAGGTCAAACGAA
TCAAGTATGAAGGAAAAACATAATTGGTCAAGGTAAAGCTTACACTCTTTTAATTGTAGAATATTCCCTATTGTATTGATTACAATAAGATAAAAGTA
TTTGAATGTATGTTCCGACATGTAGTTGACAGGAATGAAATAAGTGCACATGATAAAAAACCATGTGAA
>BK195-8_S1_55196434_Tph1_pop3_R_D12
ATTTGACTCATAGTGTATTACTTACGCTAATAGGTAGTTGACTTTGAAAAGTCAAACTAACCAAAATGAGATCAAATGGTAATTGATATAAGTCCTGAAGATGAG
GATAGTCTTAATAATGGTACTAATCTAACCGTCTTACGAATGATAAAAAGGTAGTTGACACTTGACAAGCTAGGTGAAAGCTGGGAAGGAAGCAGGTCAAACGAA
AATCAAGTATGAAGGAAAAACATAATTGGTCAAGGTAAAGCTTACACTCTTTTAATTGTAGAATATTCCCTATTGTATTGATTACAATAAGATAAAAGTA
TTTGAATGTATGTTCCGACATGTAGTTGACAGGAATGAAATAAGTGCACATGATAAAAAACCATGTGAA
>BK195-10_S1_55196434_Tph1_pop3_R
ATTTGACTCTAGTGTATTACTTACGCTAATAGGTAGTTGACTTTGAAAAGTCAAACTAACCAAAATGAGATCAAATGGTAATTGATATAAGTCCTGAAGATGAG
GATAGTCTTAATAATGGTACTAATCTAACCGTCTTACGAATGATAAAAAGGTAGTTGACACTTGACAAGCTAGGTGAAAGCTGGGAAGGAAGCAGGTCAAACGAA
AATCAAGTATGAAGGAAAAACATAATTGGTCAAGGTAAAGCTTACACTCTTTTAATTGTAGAATATTCCCTATTGTATTGATTACAATAAGATAAAAGTA
TTTGAATGTATGTTCCGACATGTAGTTGACAGGAATGAAATAAGTGCACATGATAAAAAACCATGTGAA
>BK303-3_S1_55196434_Tph1_pop3_R_H12
ATTTGACTCATAGTGTATTACTTACGCTAATAGGTAGTTGACTTTGAAAAGTCAAACTAACCAAAATGAGATCAAATGGTAATTGATATAAGTCCTGAAGATGAG
GATAGTCTTAATAATGGTACTAATCTAACCGTCTTACGAATGATAAAAAGGTAGTTGACACTTGACAAGCTAGGTGAAAGCTGGGAAGGAAGCAGGTCAAACGAA
ATCAAGTATGAAGGAAAAACATAATTGGTCAAGGTAAAGCTTACACTCTTTTAATTGTAGAATATTCCCTATTGTATTGATTACAATAAGATAAAAGTA
TTTGAATGTATGTTCCGACATGTAGTTGACAGGAATGAAATAAGTGCACATGATAAAAAACCATGTGAA
```

Результаты секвенирования

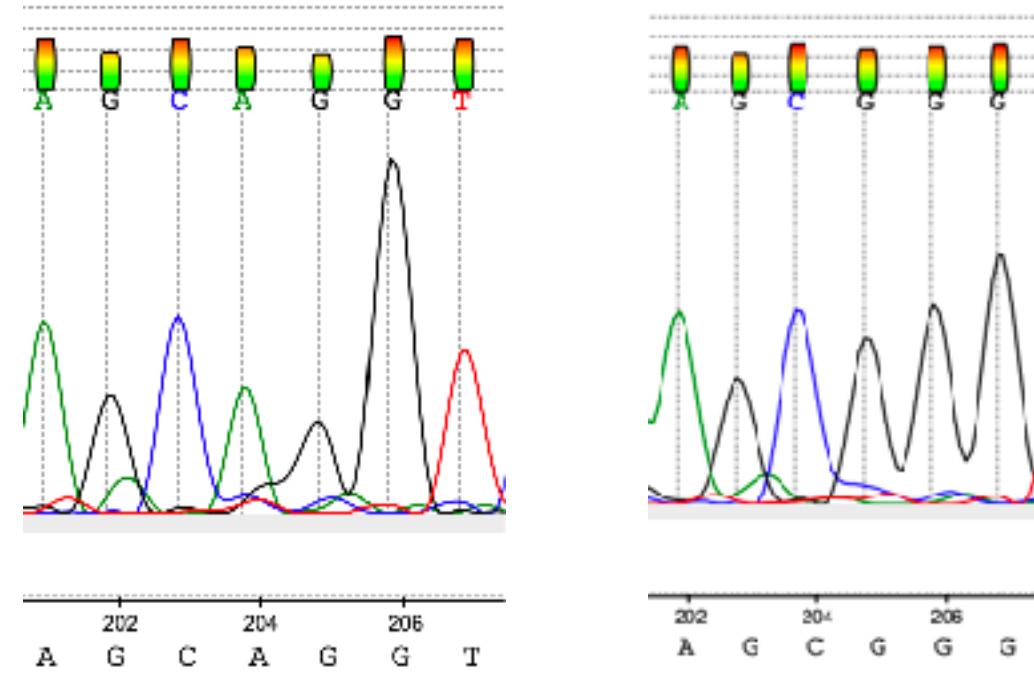
Ugene software - universal tool for a molecular geneticist <http://ugene.net/ru/download-all.html>



Результаты секвенирования

Ugene software - universal tool for a molecular geneticist <http://ugene.net/ru/download-all.html>

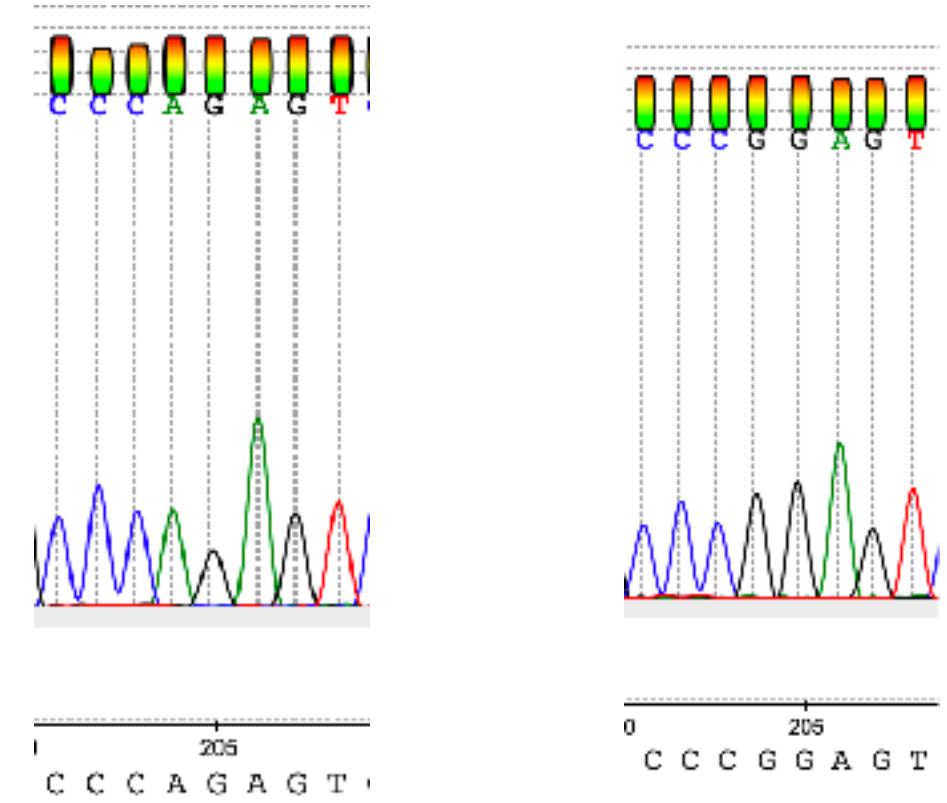
S1_55196434_Tph1_pop3 C(wt)/T(mut) GAAGCAGGTC



BK195 родитель Ж АА

BK303 родитель М ВВ

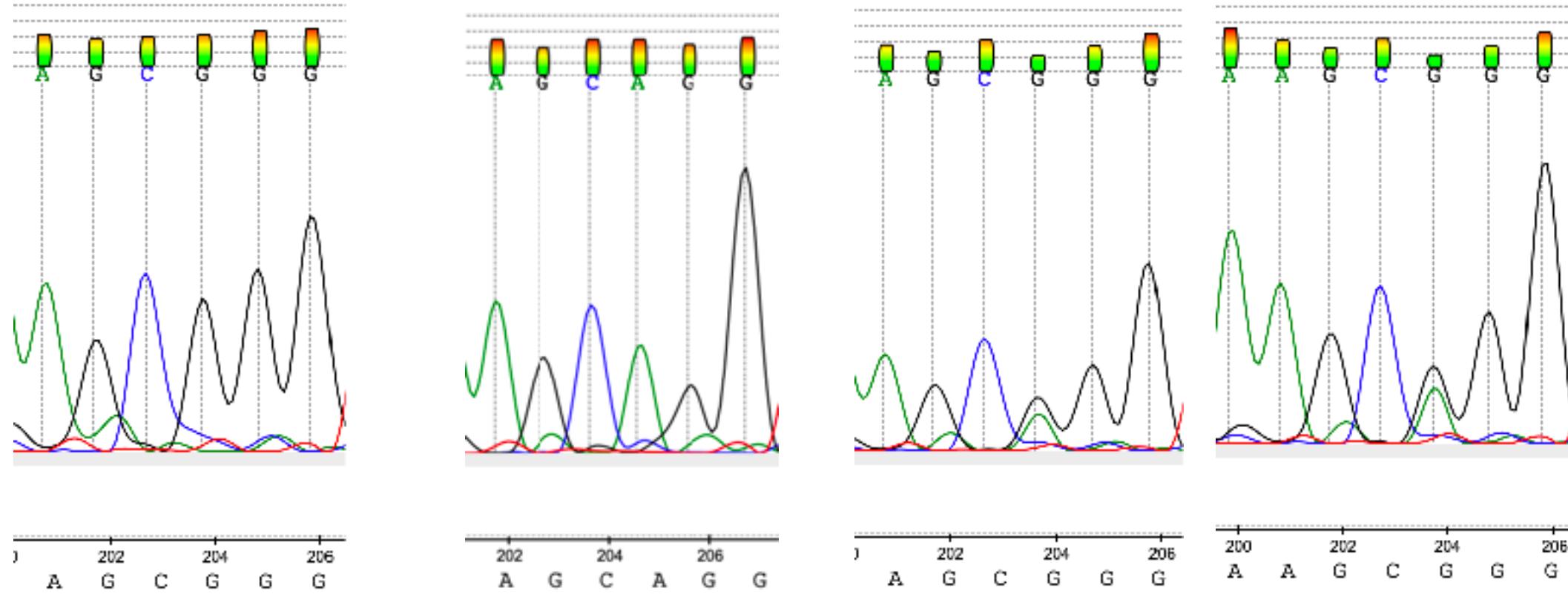
S8_30578572_Tph2_pop3 C(wt)/T(mut) GCCCGGGACT



BK195 родитель Ж АА

BK303 родитель М ВВ

Потомство F2

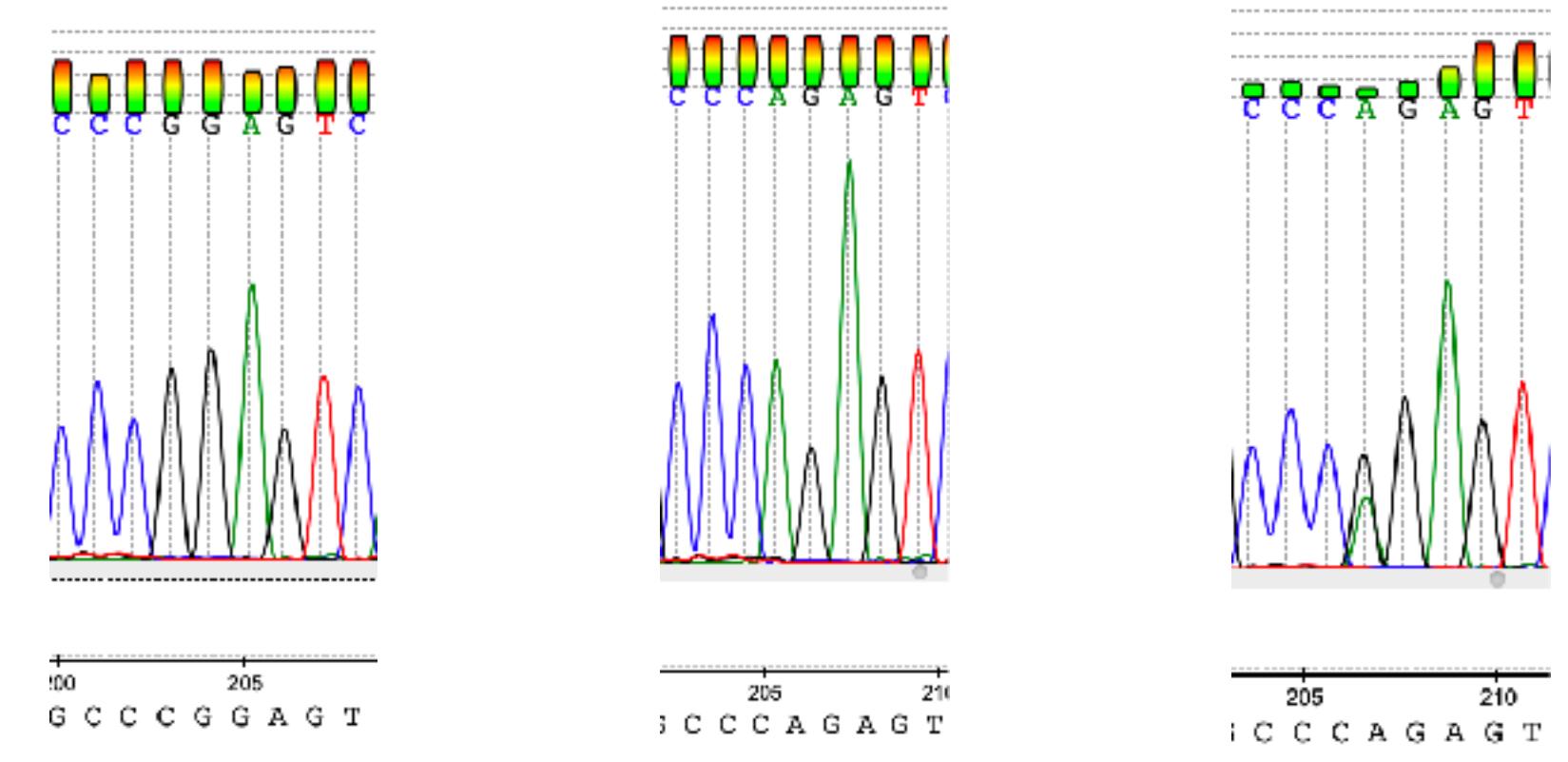


Sts107-Гомозигота ВВ

Sts10-Гомозигота АА

Sts118-Sts100-Гетерозигота АВ

Потомство F2



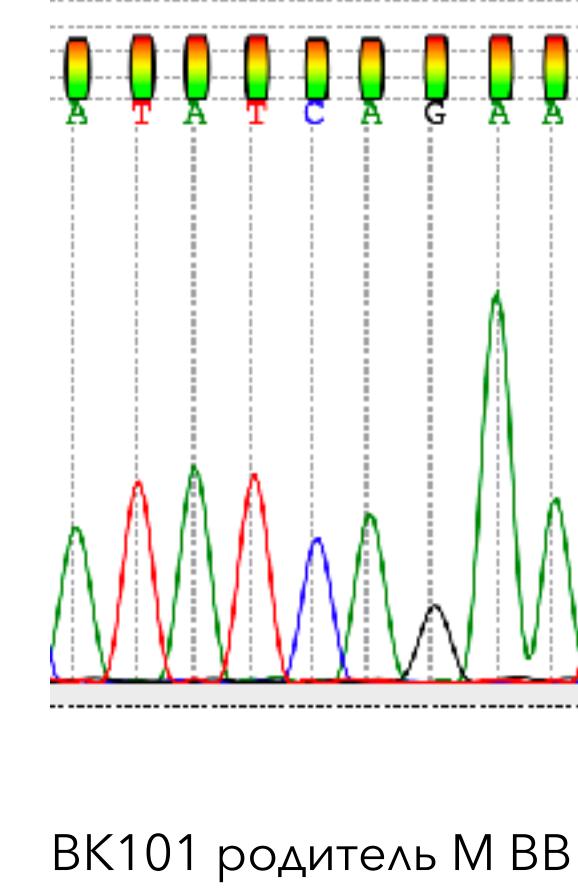
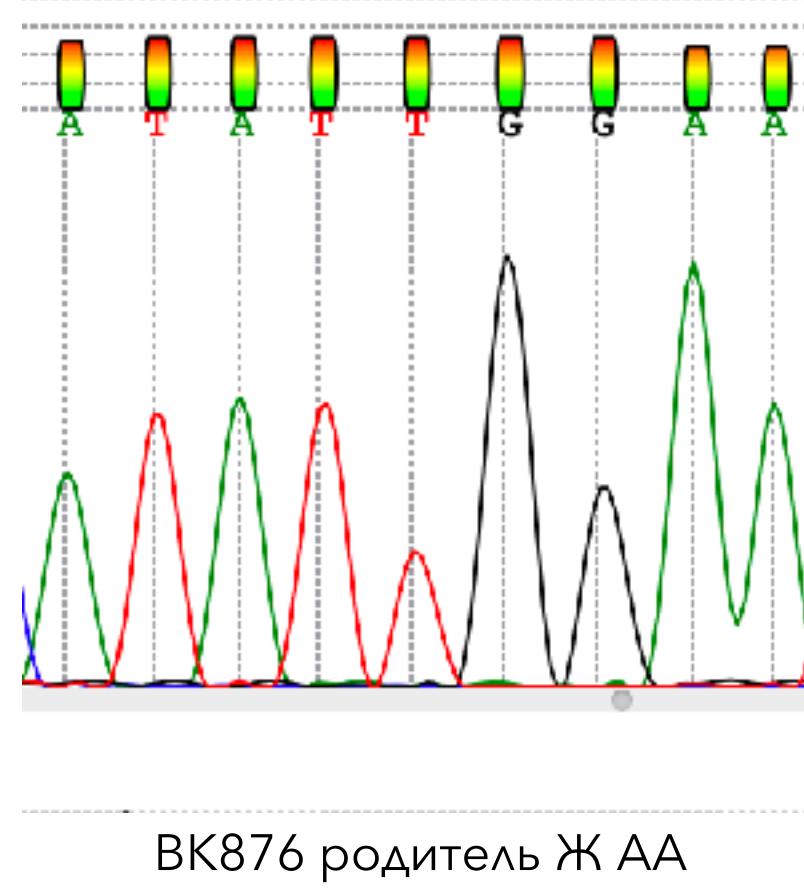
Sts112-Гомозигота ВВ

Sts115-Гомозигота АА

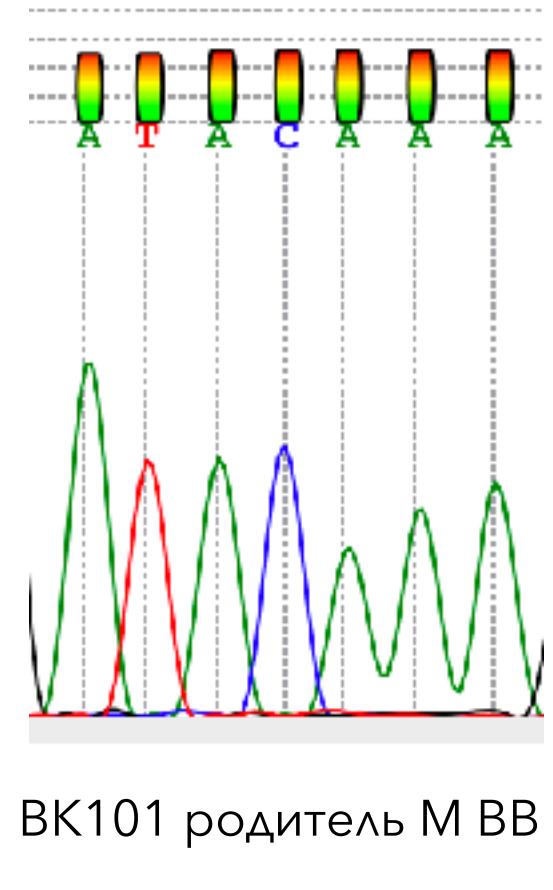
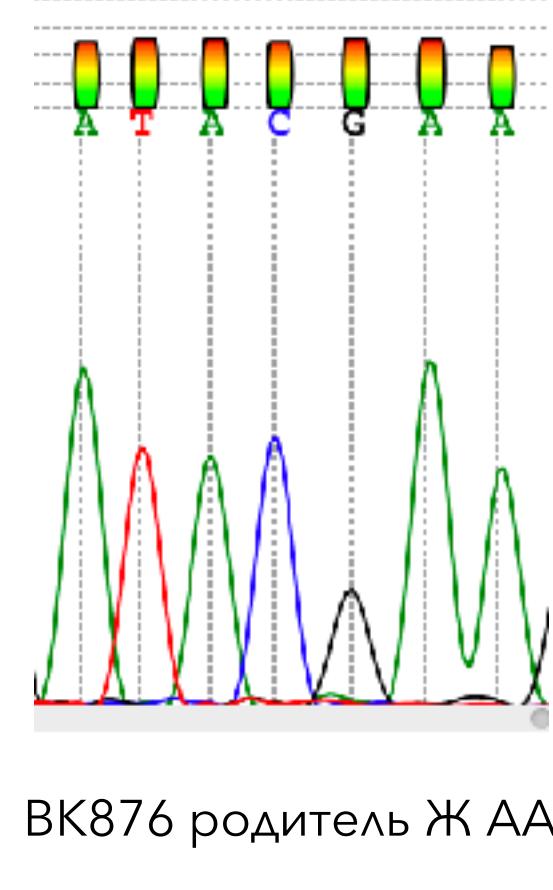
Sts112-Гетерозигота АВ

Результаты секвенирования

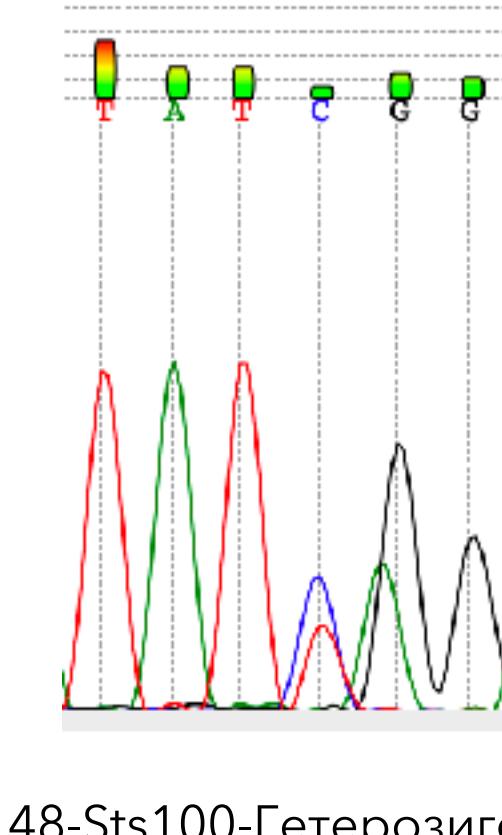
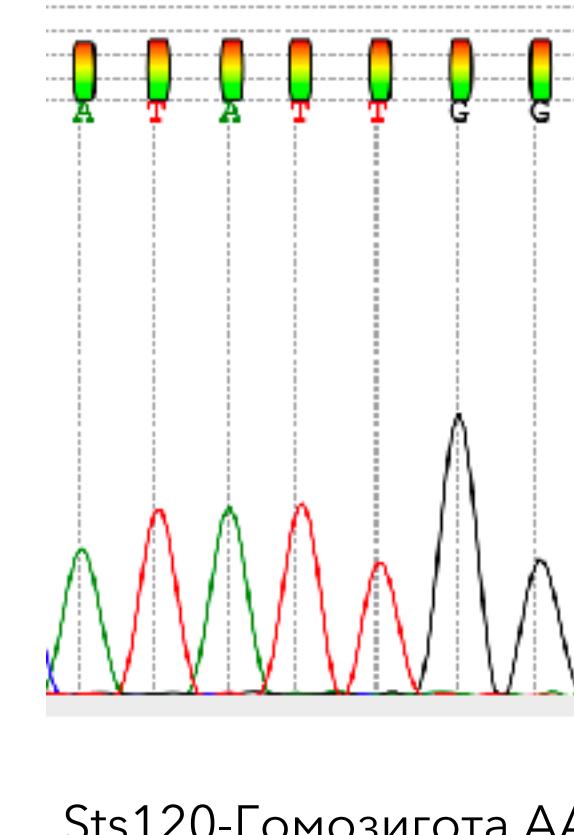
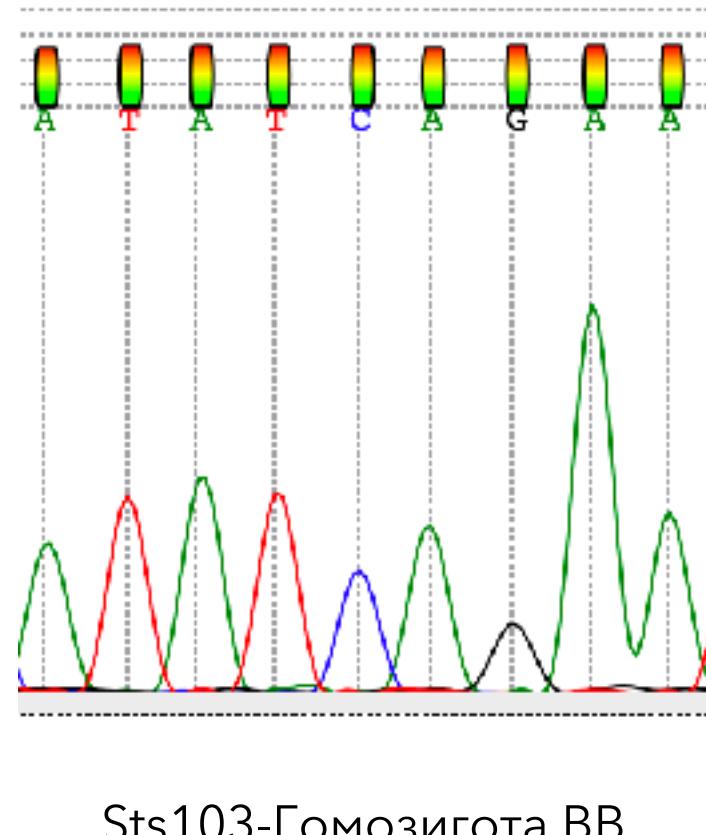
S1_71748138_Tph1_pop8 C(wt)/T(mut) ATATTGGAAT



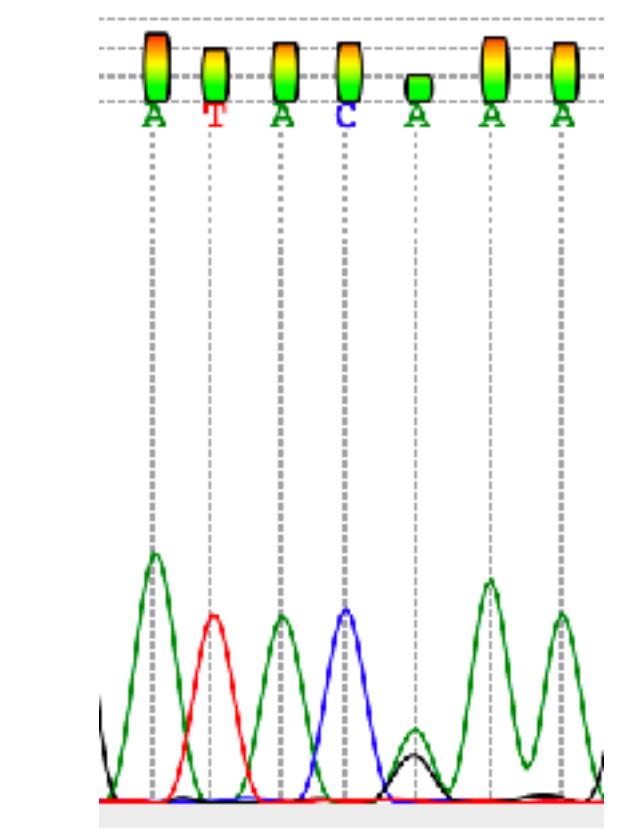
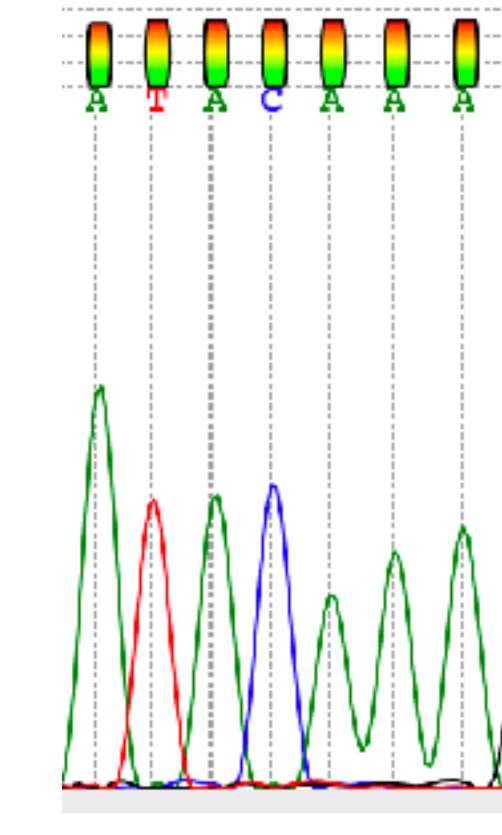
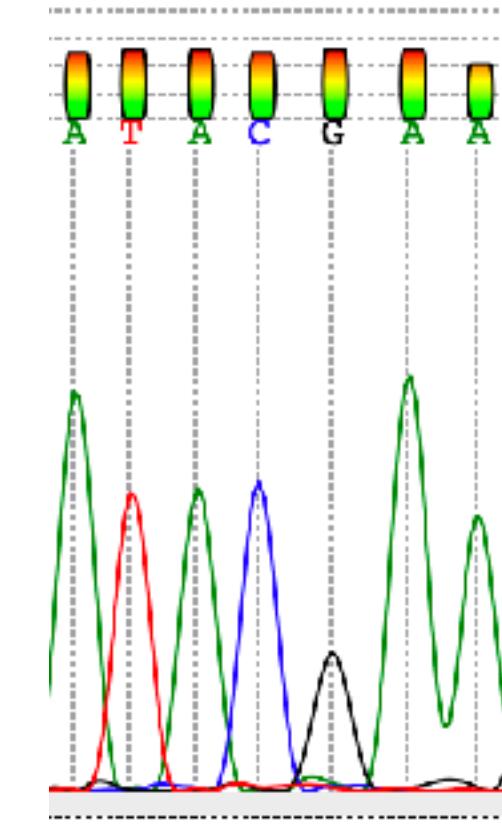
S8_23941299_Tph2_pop8 C(wt)/T(mut) ATACGAAGG



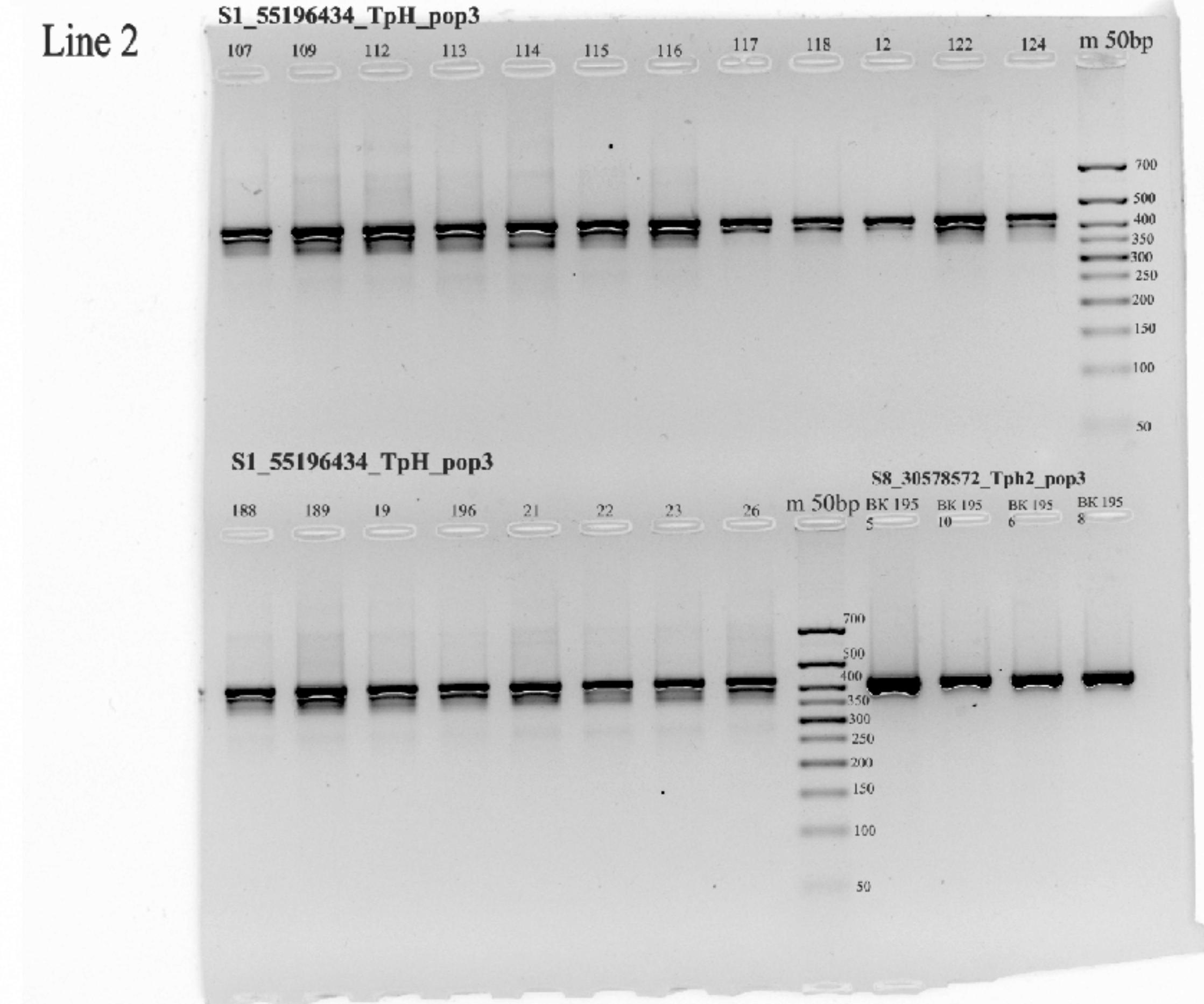
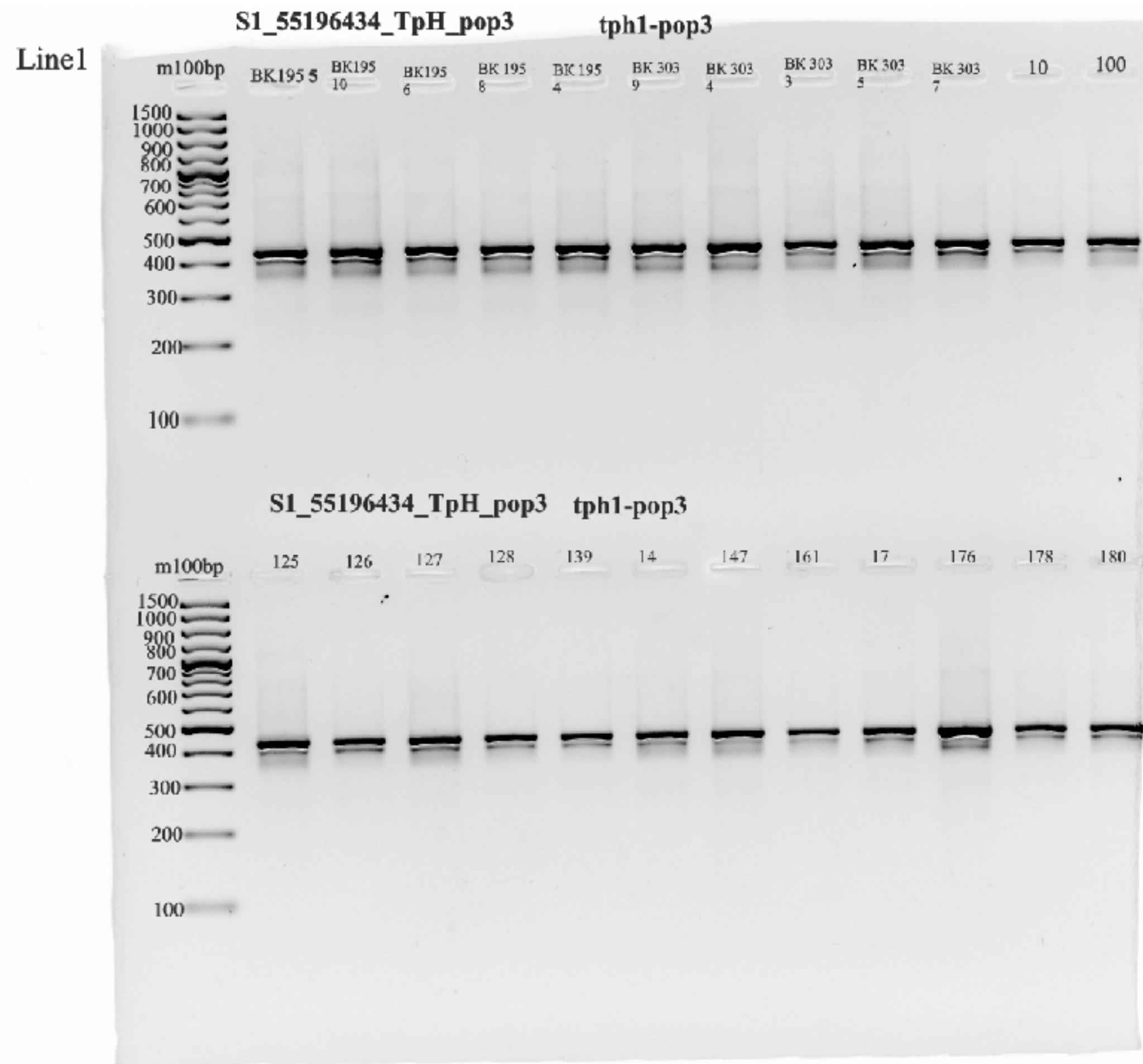
Потомство F2



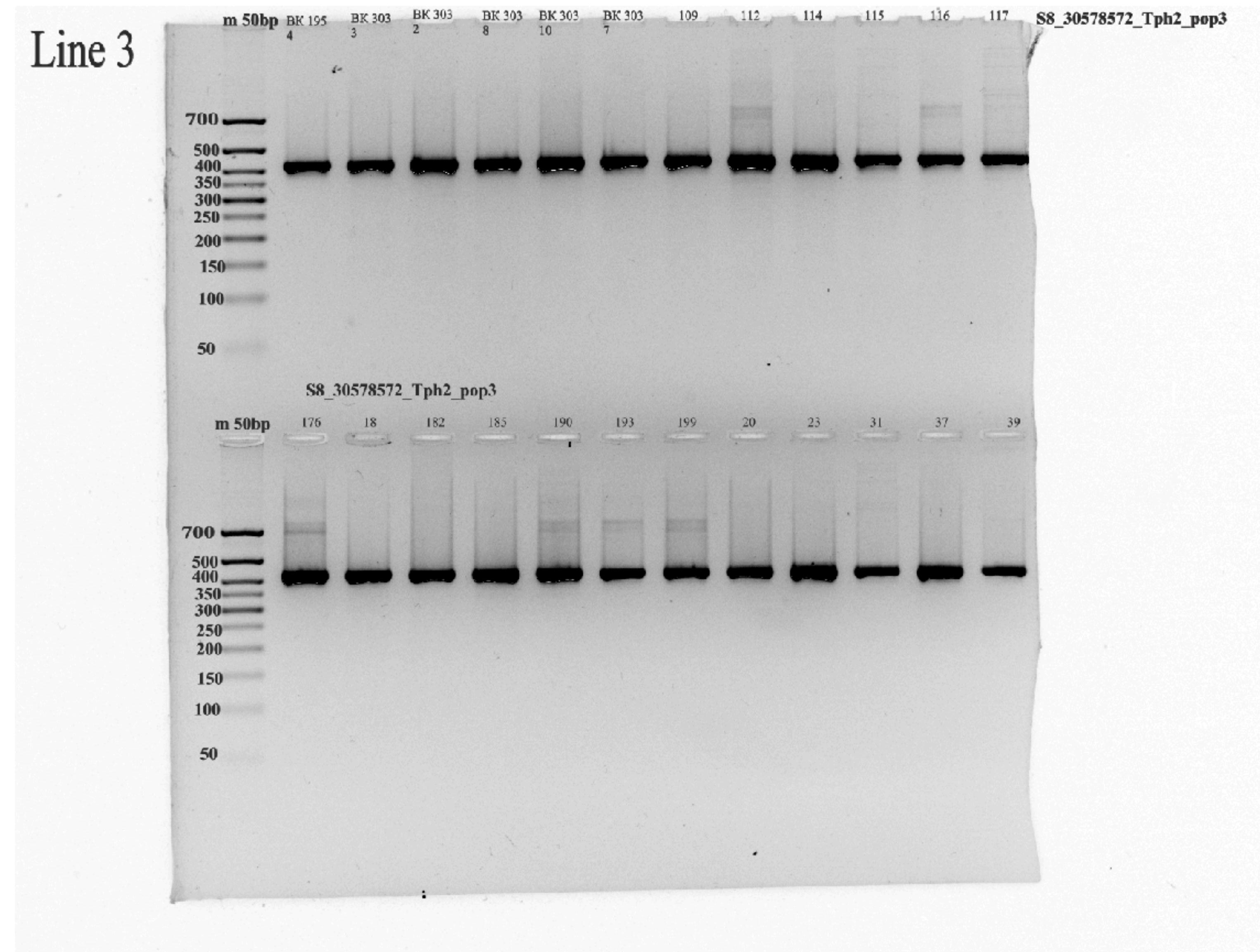
Потомство F2



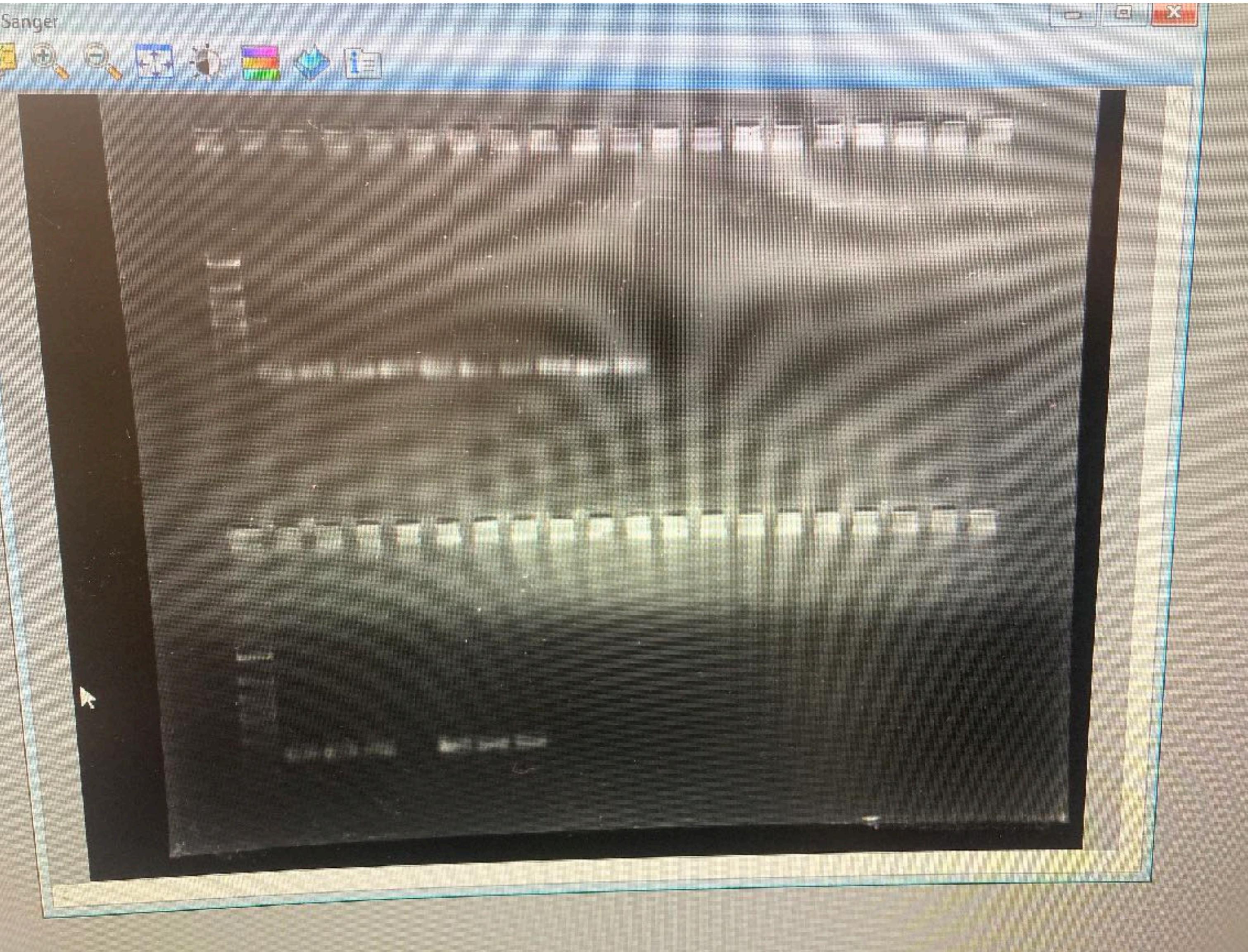
Результаты ПЦР амплификации для верификации SNP методом секвенирования по Сэнгеру



Результаты ПЦР амплификации для верификации SNP методом секвенирования по Сэнгеру



Результаты ПЦР амплификации для верификации SNP методом секвенирования по Сэнгеру



Второй вариант: качественный признак

Задание 7.

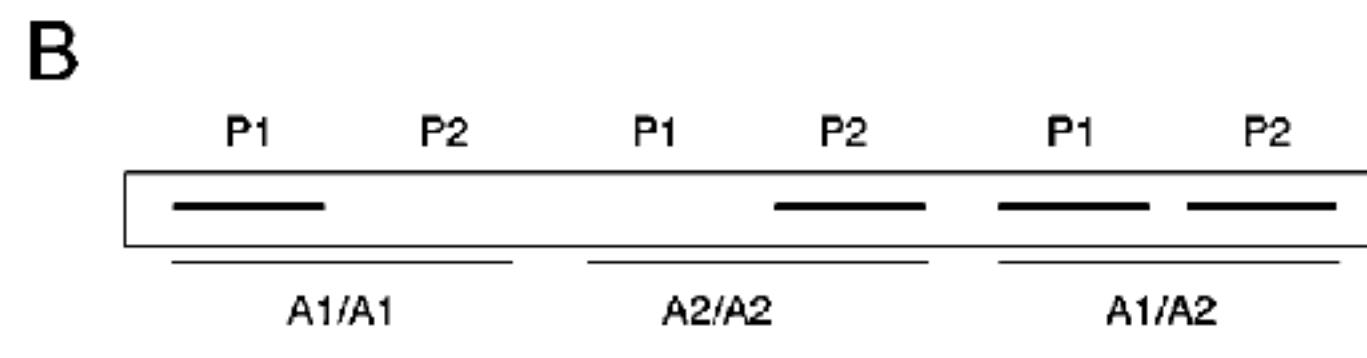
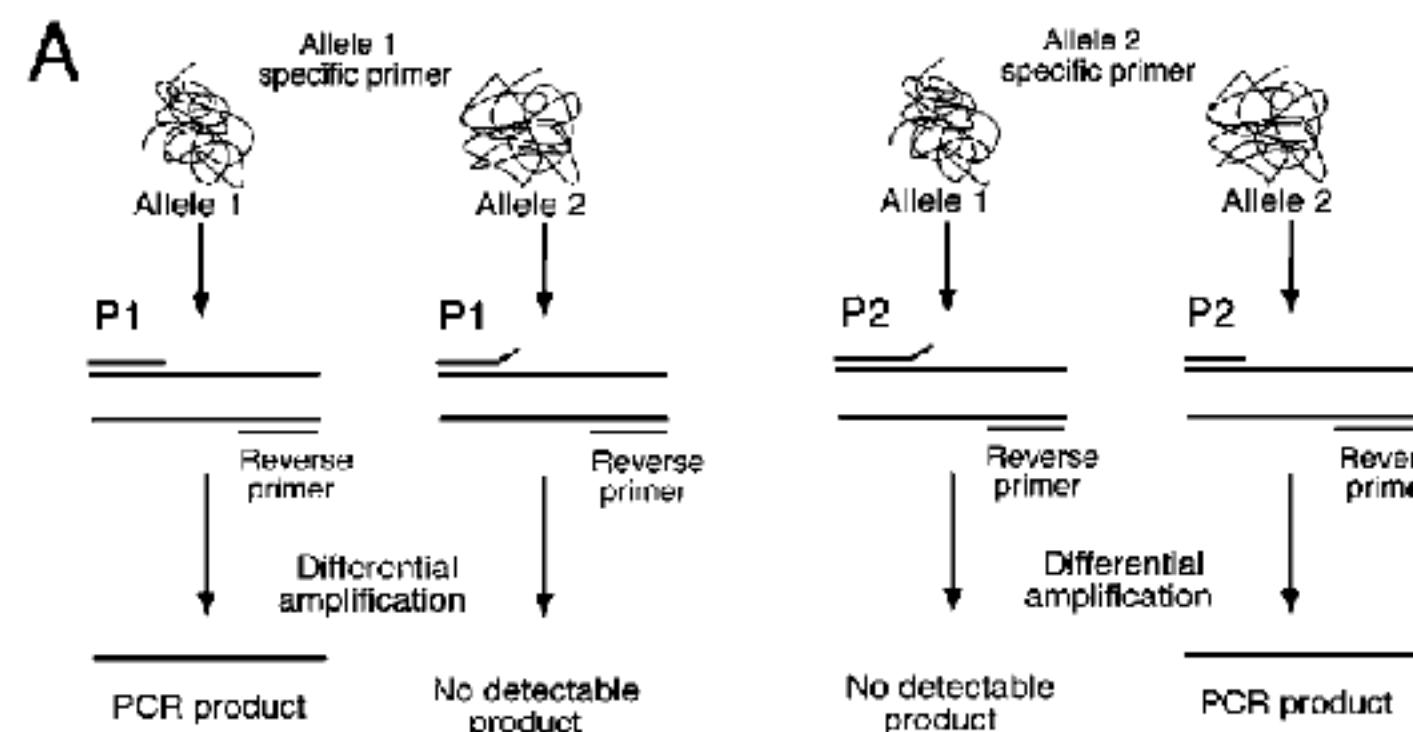
Вы получите результаты секвенирования по Сэнгеру (формат .fas) для 4 SNP (по одному для каждой комбинации популяции (pop-3 и pop-8) и признака (Trh1 - увеличенное содержание бета-токоферола и Trh2 - увеличенное содержание гамма/дельта-токоферолов)), которые были выбраны как наиболее значимые ассоциированные SNP. Вам необходимо заполнить таблицу с полученными в результате секвенирования генотипами для каждого растения в образце и сравнить с предсказанными генотипами для одного из SNP. Сделайте выводы.

Вы можете использовать программное обеспечение Ugene (<http://ugene.net/ru/download-all.html>) или ChromasPro (<https://technelysium.com.au/wp/chromaspro/>) для проведения анализа данных секвенирования.

Названия файлов:

Practicals_Sanger_S1_55196434.zip
Practicals_Sanger_S1_71748138.zip
Practicals_Sanger_S8_23941299.zip
Practicals_Sanger_S8_30578572.zip

Аллель-специфичная ПЦР



SNP types (01-88/02-12)	mismatch types at 3'end of primer in 01-88	Destabilization strength of mismatch type
A/G	TG	Strong
C/T	GT	
G/A	AC	
T/C	CA	
A/C	TC	Weak
G/T	CT	
C/A	AG	
T/G	GA	
A/T	TT	Strong and Medium
T/A	AA	
G/C	CC	Strong and Medium
C/G	GG	

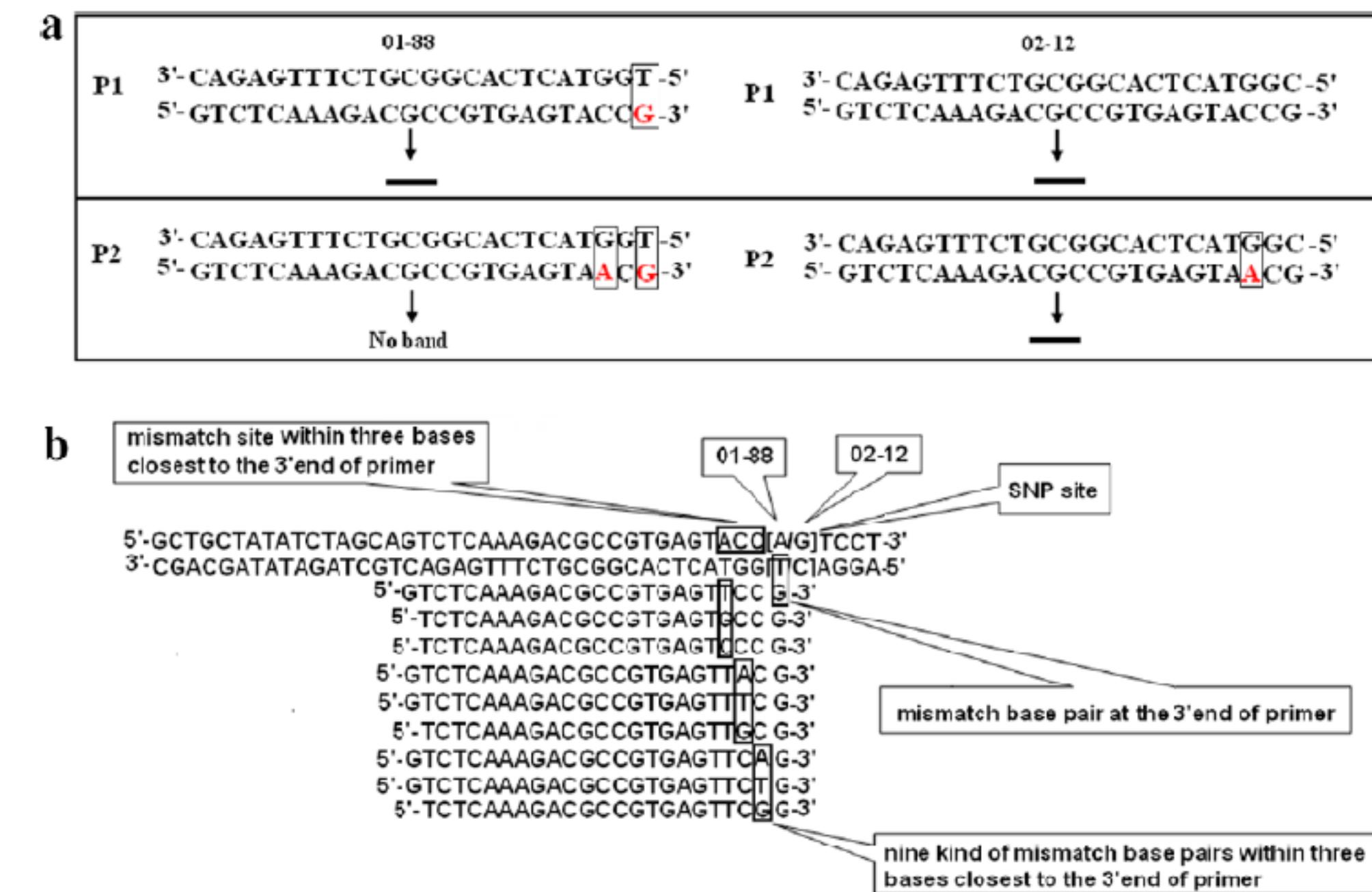


Figure 1 Schematic representation of the AS-PCR primer design. **a**, Primer P1 forms a perfect match with allele from 02-12, but a mismatch base pair at the 3'end with the DNA sequence of allele from 01-88. It could amplify the band in both of two lines 01-88 and 02-12. Primer P2 forms two mismatch base pairs with allele from 01-88 at the 3'end and in the 3rd nucleotide from the 3'end, while a mismatch base pair in the 3rd nucleotide with allele from 02-12. It amplified the band only in 02-12. **b**, Schematic representation of different mismatches during the SNP primer design.

- Liu et al. Plant Methods 2012, 8:34
<http://www.plantmethods.com/content/8/1/34>
- A Simple Procedure for the Analysis of Single Nucleotide Polymorphisms Facilitates Map-Based Cloning in Arabidopsis, Eliana Drenkard
Plant Physiology, December 2000, Vol. 124, pp. 1483–1492,

S8_30578572	Tph2	pop3	115 (93.5%)	13 (9.56%)	18,264,660,16	yes	yes	no	C	T
S1_55196434	Tph1	pop3	119 (92.25%)	7 (5.15%)	20,178,981,51	yes	yes	no	C	T

Программы для подбора АС праймеров

WebSNAPER

[About Websnaper](#) [Documentation](#)

Your email address: s.boldyrev@skoltech.ru

PCR Product Optimum Size Range: 325 - 375

PCR Product Absolute Size Range: 300 - 500

Primer Concentration(nM): 5

Primer Optimal TM: 67 °C

Primer Minimum Size: 20

Primer Maximum TM: 70 °C

Primer Maximum Size: 36

Primer Minimum TM: 62 °C

Mispriming Library: None

Sequence Name: S8_23941299

Please Enter Sequence. The SNP should be indicated as [G/C]:

```
GGCTGCGGATCAGCTATTGTCGTGGGGTCGCCGGCAGGTTCAAGGGACTGGTGC CGTGTTGCCGCAC
GTCCACCAAAACAAGTTCTTGTCTTGTCACTTCCAGTCCCCCATGTTGATATTATTGATGTTATT
ATGTAACCTTATATTTAACATGTGATTTGAAACCAAGTATAAGTATGATTAAAGGATTGT
AAACGAATAAATCAAGTATTGACGTATAGAATGTTAGAGTCACGAATTACGTTAGTTGTCGCAA
GTTAAACAAGAACATATATGAATTAAATGTATGAACACGTATAAATAACAATGTATAATA
ATAATAATTGATTACGATCGAAGTCTCGGATTACAGTGGTTGAAATGAAATATGATTACAAGGGTT
ACAAGTTTCAAAAATAATTACAAGACAAGCTTCTAAATATGAAAGTTACGAAAAGCGAGGCCTT
ACACAAACCGCATCCTT[C/T]GTATCCTGACCGCCCTCACCTCCGTTAACAACTCTAACCAAC
CATCATGGATCCCCATGCTCAGATCGAGAGACTCGAATGATCTATGCAGATCTGACCAACAAATTAT
CAAAAATAACCTGAAACTGCCAGAACCCATGTTTCAAAATTAACAGTGCCTTTCAGATG
AGTAAACTGAAATCGGGGTAGGGTTCCAGGGAAAAAGAAAAAGAGTGGTTATGGTGGTGGTCA
TGGGATTCGGGTAGTCGATGGACAGTGGTGGGAGGTGGCGTGGAGTGGCATATGGTTGGTTG
```

Output Options:

[Include Sequence](#)

[Include Conditions](#)

[Output Tab Separated Values](#)

[Interpreting Results](#)

[Questions or comments?](#)

Submit! RESET

File Edit Tools Run Help

PCR primers design In silico PCR Oligo test Primers list analysis Polymerase cycling assembly Restriction Tools

PCR Primer Design Options: Length range (12-200 nt): 20-30; 30-60; 60-120; 120-200; Quality Int (0-80%): 20-27; 27-30; 30-33; 33-36; 36-40; 40-44; 44-48; 48-52; 52-56; 56-60; 60-64; 64-68; 68-72; 72-76; 76-80; 80-84; 84-88; 88-92; 92-96; 96-100. Special Settings for Primer Design: Default PCR mode; Long distance PCR; Random PCR; Random Chain Reaction (RCR); Inverse PCR; Different enzymes; Multiplex PCR; Polymerase extension chaining (PEC PCR); Inverted PCR; Circular sequencing; Unique PCR; Group-specific PCR. Parameters for PCR Product Analysis: Symmetrically Tm (°C) and GC (valence) for primer pair (%): 2; Len for consecutive combination of our primers: 5; Max difference between PCR products (bp): 0; Max difference between Tm of PCR products (°C): 0.

Sequences: 7/2014 PCR designed primers (probe) list PCR primer design result

1-Sequence (15-3761) : S8_23941299

2-Sequence (15-3761) : S8_23941299

3-Sequence (15-3761) : S8_23941299

4-Sequence (15-3761) : S8_23941299

5-Sequence (15-3761) : S8_23941299

6-Sequence (15-3761) : S8_23941299

7-Sequence (15-3761) : S8_23941299

8-Sequence (15-3761) : S8_23941299

9-Sequence (15-3761) : S8_23941299

10-Sequence (15-3761) : S8_23941299

11-Sequence (15-3761) : S8_23941299

12-Sequence (15-3761) : S8_23941299

13-Sequence (15-3761) : S8_23941299

14-Sequence (15-3761) : S8_23941299

15-Sequence (15-3761) : S8_23941299

16-Sequence (15-3761) : S8_23941299

17-Sequence (15-3761) : S8_23941299

18-Sequence (15-3761) : S8_23941299

19-Sequence (15-3761) : S8_23941299

20-Sequence (15-3761) : S8_23941299

21-Sequence (15-3761) : S8_23941299

22-Sequence (15-3761) : S8_23941299

23-Sequence (15-3761) : S8_23941299

24-Sequence (15-3761) : S8_23941299

25-Sequence (15-3761) : S8_23941299

26-Sequence (15-3761) : S8_23941299

27-Sequence (15-3761) : S8_23941299

28-Sequence (15-3761) : S8_23941299

29-Sequence (15-3761) : S8_23941299

30-Sequence (15-3761) : S8_23941299

31-Sequence (15-3761) : S8_23941299

32-Sequence (15-3761) : S8_23941299

33-Sequence (15-3761) : S8_23941299

34-Sequence (15-3761) : S8_23941299

35-Sequence (15-3761) : S8_23941299

36-Sequence (15-3761) : S8_23941299

37-Sequence (15-3761) : S8_23941299

38-Sequence (15-3761) : S8_23941299

39-Sequence (15-3761) : S8_23941299

40-Sequence (15-3761) : S8_23941299

41-Sequence (15-3761) : S8_23941299

42-Sequence (15-3761) : S8_23941299

43-Sequence (15-3761) : S8_23941299

44-Sequence (15-3761) : S8_23941299

45-Sequence (15-3761) : S8_23941299

46-Sequence (15-3761) : S8_23941299

47-Sequence (15-3761) : S8_23941299

48-Sequence (15-3761) : S8_23941299

49-Sequence (15-3761) : S8_23941299

50-Sequence (15-3761) : S8_23941299

51-Sequence (15-3761) : S8_23941299

52-Sequence (15-3761) : S8_23941299

53-Sequence (15-3761) : S8_23941299

54-Sequence (15-3761) : S8_23941299

55-Sequence (15-3761) : S8_23941299

56-Sequence (15-3761) : S8_23941299

57-Sequence (15-3761) : S8_23941299

58-Sequence (15-3761) : S8_23941299

59-Sequence (15-3761) : S8_23941299

60-Sequence (15-3761) : S8_23941299

61-Sequence (15-3761) : S8_23941299

62-Sequence (15-3761) : S8_23941299

63-Sequence (15-3761) : S8_23941299

64-Sequence (15-3761) : S8_23941299

65-Sequence (15-3761) : S8_23941299

66-Sequence (15-3761) : S8_23941299

67-Sequence (15-3761) : S8_23941299

68-Sequence (15-3761) : S8_23941299

69-Sequence (15-3761) : S8_23941299

70-Sequence (15-3761) : S8_23941299

71-Sequence (15-3761) : S8_23941299

72-Sequence (15-3761) : S8_23941299

73-Sequence (15-3761) : S8_23941299

74-Sequence (15-3761) : S8_23941299

75-Sequence (15-3761) : S8_23941299

76-Sequence (15-3761) : S8_23941299

77-Sequence (15-3761) : S8_23941299

78-Sequence (15-3761) : S8_23941299

79-Sequence (15-3761) : S8_23941299

80-Sequence (15-3761) : S8_23941299

81-Sequence (15-3761) : S8_23941299

82-Sequence (15-3761) : S8_23941299

83-Sequence (15-3761) : S8_23941299

84-Sequence (15-3761) : S8_23941299

85-Sequence (15-3761) : S8_23941299

86-Sequence (15-3761) : S8_23941299

87-Sequence (15-3761) : S8_23941299

88-Sequence (15-3761) : S8_23941299

89-Sequence (15-3761) : S8_23941299

90-Sequence (15-3761) : S8_23941299

91-Sequence (15-3761) : S8_23941299

92-Sequence (15-3761) : S8_23941299

93-Sequence (15-3761) : S8_23941299

94-Sequence (15-3761) : S8_23941299

95-Sequence (15-3761) : S8_23941299

96-Sequence (15-3761) : S8_23941299

97-Sequence (15-3761) : S8_23941299

98-Sequence (15-3761) : S8_23941299

99-Sequence (15-3761) : S8_23941299

100-Sequence (15-3761) : S8_23941299

101-Sequence (15-3761) : S8_23941299

102-Sequence (15-3761) : S8_23941299

103-Sequence (15-3761) : S8_23941299

104-Sequence (15-3761) : S8_23941299

105-Sequence (15-3761) : S8_23941299

106-Sequence (15-3761) : S8_23941299

107-Sequence (15-3761) : S8_23941299

108-Sequence (15-3761) : S8_23941299

109-Sequence (15-3761) : S8_23941299

110-Sequence (15-3761) : S8_23941299

111-Sequence (15-3761) : S8_23941299

112-Sequence (15-3761) : S8_23941299

113-Sequence (15-3761) : S8_23941299

114-Sequence (15-3761) : S8_23941299

115-Sequence (15-3761) : S8_23941299

116-Sequence (15-3761) : S8_23941299

117-Sequence (15-3761) : S8_23941299

118-Sequence (15-3761) : S8_23941299

119-Sequence (15-3761) : S8_23941299

120-Sequence (15-3761) : S8_23941299

121-Sequence (15-3761) : S8_23941299

122-Sequence (15-3761) : S8_23941299

123-Sequence (15-3761) : S8_23941299

124-Sequence (15-3761) : S8_23941299

125-Sequence (15-3761) : S8_23941299

126-Sequence (15-3761) : S8_23941299

127-Sequence (15-3761) : S8_23941299

128-Sequence (15-3761) : S8_23941299

129-Sequence (15-3761) : S8_23941299

130-Sequence (15-3761) : S8_23941299

131-Sequence (15-3761) : S8_23941299

132-Sequence (15-3761) : S8_23941299

133-Sequence (15-3761) : S8_23941299

134-Sequence (15-3761) : S8_23941299

135-Sequence (15-3761) : S8_23941299

136-Sequence (15-3761) : S8_23941299

137-Sequence (15-3761) : S8_23941299

138-Sequence (15-3761) : S8_23941299

139-Sequence (15-3761) : S8_23941299

140-Sequence (15-3761) : S8_23941299

141-Sequence (15-3761) : S8_23941299

142-Sequence (15-3761) : S8_23941299

143-Sequence (15-3761) : S8_23941299

144-Sequence (15-3761) : S8_23941299

145-Sequence (15-3761) : S8_23941299

146-Sequence (15-3761) : S8_23941299

147-Sequence (15-3761) : S8_23941299

148-Sequence (15-3761) : S8_23941299

149-Sequence (15-3761) : S8_23941299

150-Sequence (15-3761) : S8_23941299

151-Sequence (15-3761) : S8_23941299

152-Sequence (15-3761) : S8_23941299

153-Sequence (15-3761) : S8_23941299

154-Sequence (15-3761) : S8_23941299

155-Sequence (15-3761) : S8_23941299

156-Sequence (15-3761) : S8_23941299

157-Sequence (15-3761) : S8_23941299

158-Sequence (15-3761) : S8_23941299

159-Sequence (15-3761) : S8_23941299

160-Sequence (15-3761) : S8_23941299

161-Sequence (15-3761) : S8_23941299

162-Sequence (15-3761) : S8_23941299

163-Sequence (15-3761) : S8_23941299

164-Sequence (15-3761) : S8_23941299

165-Sequence (15-3761) : S8_23941299

166-Sequence (15-3761) : S8_23941299

167-Sequence (15-3761) : S8_23941299

168-Sequence (15-3761) : S8_23941299

169-Sequence (15-3761) : S8_23941299

170-Sequence (15-3761) : S8_23941299

171-Sequence (15-3761) : S8_23941299

172-Sequence (15-3761) : S8_23941299

173-Sequence (15-3761) : S8_23941299

174-Sequence (15-3761) : S8_23941299

175-Sequence (15-3761) : S8_23941299

176-Sequence (15-3761) : S8_23941299

177-Sequence (15-3761) : S8_23941299

178-Sequence (15-3761) : S8_23941299

179-Sequence (15-3761) : S8_23941299

180-Sequence (15-3761) : S8_23941299

181-Sequence (15-3761) : S8_23941299

182-Sequence (15-3761) : S8_23941299

183-Sequence (15-3761) : S8_23941299

184-Sequence (15-3761) : S8_23941299

185-Sequence (15-3761) : S8_23941299

186-Sequence (15-3761) : S8_23941299

187-Sequence (15-3761) : S8_23941299

188-Sequence (15-3761) : S8_23941299

189-Sequence (15-3761) : S8_23941299

190-Sequence (15-3761) : S8_23941299

191-Sequence (15-3761) : S8_23941299

192-Sequence (15-3761) : S8_23941299

193-Sequence (15-3761) : S8_23941299

194-Sequence (15-3761) : S8_23941299

195-Sequence (15-3761) : S8_23941299

196-Sequence (15-3761) : S8_23941299

197-Sequence (15-3761) : S8_23941299

198-Sequence (15-3761) : S8_23941299

199-Sequence (15-3761) : S8_23941299

200-Sequence (15-3761) : S8_23941299

201-Sequence (15-3761) : S8_23941299

202-Sequence (15-3761) : S8_23941299

203-Sequence (15-3761) : S8_23941299

204-Sequence (15-3761) : S8_23941299

205-Sequence (15-3761) : S8_23941299

206-Sequence (15-3761) : S8_23941299

207-Sequence (15-3761) : S8_23941299

208-Sequence (15-3761) : S8_23941299

209-Sequence (15-3761) : S8_23941299

210-Sequence (15-3761) : S8_23941299

211-Sequence (15-3761) : S8_23941299

212-Sequence (15-3761) : S8_23941299

213-Sequence (15-3761) : S8_23941299

214-Sequence (15-3761) : S8_23941299

215-Sequence (15-376

FastPCR

>S8_30578572

- GCGTGTGTTGAACCTCGACCCTTATGACTAC
- TGTTTGAACCTCGACCCTTATGACGCC
- GCGTGTGTTGAACCTCGACCCTTATGACTGT
- CGCGTGTGTTGAACCTCGACCCTTATGACTAT
- AATTGGGTGCTTACGCTGCTAACCGCGTGTGTTGAACCTCGACCCTTATGACTC[C/T]

>S8_23941299

- GAGGCCTTACACAAACCGCATCCTAC
- GCGTTACACAAACCGCATCCGTC
- AGGCGTTACACAAACCGCATCGTT
- AGGCGTTACACAAACCGCATCCATT
- TTCTAAATATGGAAAGTTACGAAAAGCGAGGCGTTACACAAACCGCATCCTT[C/T]

WebSNAPER

>S8_30578572

- TCGGGGCCACATTTGTTATATAAA
- TCAGGGCCCCACATTTGTTATATAACA
- GCGTGTGTTGAACCTCGACCCTTATGACTC[C/T]GGGCCCCACATTTGTTATATAACAT

>S8_23941299

- TTAGTATCCTTGCACCGCCCTCACCTCCG
- TTGGTATCCTTGCACCGCCCTCACCT
- CGTTACACAAACCGCATCCTT[C/T]GTATCCTTGCACCGCCCTCACCTCCGTTAACAA



PCR Primer Design Options

Length range (12-500 nt):	23	30
Tm range (40-80°C):	56	61
dG range (-18..-40 kcal/mol):	-26	-27
Quality limit (30-95%):	75	
Linguistic complexity (30-95%):	80	
3'-end composition (5'-3'): swh ssw wsh www		
Primer 5'-end tail:		

- Alternative amplification (non-specific binding) control
- Overlapping primers
- Primers design with dG priority
- C >> T bisulphite conversion

Special Settings for Primer Design

- Default PCR criteria
- Long distance PCR
- Quantitative PCR
- Xtreme Chain Reaction (XCR)
- Microarray
- Difficult sequence
- Inverted PCR
- Circular sequence
- Unique PCR
- Group-specific PCR

Parameters for PCR Product Analysis

Synchronizing Tm (°C) and dG (kcal/mol) for primer pair (±):
 Limit for compatible combination of pair primers:

Multiplex PCR
 Polymerase extension cloning (OE-PCR)

Minimal difference between PCR products, bp:
 Maximal difference between Ta of PCR products (±°C):

AS-PCR primers sets for: s8_30578572									
PrimerID	Sequence(5'-3')	Length(nt)	Tm (°C)	dG (kcal/mol)	CG (%)	Linguistic_complexity (%)	Primer_Quality (%)	PCR_Fragment_Size (bp)	Topt (°C)
R_2_504-529	ttatataacaaaatgtggggcccgaa	26	58,1	-31,3	42,3	81	56		
R_2_504-531	tgttatataacaaaatgtggggccctga	28	58,2	-32,9	39,3	82	36		
F_2_381-406	accgatcaaaccacttcggccatat	26	60,6	-32,9	46,2	84	84	174	63
R_2_504-529	ttatataacaaaatgtggggcccgaa	26	58,1	-31,3	42,3	81	56		
R_2_504-531	tgttatataacaaaatgtggggccctga	28	58,2	-32,9	39,3	82	36		
F_2_397-419	tgcgccatatgggtgcgttact	23	62,0	-31,1	56,5	90	79	158	63
AS-PCR primers sets for: s8_23941299									
PrimerID	Sequence(5'-3')	Length(nt)	Tm (°C)	dG (kcal/mol)	CG (%)	Linguistic_complexity (%)	Primer_Quality (%)	PCR_Fragment_Size (bp)	Topt (°C)
R_2_504-529	aggtgaggcggtgcaggataccaa	26	65,2	-35,6	57,7	81	81		
R_2_504-532	cggaggtgaggcggtgcaggataactaa	29	66,3	-39,6	58,6	78	78		
F_2_360-384	tacgatcgaagtctcgattacagt	25	57,0	-30,1	44,0	81	81	195	62
R_2_504-529	aggtgaggcggtgcaggataccaa	26	65,2	-35,6	57,7	81	81		
R_2_504-532	cggaggtgaggcggtgcaggataactaa	29	66,3	-39,6	58,6	78	78		
F_2_361-387	acgatcgaagtctcgattacagtgt	27	61,0	-34,2	48,1	86	86	194	66
AS-PCR primers sets for: s8_24016228									
PrimerID	Sequence(5'-3')	Length(nt)	Tm (°C)	dG (kcal/mol)	CG (%)	Linguistic_complexity (%)	Primer_Quality (%)	PCR_Fragment_Size (bp)	Topt (°C)
R_2_504-529	ttgcccccttcttaagcttccatat	26	59,1	-31,7	46,2	77	67		
R_2_504-529	ttgcccccttcttaagcttccatat	26	61,4	-33,2	50,0	79	69		
F_2_456-485	aaaagtagatgaacaagcaatgccttt	30	58,1	-34,6	33,3	87	81	99	62
AS-PCR primers sets for: s8_20164211									
PrimerID	Sequence(5'-3')	Length(nt)	Tm (°C)	dG (kcal/mol)	CG (%)	Linguistic_complexity (%)	Primer_Quality (%)	PCR_Fragment_Size (bp)	Topt (°C)
AS-PCR primers sets for: s1_71748138									
PrimerID	Sequence(5'-3')	Length(nt)	Tm (°C)	dG (kcal/mol)	CG (%)	Linguistic_complexity (%)	Primer_Quality (%)	PCR_Fragment_Size (bp)	Topt (°C)
R_2_506-535	caaaggcggttcttagactaattccatat	30	56,5	-33,5	36,7	91	66		
R_2_506-533	aaggcggttcttagactaattccatat	28	57,4	-32,4	39,3	95	66		
F_2_385-408	gtgggttggaatctcaaattgtgac	24	56,2	-28,8	45,8	80	80	180	61

Your WebSNAPER results - Message (Plain Text)

File Message КонсультантПлюс Tell me what you want to do...

Ignore Delete Reply Reply All Forward More New Meeting w... Move to: ? To Manager Team Email Done Reply & Delete Rules CneNote Move Mark Unread Actions Categorize Follow Up Translate Find Related Select Find Contact Persons Save Assign Recipients Send Message as Task Консультант Плюс Введите текст для поиска: DIRECTUM КонсультантПлюс

WebSNAPER@pga2.mgh.harvard.edu Stepan Boldyrev 11:51

Your WebSNAPER results

Here are the WebSNAPER results for sequence S1_4480460:

Conditions:

```
PRIMER_PICK_ANYWAY=1
PRIMER_MAX_TM=70
PRIMER_OPT_TM=67
PRIMER_DNA_CONC=5
PRIMER_MIN_TM=62
PRIMER_PRODUCT_SIZE_RANGE=325-375 300-500
PRIMER_PAIR_WT_PRODUCT_SIZE_GT=.0
PRIMER_PAIR_WT_PRODUCT_SIZE_LT=.0
PRIMER_MIN_SIZE=20
PRIMER_MAX_SIZE=36
PRIMER_WT_SIZE_LT=0
PRIMER_WT_SIZE_GT=0
PRIMER_EXPLAIN_FLAG=1
PRIMER_SELF_END=5
PRIMER_SELF_ANY=10
```

Sequence: S1_4480460

```
TGTTTCGTATGGTCAGCGACGAAACAGACCTGTTCGTCTGATGGGCTTAGTGTTGGGCCAGTCCG
TTTCGTCGAGCCTAGTTGGTTCTGAGTTCTATGGCCAGCTGCTATAAATAGACAACAGATGTTCTG
TTACAACCTAGAGATAAGAGCATACCCAAGGTCAATTCTGTCTAGACTGTGTTGTATCAGTTGTACTCTA
TCCAGTTAACATTGAGTGTGTTCATGGTAAACCAACTGTGTTACTGCATTGTATGTTGATTGGC
ATAGTTACGGGGATTCCGACCCGTAACATTGTTATAAAACAAGGATTAGGTTGTGATATCGATCCTC
CGATTCGGGACCCACAAAAACACTTATCTAAATTAAATTCTTCAATCTTGAACACAGTTGTCTA
TATTCTTATCTTAAAGATTGAACAAACCTCATCCATTCTGGCCCGTGTAGTGTGTTACTCTCGGA[C/T
]CCCAGCTTCGACTGCAACCTGTACTCATCCCCATATAATCAACAGCTTAATTATTCACATTTTTT
GGTAGTTGGTTACTCAATATCTTGGGAGTACCTATATACAATATCTTCTTACAAACACAAATTITA
GCATACAATTGTCAATTCAAGTTGTGTAAGAATGTATCTTACCGGTCCCTTTAGGTTTTGTAA
TCAGCCACATTGGTTACTATTGAAAATTCTCAAAGTATAAAAAAAATCAAATACTTAAAGA
GACGAAATCAGATTAGGTCAAGGATGGGTCAGGTAAGGATGAGTTCGGGTAAGATGAGATAAGAAAAAA
TTAAAAAAATGAGCATCACACAGGCCCAATGAACAAAGTACTGAATAAAATTCAAACATGAACAA
AGAATGAACCACAAATCTGGAGATTAACTGCTGGCTAAAGAAAACAAACATCTCATATAGACATTAT
```

Primer Id	Primer Sequence	Tm	Primer Length	Product Size	Warnings
S1_4480460_L_REF_1	GGCCCGTGTAGT6TTTACTTCTCGCAC	68.023	28		
S1_4480460_L_REF_1_REVERSE	CATCTTACCCGAACTCACCTTACCTGAA	67.016	31	374	
S1_4480460_L_REF_2	GGCCCGTGTAGT6TTTACTTCTCGGT	67.224	28		
S1_4480460_L_REF_2_REVERSE	CATCTTACCCGAACTCACCTTACCTGAA	67.016	31	374	
S1_4480460_L_REF_3	GCCCGTGTAGT6TTTACTTCTCGGCC	68.033	27		

Аллель-специфичная ПЦР

Результаты верификации молекулярного маркера на повышенное содержание гамма/дельта-токоферола в семенах подсолнечника (tph2, S1_30578572)

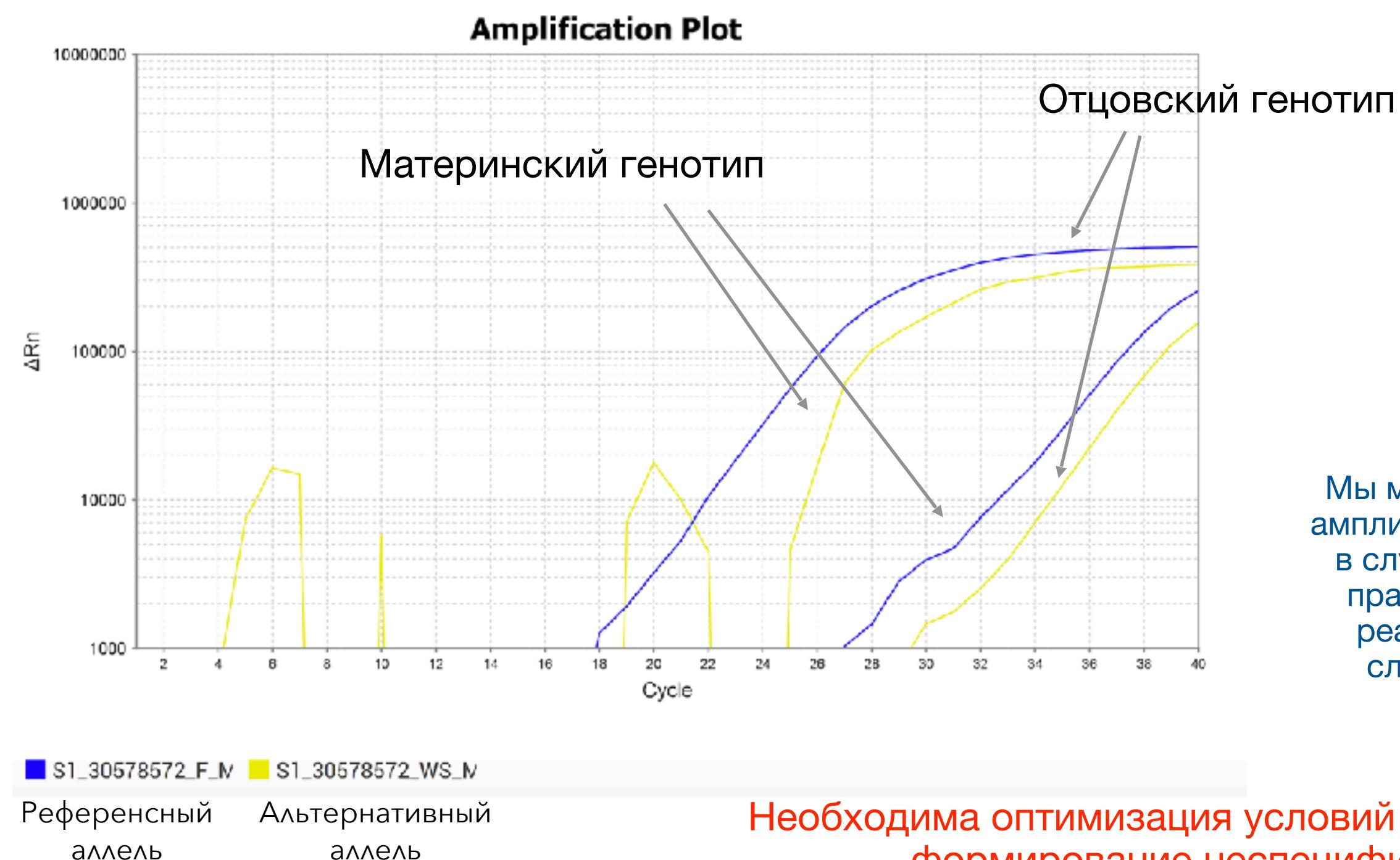
Для данного локуса были использованы два аллель-специфических прямых праймера - праймер, специфичный к референсному аллелю (дикий тип) и праймер, специфичный к альтернативному аллелю (мутант) и универсальный обратный праймер.

В случае наличия у растения двух копий аллея дикого типа амплификация с праймером, специфичным к референсному аллелю приведет к появлению продукта амплификации, а амплификация с праймером, специфичным к мутантому аллелю не даст продукта амплификации.

В случае наличия у растения двух копий мутантного аллея, ситуация будет противоположная, амплификация с праймером, специфичным к референсному аллелю не даст продукта, а амплификация с праймером, специфичным к мутантому аллелю приведет к появлению продукта амплификации. В случае гетерозиготного растения, амплификация с обоими праймерами приведет к появлению продуктов амплификации.

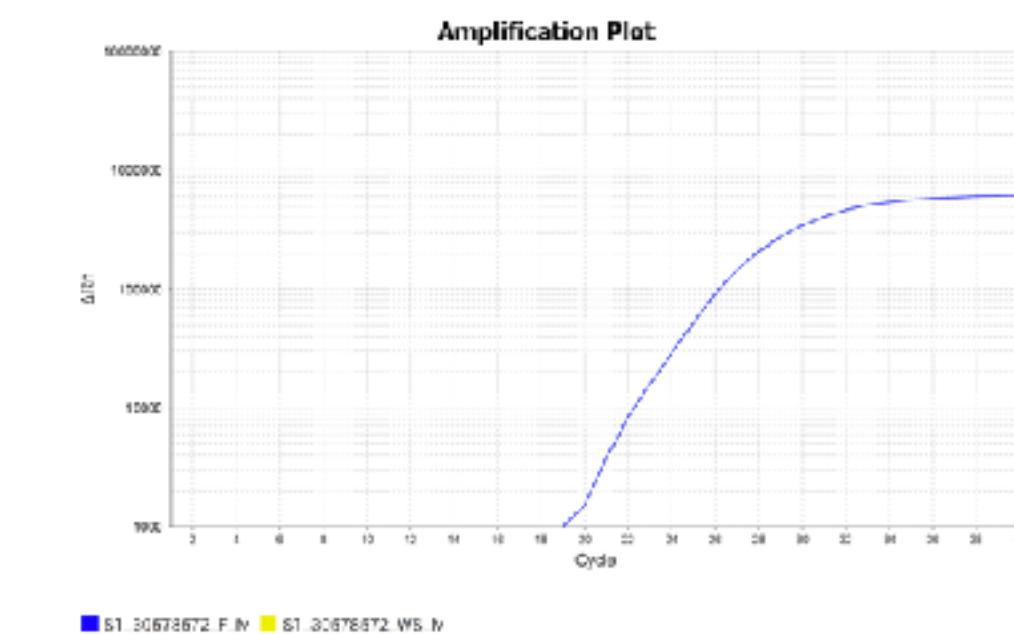
Таким образом, использование сразу двух праймеров делает систему аллель-специфичной ПЦР кодоминантной.

Результаты амплификации родителей - материнский (мутантный генотип) и отцовский (дикий тип)

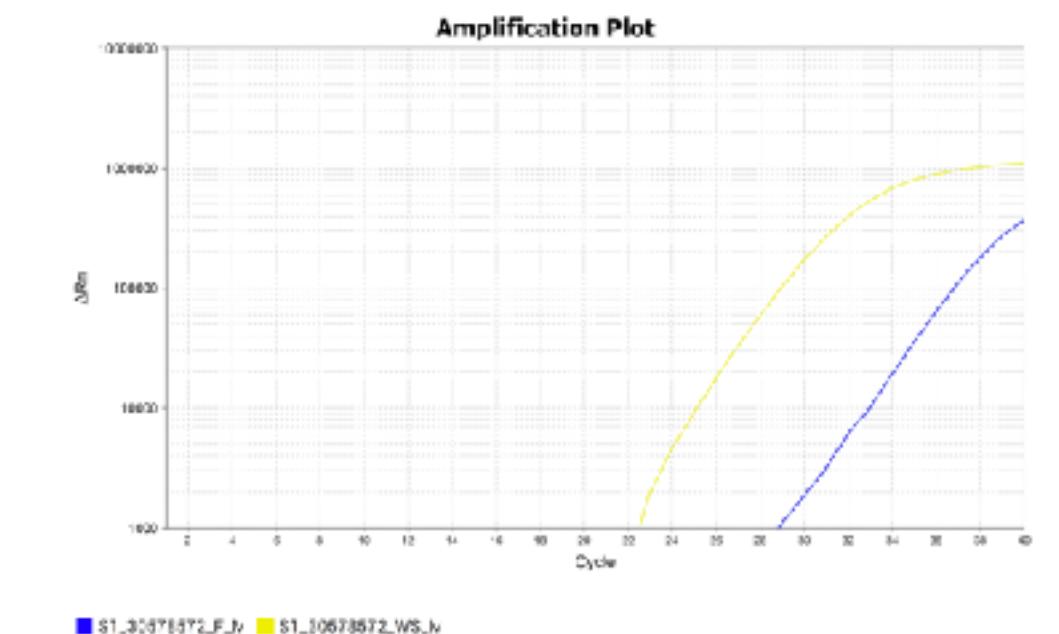


Результаты амплификации растений из картирующей популяции F1

Дикий (отцовский) тип - гомозигота

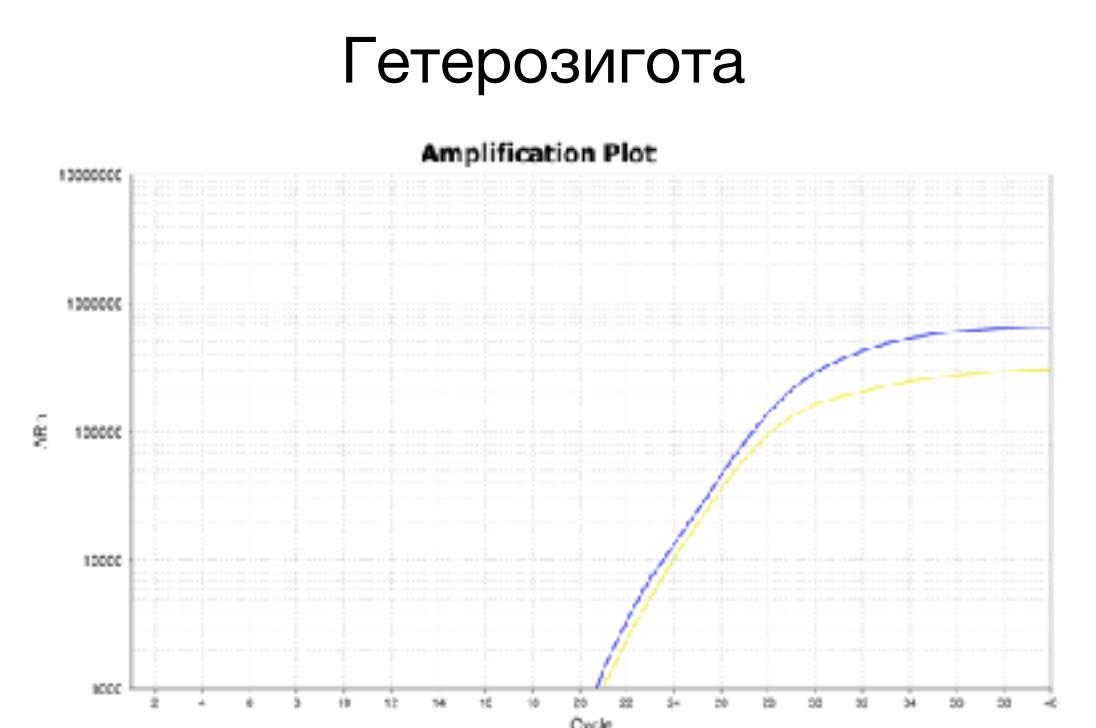


Мутант (материнский тип) - гомозигота



Мы можем видеть наличие продуктов амплификации в обоих случаях, однако в случае наличия неспецифического праймера формирование продукта реакции происходит позже, чем в случае наличия специфического праймера

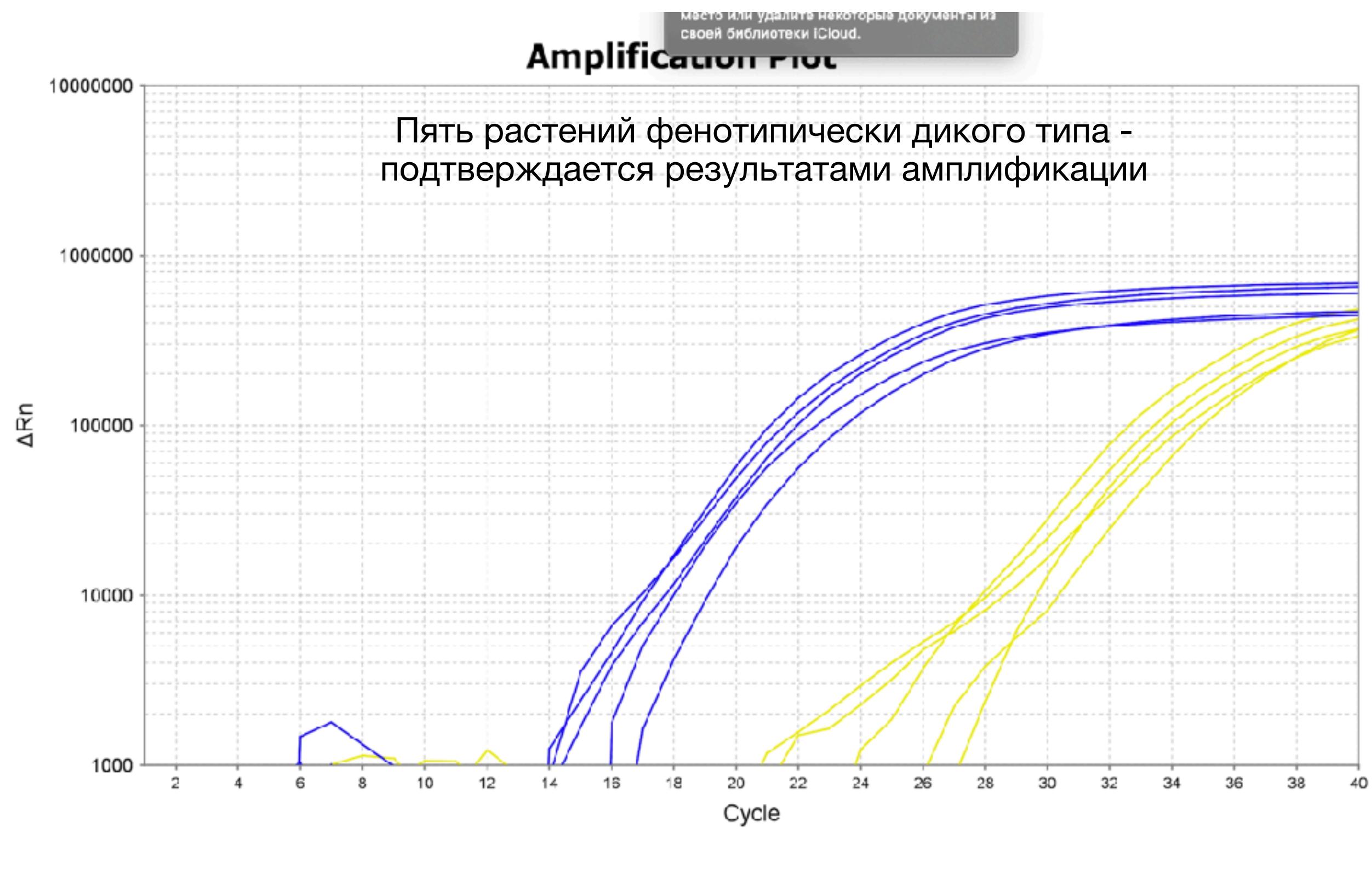
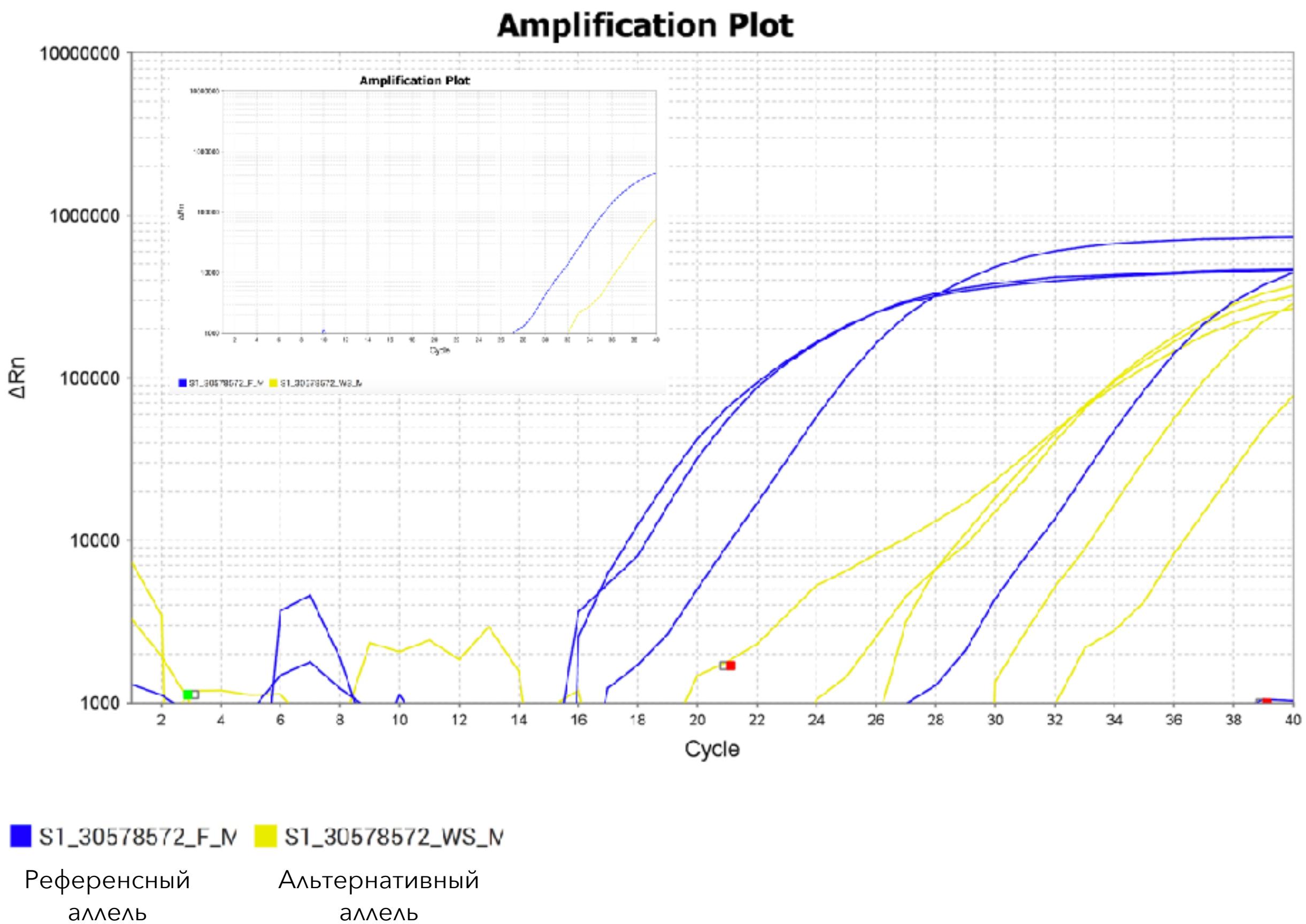
Необходима оптимизация условий ПЦР, чтобы предотвратить формирование неспецифического продукта!



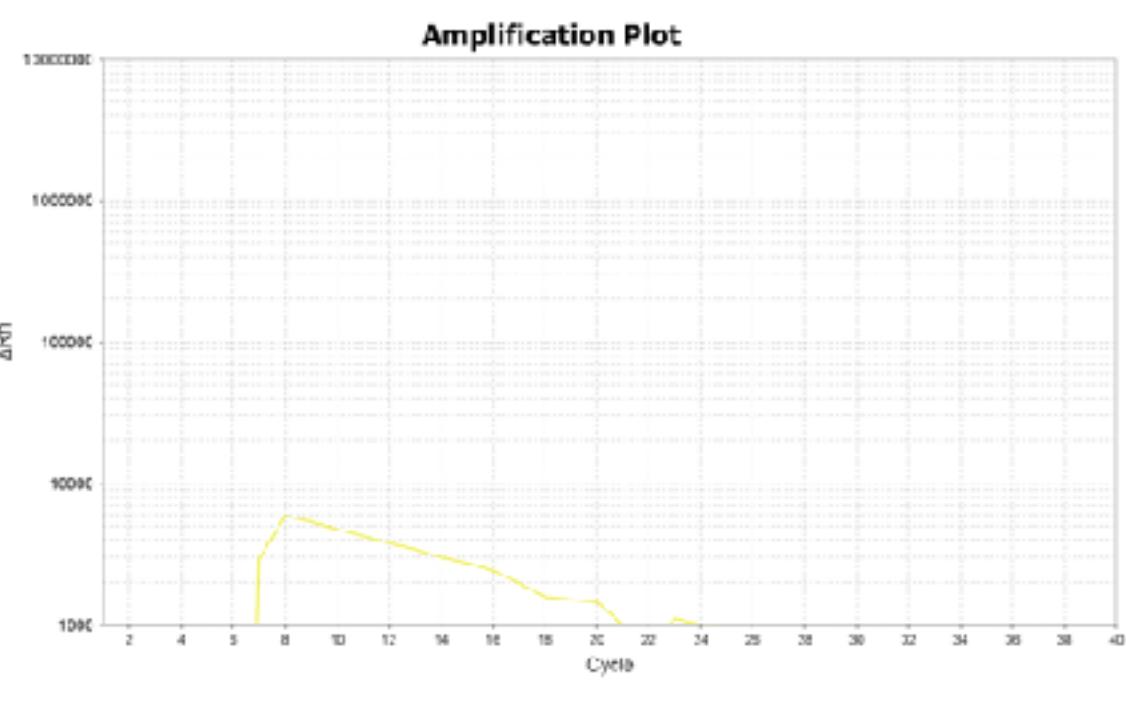
Аллель-специфичная ПЦР

Результаты использования аллель-специфичного маркера (S1_30578572) на независимой выборке линий с известными генотипами

Пять растений фенотипически дикого типа - подтверждается результатами амплификации. Однако амплификация прошла неудачно. Для одного образца не были получены продукты амплификации. Для второго (на врезке) продукты сформировались позднее, чем для других образцов.

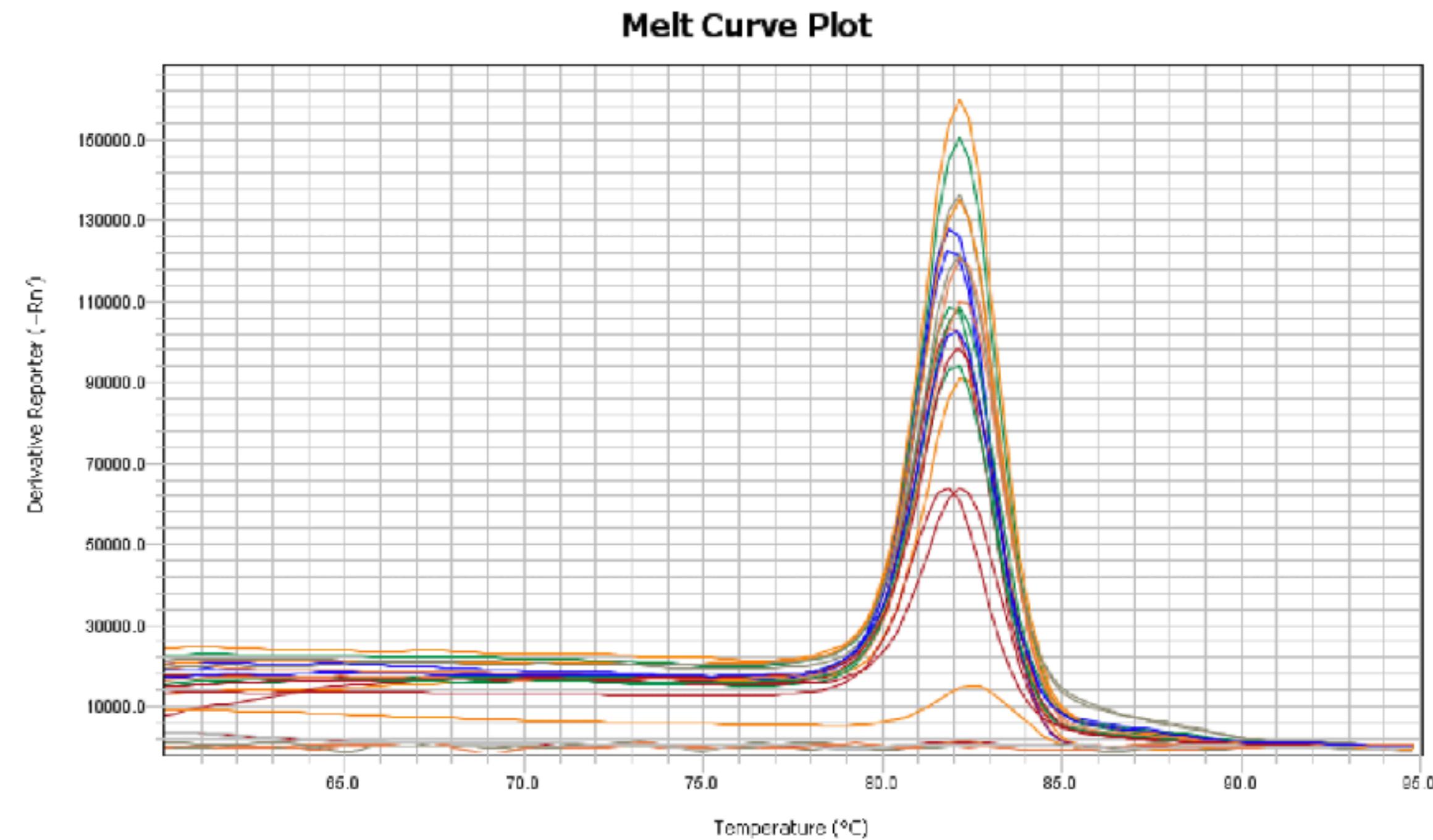
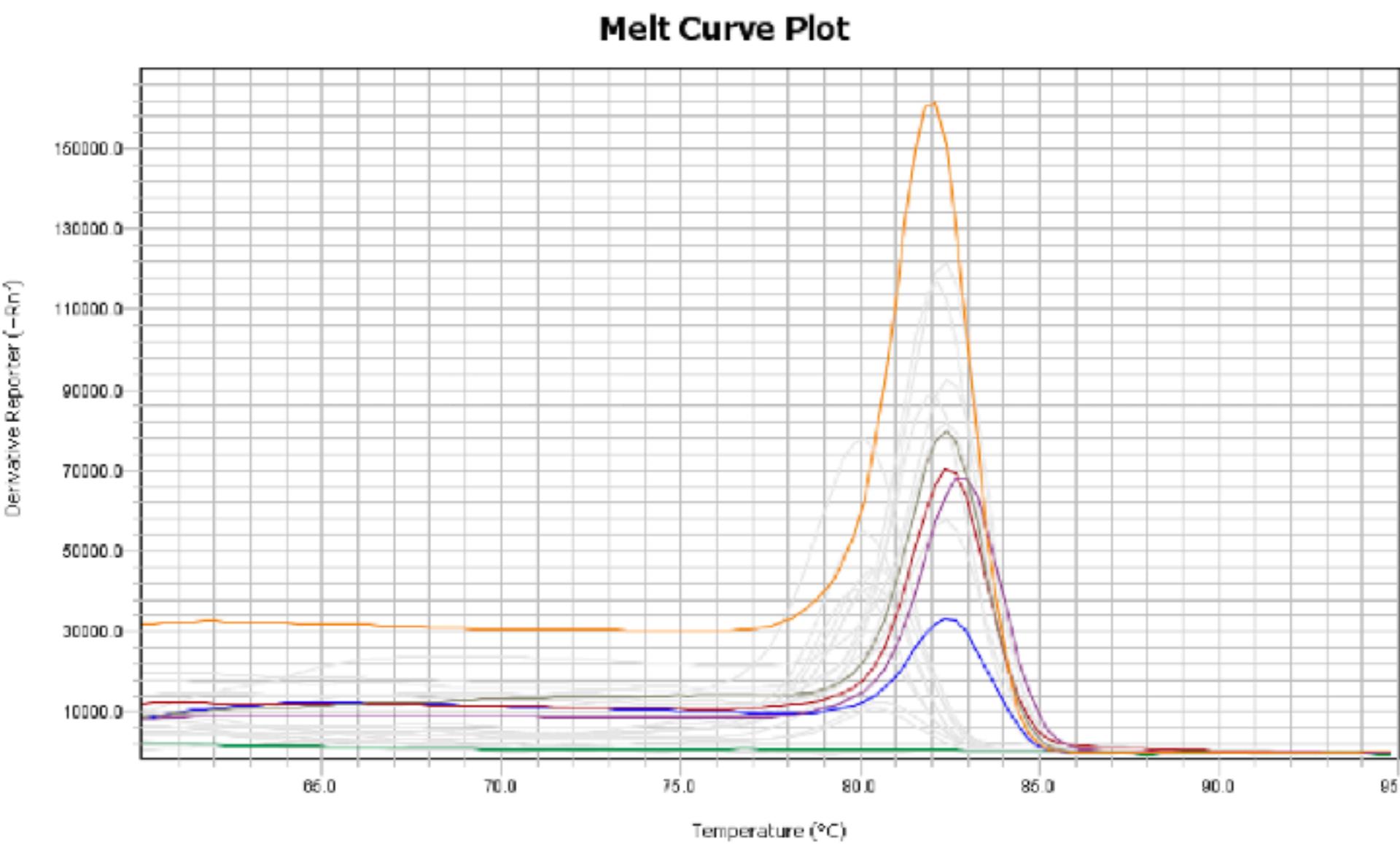
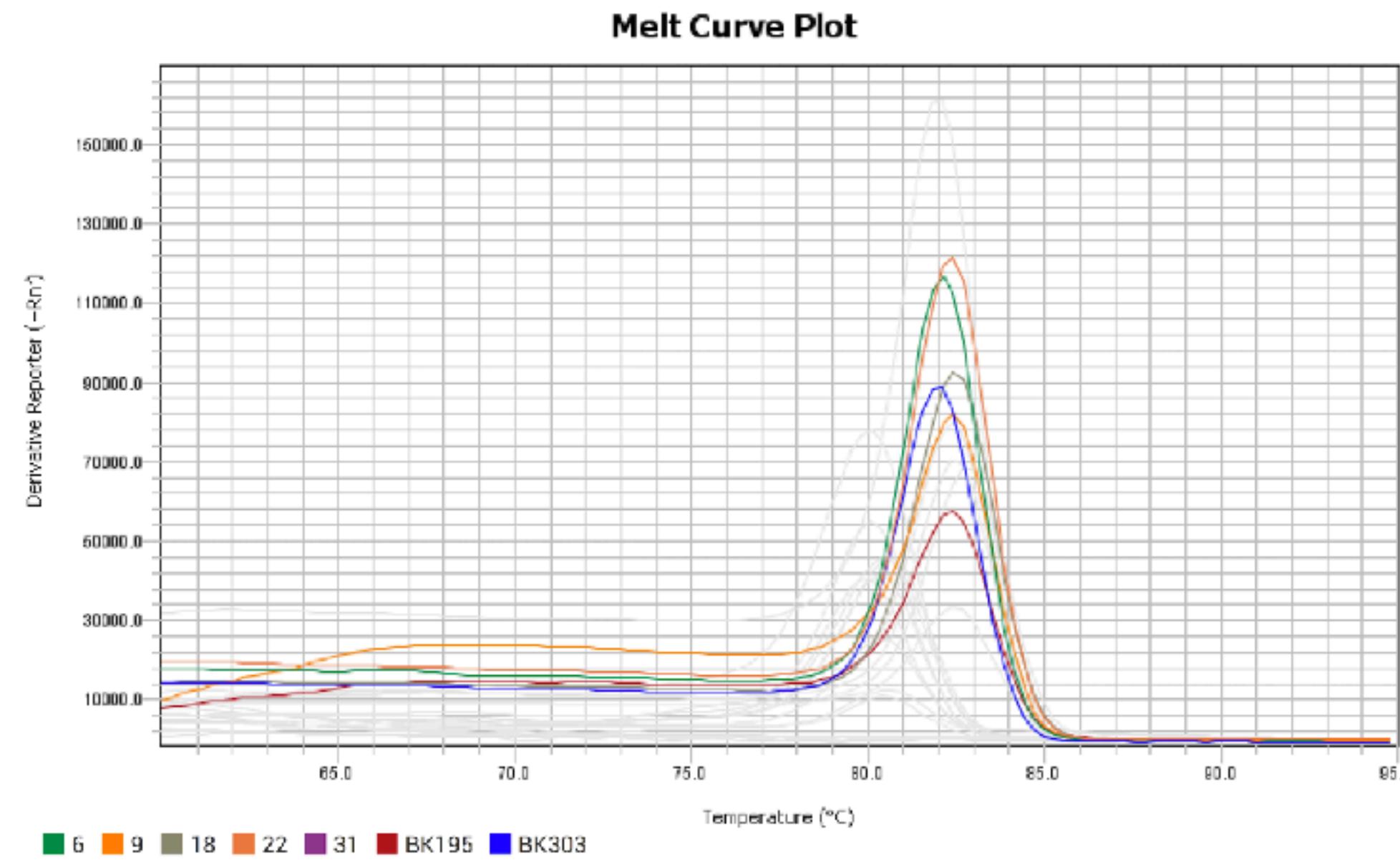


Отрицательный контроль



Аллель-специфичная ПЦР

Оценка специфичности ПЦР - анализ кривых плавления



Все продукты амплификации имеют одинаковый, в пределах допустимого отклонения, пик температуры плавления

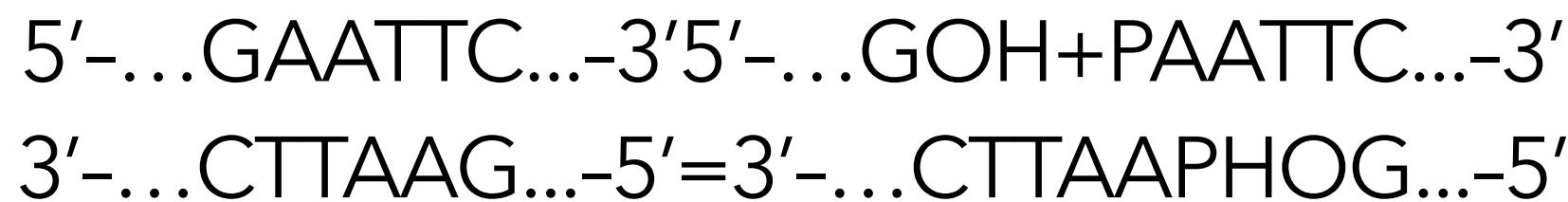
ПЦР прошел специфично!

Вспомним о рестрикции

Эндонуклеазы рестрикции (рестриктазы) широко используются в работах по картированию геномов, клонированию генов, генотипированию и других молекулярно-генетических исследованиях в качестве «молекулярных ножниц». Рестриктазы расщепляют молекулы ДНК в определённых нуклеотидных последовательностях, называемых сайтами узнавания. Эти сайт-специфические эндонуклеазы способны делать разрывы внутри цепи ДНК в отличие от экзонуклеаз, деградирующих ДНК с концов.

Первые рестриктазы были выделены Смитом (Smith, 1970; Smith, Nathans, 1973), само же явление рестрикции-модификации ДНК впервые было обнаружено Лурия (Luria, 1952) при изучении размножения фагов на разных штаммах бактерий.

Рестриктазы II типа узнают палиндромные последовательности нуклеотидов-последовательности, имеющие ось симметрии второго порядка. Так, сайт узнавания рестриктазы EcoRI из шести нуклеотидов-GAATTС-на комплементарной цепи ДНК в том же 5'→3' направлении читается точно так же. Точка разрезания находится между G и первым A, именно там происходит разрыв фосфодиэфирной связи. Образующийся при рестрикции 5'-выступающий «липкий конец» имеет длину 4 нуклеотида:



Эти ферменты, как правило, работают в форме димера, причем каждая субединица белка гидролизует одну цепь ДНК независимо от другой. Однако на одноцепочечной ДНК они работают значительно медленнее. В зависимости от относительного расположения разрыва на комплементарных цепях ДНК получившиеся рестрикты обладают либо 5'-выступающими (EcoRI, BamHI, HindIII), либо 3'-выступающими (PstI) липкими концами. Большинство рестриктаз дают 5'-выступающие липкие концы.

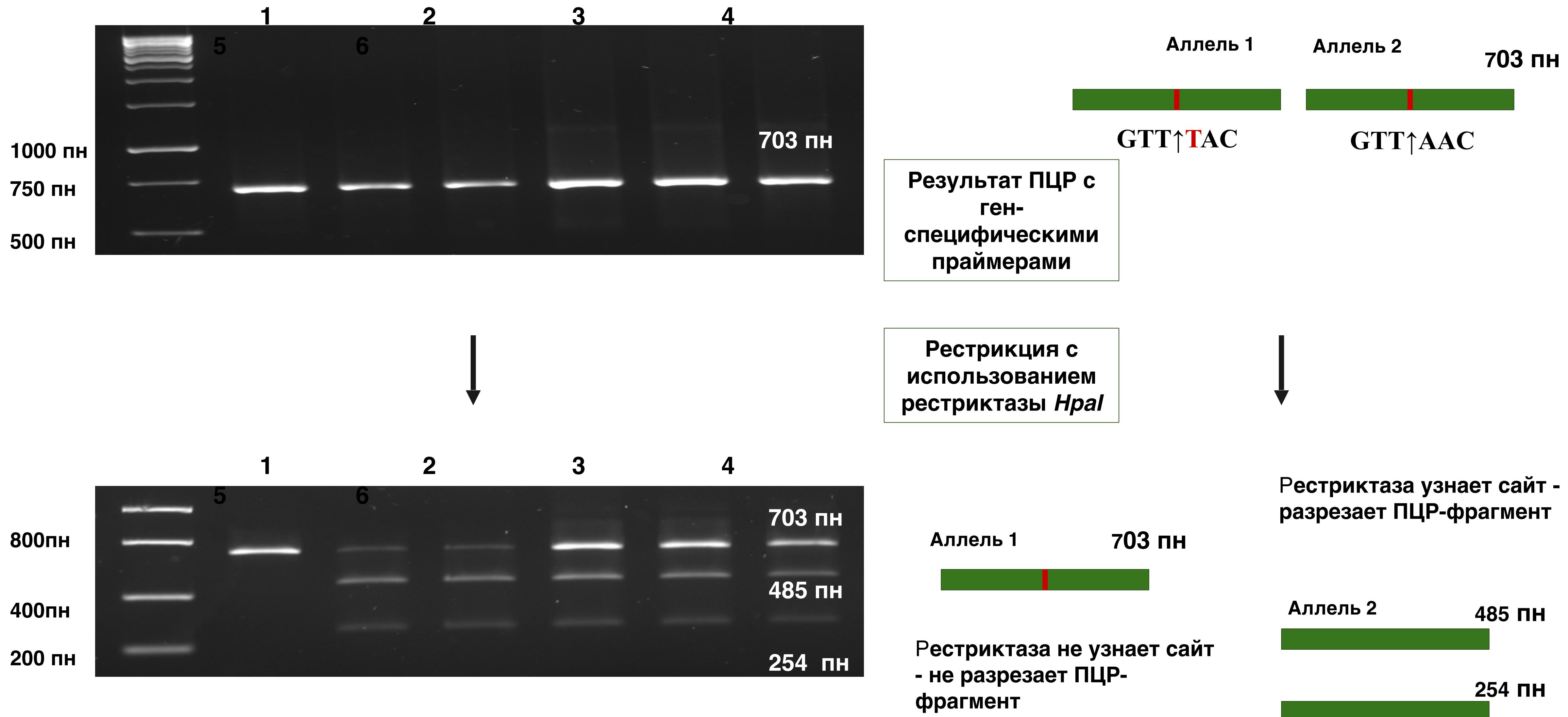
Поговорим о рестрикции

Сайты узнавания некоторых рестрикций

Фермент	Сайт узнавания и точка разрезания (указана стрелкой)	Количество сайтов в ДНК		
		pBR322	M13 mp18	Lambda
BamHI	G [↓] GATCC	1	1	5
BglII	A [↓] GATCT	—	1	6
Sau3A	↑GATC	22	7	16
EcoRI	G [↓] AATTTC	1	1	5
EcoRII	↑CC(A/T)GG	6	7	71
EcoRV	GAT [↓] ATC	1	—	21
HindIII	A [↓] AGCTT	1	1	7
KpnI	G [↓] GTACC	—	1	2
NotI	GC [↓] GGCCGC	—	—	—
PstI	CTGCA [↓] G	1	1	28
SmaI	CCC [↓] GGG	—	1	3
XbaI	C [↓] TCGAG	—	—	1

Для любой ДНК с известной первичной структурой можно выяснить число и локализацию всех сайтов рестрикции с помощью специальных программ для рестрикционного картирования, к примеру, программы **RestrictionMapper**(<http://www.restrictionmapper.org>).

Оценка генетического разнообразия с использованием CAPS маркера - как это работает



Создание CAPS маркера

Этапы создания CAPS-маркера

1. Поиск рестрикционной эндонуклеазы, сайт которой нарушается в результате мутации (либо наоборот, появляется в результате мутации). Исследуемая мутация должна находиться непосредственно в сайте узнавания/разрезания фермента.
2. Проверка адекватности найденной рестриктазы для создания CAPS маркера (размер и количество потенциальных фрагментов рестрикции, стоимость и доступность рестриктазы, возможность легкой детекции различий между аллеем дикого типа и мутантным аллелем). Например, разница в длине между фрагментами не должна быть меньше возможности детекции методом электрофореза (50 нт), минимальная длина фрагмента не менее 100 пн, общее количество фрагментов рестрикции не более 3-4 и т.д.
3. Подбор праймеров и условий для первичной амплификации и визуализации результатов амплификации фрагмента, содержащего полиморфную позицию. Рекомендации для подбора праймеров стандартные. Рекомендуется, чтобы размер фрагмента был ориентировочно не меньше 400 пн и при этом не больше 1000 пн.
4. Тестирование праймеров на специфичность для исследуемого растения. Праймеры должны быть максимально специфичны (идеально -уникальны) для исследуемой культуры, то есть отжигаться только с одним локусом в геноме.
5. Тестирование системы генотипирования с использованием выборки с известными генотипами. Необходимо убедиться, что система работает так, как Вы ожидаете!
5. Верификация маркера на независимой выборке с известными фенотипами.

Создание CAPS маркера

>S8_30578572_Tph2_pop3

```
TTCATCGGAAACACTCTTGTACGCAAAATCAGGTAGCTTGATGCAATACCGTATTGATAGGCTCGTAAACCAGATGTTGGATTCTATGCTTACCGCATATGCACAACACGGTTAGGAGAAAAACGCTTGCATTTGAGAAAATGAA  
GAAAACCGGAATCGAACCCAATTCACTTTTATCGCTTCACCGCTGTAGCCATGGTGGACTCTGGATAAAGGTCAATATTACTTGAGTTAATGAAAAAGTAAACATCGAACCCGACGTTCTCATTACGTAACAATGGTTGACTTACTGGTCGAGCCGGTC  
AACTTGATCACGCCCTAAATTATTAAACGAATTACCGATCAAACCAACTCGGCCATATGGGGTGCCTACTTGGGGCTTGCAGGATGCATAAAATATGGAATTGGTGCTTACGCTGTAACCGTGTGTTGAACTCGACCCTATGACTCCGGCCCCACATTGTTA  
TATAACATCTACCGTCTGGAAAATGGAACGAAGCTCGAAAGTCCGAAAATTATGAAAGAAATCGGTGAAAGAAGAACCTCGGTGAGTGGATATTGAAACTCTGTGATATGACGATAATCACCTATGCAGAACGAAATTAAATAA  
AATGTGGGGTAAGATTAGTGAAGAGATCAAGGAAATCGGGTACGTGCCTGACACGAGTCATGTTGTTATATGTCATGAGCAAGAAAGAGAAGTTACAATTTCATAGTGAAGAAACTGCTTGGCTTGCACTTAAAGACGCCACCGGGTCCACAATTAGGA  
TTAAGAAGAACATTAGAGTTGTGGCATTGCCATTGGCTTTAAGTATGCTCGCTGGTGCAGGAGAAATTATTTGAGGGACACGAATCGGTTCATTTTATCGTGGATCTGTTCTGTGGCGA
```

Программы для поиска сайтов узнавания рестриктаз

Web Cutter 2.0

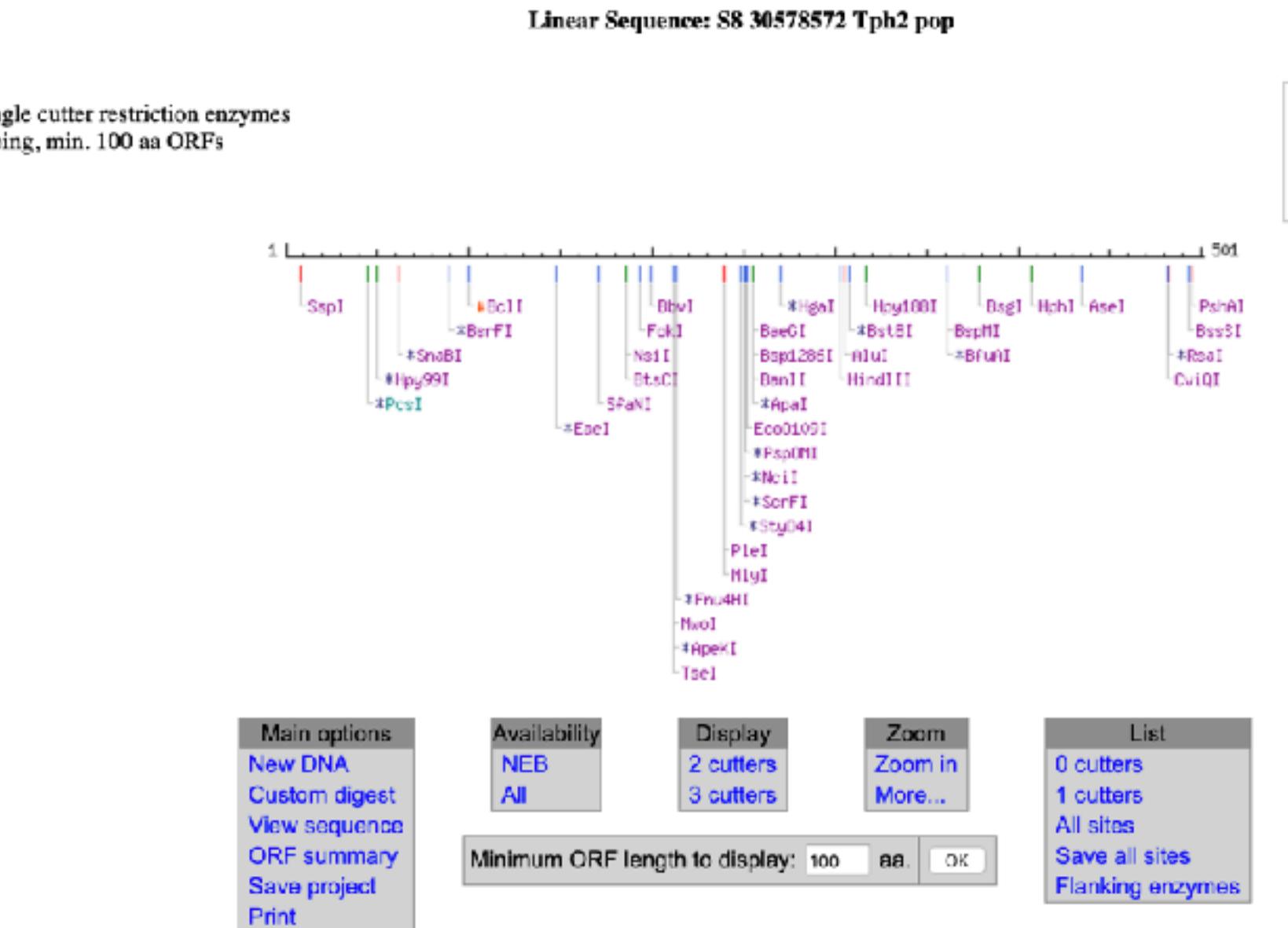
<http://heimanlab.com/cut2.html>

NebCutter2.0

<http://nc2.neb.com/NEBcutter2/>

RestrictionMapper

<http://www.restrictionmapper.org/>



NEBcutter V2.0

Program Guide Help

This tool will take a DNA sequence and find the large, non-overlapping open reading frames using the E.coli genetic code and the sites for all Type II and commercially available Type III restriction enzymes that cut the sequence just once. By default, only enzymes available from NEB are used, but other sets may be chosen. Just enter your sequence and "submit". Further options will appear with the output. The maximum size of the input file is 1 MByte, and the maximum sequence length is 300 KBases.

[What's new in V2.0](#) [Using NEBcutter](#)

Local sequence file: Выбрать файл / файл не выбран [Browse GenBank] Standard sequences: Plasmid vectors Viral + phage

GenBank number: or paste in your DNA sequence: (plain or FASTA format)

Submit

The sequence is: Linear Circular Enzymes to use:

NEB enzymes All commercially available specificities All specificities All + defined oligonucleotide sequences Only defined oligonucleotide sequences [define oligos] Set colors

Minimum ORF length to display: 100 a.a.

Name of sequence: (optional)

Earlier projects:

Note: Your earlier projects will be deleted 2 days after they were last accessed. You need to have cookies enabled in your browser for this feature to work. Delete projects Disable NEBcutter cookies

Создание CAPS маркера

Задание 8.

Используйте полученные ранее файлы, содержащие нуклеотидные последовательности фланирующие изучаемые SNP (501 нт или 1001 нт). С помощью программ WebCutter или NEBCutter исследуйте данные последовательности и выявите наличие рестрикционных эндонуклеаз, сайты узнавания которых содержат исследуемый SNP.

В случае обнаружения подходящей рестриктазы, оцените наличие дополнительных сайтов узнавания вблизи SNP, количество и длину фрагментов, получаемых при обработке рестриктазой исследуемого фрагмента последовательности ДНК подсолнечника.

Подберите оптимальную пару праймеров, фланкирующих участок, содержащий SNP. Проведите пробную амплификацию и рестрикцию «на бумаге», чтобы оценить удобство визуализации результатов генотипирования и возможность дифференциации между двумя альтернативными аллелями.

Создание CAPS маркера

>S8_30578572_Tph2_pop3

TTCATCGGCAACAACTCTTGCATGTACGCAAAATCAGGTAGCTTGATGCAATACCGTATTGATGGCTCGTAAACCAGATGTTGTTGTGAAATTCTATGCTTACCGCATATGCACAACAC
GGGTTAGGAGAAAAACGCTTGATCATTGAGAAAATGAAGAAACCGGAATCGAACCCATTACAGTAACCTTTATGCCTCACCCTGTAGCCATGGTGGACTCTGGATAAAGGTCAATATTAC
TTTGAGTTAACGAAACATCGAACCCGACGTTCTCATTACGTAACAATGGTTGACTTACTGGTCGAGCCGGTCAACTGATCACGCCCTAAATTATTAAACGAATTACGATCAAACCAACT
TCGGCCATATGGGTGCGTTACTTGGGCTTGCAGGATGCATAAAATATGGAATTGGGTGCTTACGCTGCTGAACCGTGTGAACTCGACCCCTATGACTCCGGGGCCCCACATTGTTATATAACATC
TACCGCTGCTGGAAAATGGAACGAAGCTCGAAAGTCCGAAAATTATGAAAGAAATCGGTGTAAAGAAAGAACCTGCGTGCAGTTGGGTGGATATTGAAAACCTGTGCATATGTTGGCTAATGAC
GATAATCACCCATGCAGAACGAAATTAAATAAAATGTGGGTAAGATTAGTGAAGAGATCAAGGAAATCGGTACGTGCCTGACACGAGTCATGTTGTTATGTCGATGAGCAAGAAAGAGAAGTGAAG
TTACAATTTCATAGTGAACAAACTGCTTGGCTTGCACTTAAAGACGCCCCGGGGTCCACAATTAGGATTAAGAAGAACATTAGAGTTGTGGCGATTGCCATTGGCTTTAAGTATGCTCGCTG
GTGGTCGAGCGAGAAATTATTTGAGGGACACGAATCGGTTCATCTTATCGGGATCTGTTCTGGCGA

Подбор ПЦР праймеров
Размер ампликона - 636 пн

Оптимальная рестриктаза

Количество сайтов в ампликоне
MspI (2)

Сайт узнавания (подчеркнут)
c/cgg

SNP аллель дикого типа - C (CCGG) узнается ферментом 3 фрагмента (127 + 159 + 350)

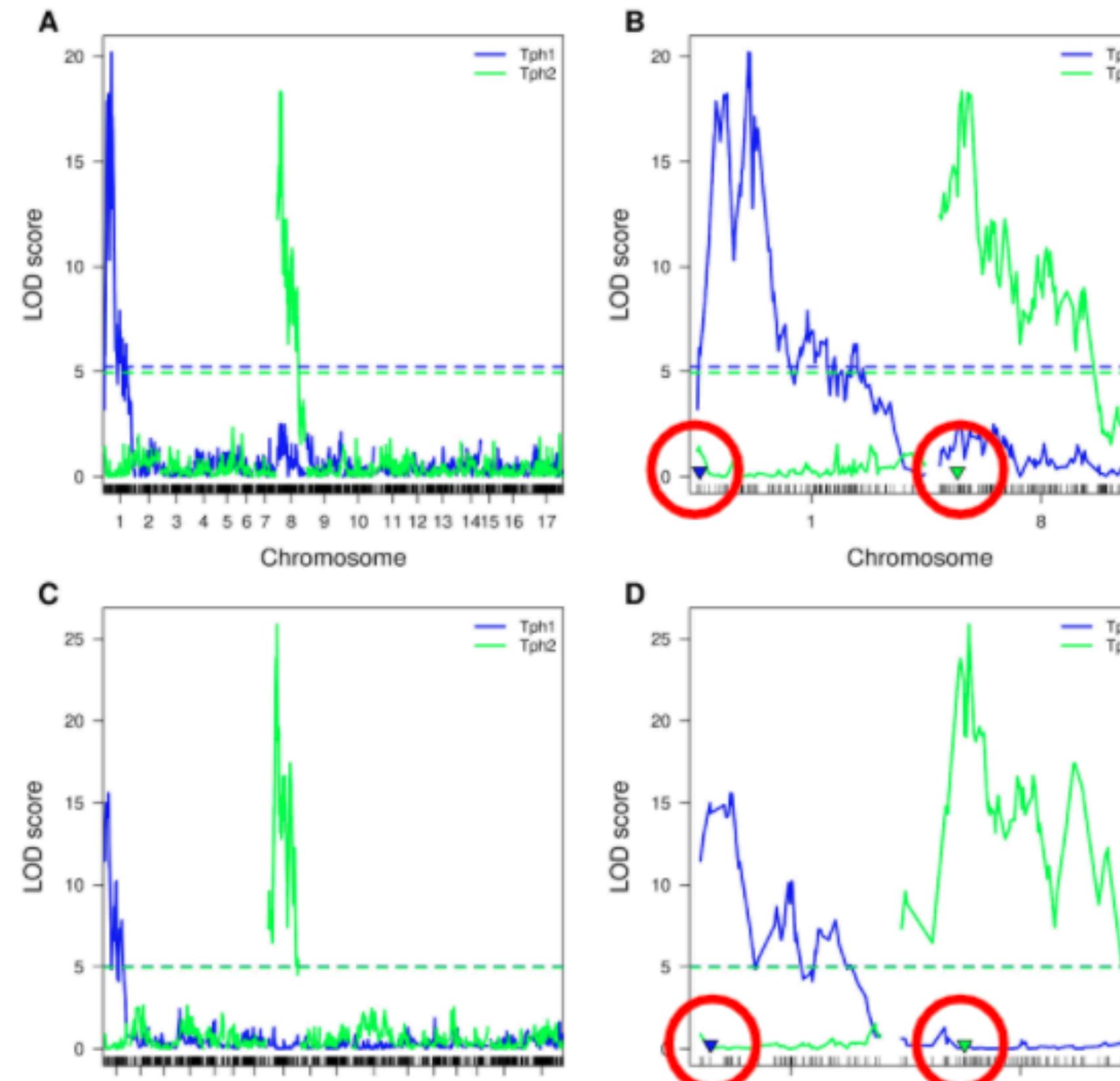
Мутантный SNP аллель - T (сайт становится CTGG) не узнается 2 фрагмента (127 + 509)

Создание CAPS маркера



Тестирование маркеров

Оценка предсказательной силы маркера



Tocopherol composition:

Wild-type - 5 lines

Alpha-beta type - 5 lines

Gamma type - 5 lines

Gamma-delta type - 5 lines

Пять независимых растений с известными генотипами были
отобраны для каждой линии

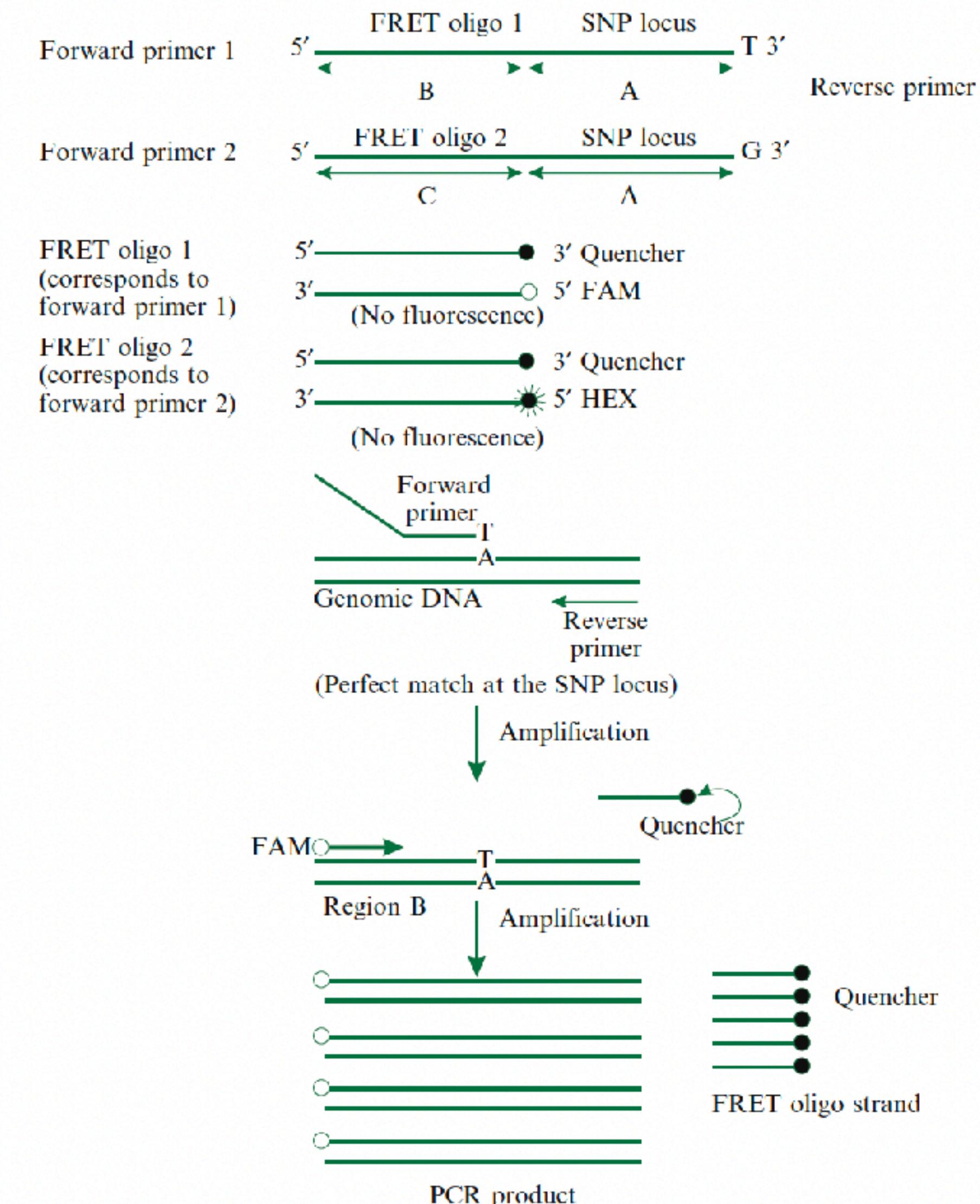
Всего 100 растений

Новые методы генотипирования

KASPTM - KASPar technology (by LCG Genomics)

Конкурентная аллель-специфическая ПЦР

Один из вариантов аллель-специфической ПЦР позволяет различать два аллеля благодаря конкурентному связыванию одного из двух аллель-специфических праймеров. Если генотип для данного SNP гомозиготный, генерируется только один из двух возможных флуоресцентных сигналов. Если генотип гетерозиготный, генерируется смешанный флуоресцентный сигнал.



<https://www.biostrechtech.com/how-does-kasp-work>

<https://www.youtube.com/watch?v=Uq9HhmzOqUQ&t=546s>

KASPTM-KASPar technology (by LCG Genomics)



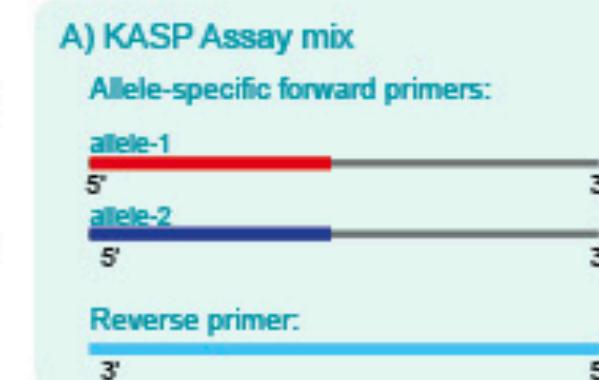
- 1 Array DNA samples into the reaction plate
- 2 Prepare the genotyping mix
- 3 Dispense genotyping mix onto the reaction plate
- 4 Seal and centrifuge the plate
- 5 Run the thermal cycle
- 6 Read the plate and analyse the data

Ключевые компоненты анализа:

1. два немеченых прямых праймера, 3'-конец одного из которых комплементарен SNP аллелю 1, а 3'-конец другого праймера комплементарен альтернативному SNP аллелю 2;
2. общий немеченый обратный праймер (универсальная FRET-кассета);
3. два FRET-олигонуклеотида (FRET-кассеты), одна из которых мечена на 5'-конце одной цепи флуоресцентным красителем FAM (синий), а другая - красителем HEX (зеленый). Комплémentарные цепи обоих олигонуклеотидов несут гаситель флуоресценции.
4. KASP Master Mix содержит пассивный референсный краситель ROXTM, тау-полимеразу, свободные нуклеотиды и MgCl2 в буферном растворе.

1) Assay components:

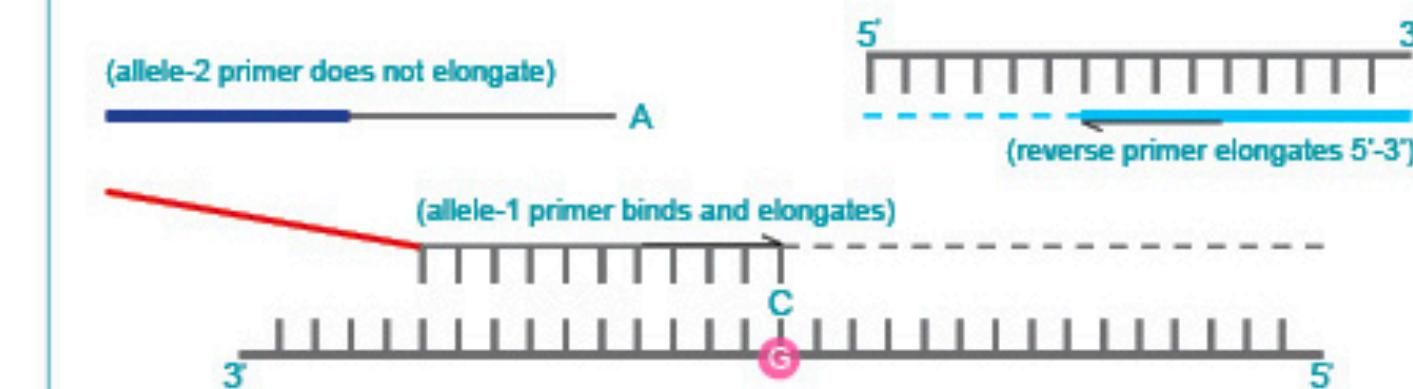
KASP uses three components: test DNA with the SNP of interest; KASP Assay mix containing two different, allele-specific, competing forward primers with unique tail sequences and one reverse primer; the KASP Master mix containing FRET cassette plus Taq polymerase in an optimised buffer solution.



C) DNA template (sample)

5' ... GAT ... 3'

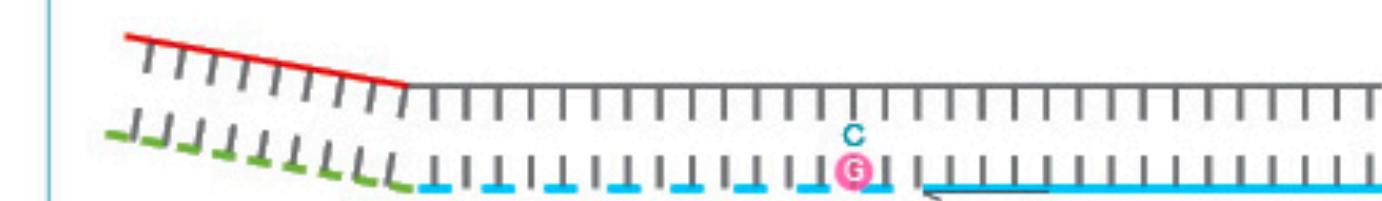
2) Denatured template and annealing components – PCR round 1:



In the first round of PCR, one of the allele-specific primers matches the target SNP and, with the common reverse primer, amplifies the target region.

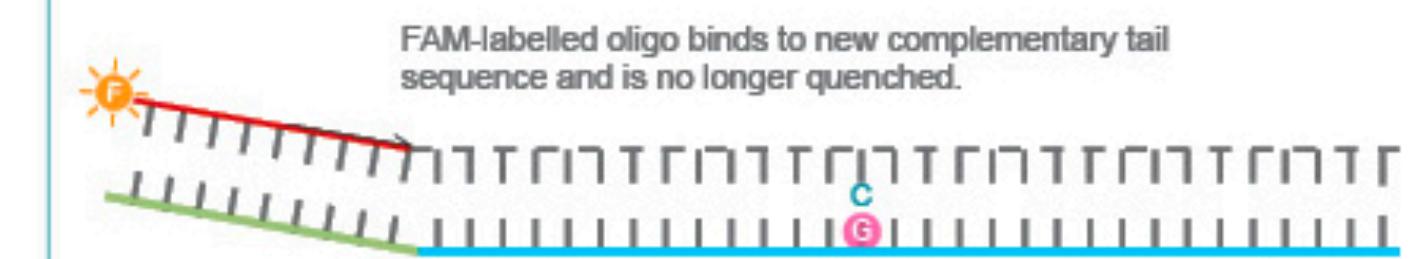
Legend
Allele-1 tail FAM-labelled oligo sequence
Allele-2 tail HEX-labelled oligo sequence
Common reverse primer
FAM dye
HEX dye
Target SNP
Quencher

3) Complement of allele-specific tail sequence generated – PCR round 2:



(Reverse primer binds, elongates and makes a complementary copy of the allele-1 tail.)

4) Signal generation – PCR round 3:

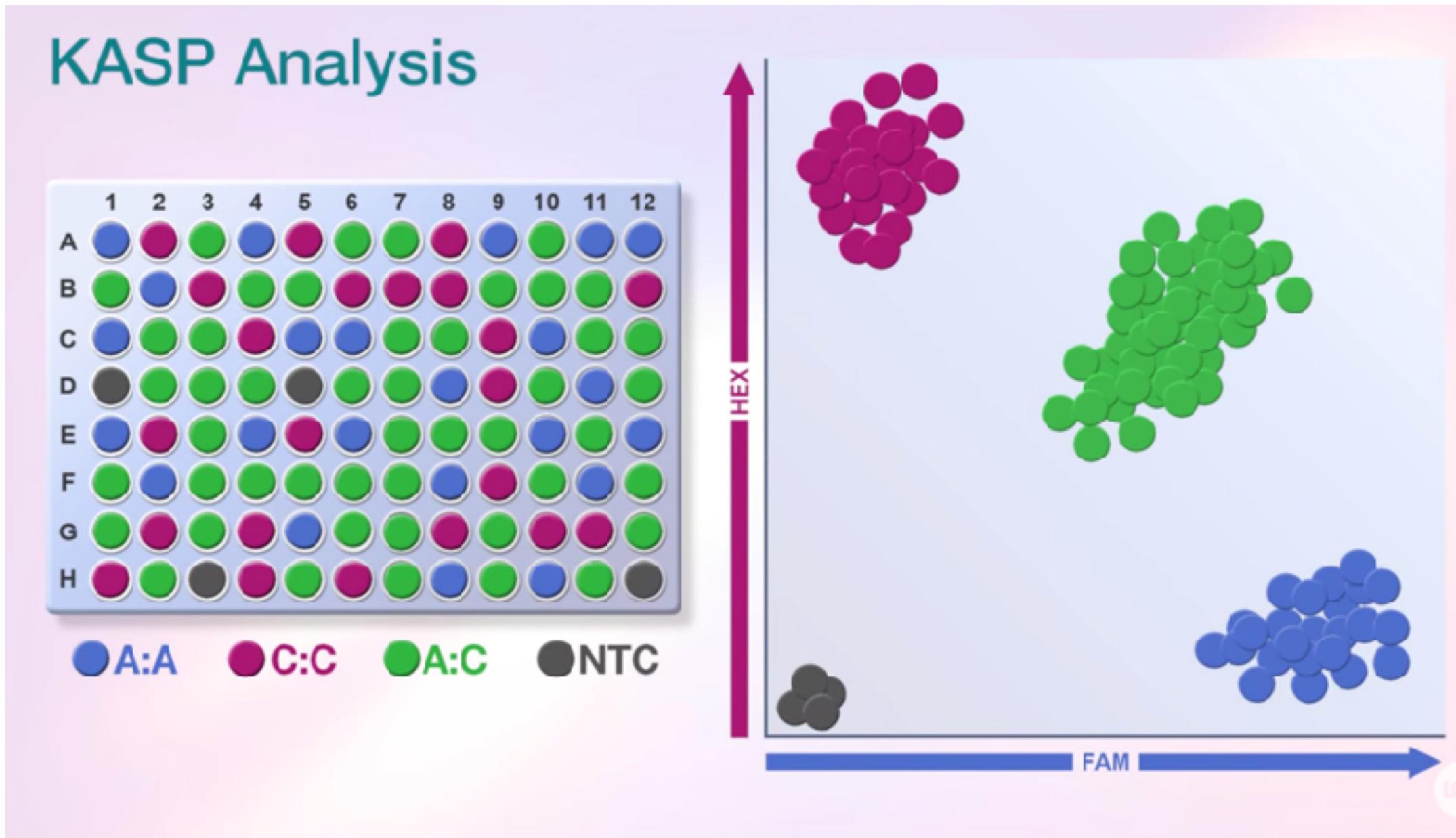


In further rounds of PCR, levels of allele-specific tail increase. The fluor labelled part of the FRET cassette is complementary to new tail sequences and binds, releasing the fluor from the quencher to generate a fluorescent signal.

Fluor for incorporated G allele no longer quenched.	Fluor for non-incorporated T allele remains quenched.
---	---

KASPTM- KASPar technology (by LCG Genomics)

Анализ результатов KASP с помощью кластерного анализа



- Каждая точка данных на кластерном графике представляет собой индивидуальный образец ДНК
- Образцы с одинаковым генотипом группируются вместе

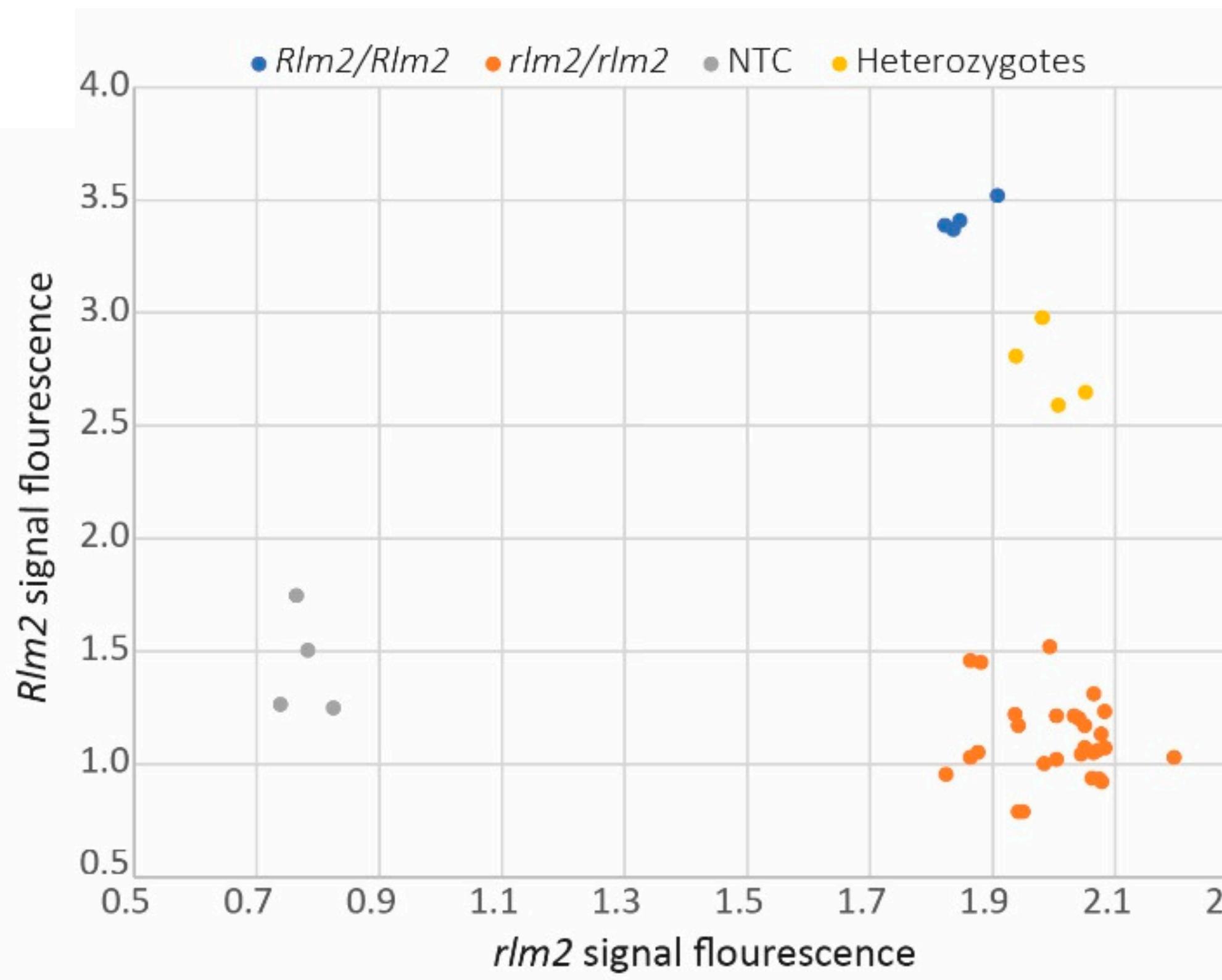
Красный цвет	Образец гомозиготен по аллелю, меченному HEX-меткой
Зеленый цвет	Образец гетерозиготен: один аллель FAM и один аллель HEX
Синий	Образец гомозиготен по аллелю, меченному FAM
Черный	Отрицательный контроль

KASP это анализ по конечной точке, что отличает его от реал-тайм ПЦР

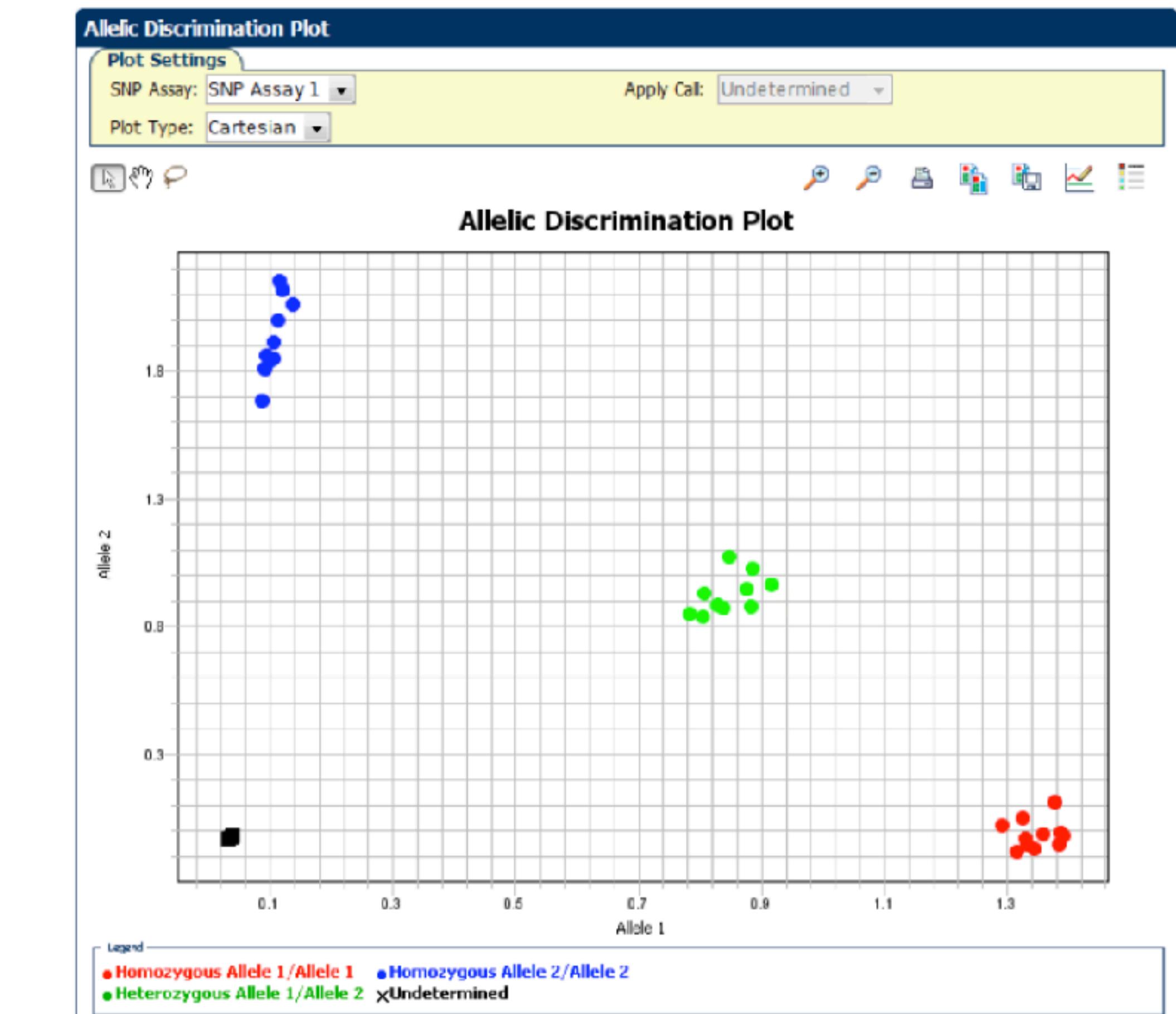
KASP™ - KASPar technology (by LCG Genomics)

Анализ результатов KASP с помощью кластерного анализа

Real data



Ideal data from QuantStudio5



KASPTM- KASPar technology (by LCG Genomics)

Преимущества технологии KASP

- Удобная и простая в практической реализации;
- Возможность масштабирования (минимум до 96 или до 384 образцов за один цикл прибора) и автоматизации;
- Легко модулируемая система - как в случае обычной ПЦР существует возможность создавать системы генотипирования под любой SNP-локус интереса. Необходимо только подобрать праймеры как для обычной аллель-специфичной ПЦР и пометить флуоресцентной меткой;
- Легко интерпретировать результат;
- Высокая точность, специфичность, чувствительность.

Недостатки технологии KASP

- Относительно высокая стоимость реактивов (KASP MasterMix), необходимость заказывать реактивы у единственной компании-производителя (США);
- Зависимость результата от качества подбора праймеров (специфичность ПЦР);
- Для получения достоверного результата необходимо использовать выборки с не менее 20 растениями;
- Необходима оптимизация системы визуализации результата - SNP calling'a.

Новые методы генотипирования

HRM - High Resolution Melting (анализ кривых плавления)

Данная технология основана на двух методических подходах: real-time ПЦР для амплификации участка, содержащего полиморфизм, и высокоточного анализа кривых плавления полученных ПЦР-продуктов для определения генотипа.

Приборы для проведения HRM

HRM был разработан учеными из Университета Юты совместно с сотрудниками компании Idaho Technology (США) и представлен ученому сообществу в 2003 г

Оригинальная публикация Wittwer CT, Reed GH, Gundry CN et al (2003) High-resolution genotyping by amplicon melting analysis using LCGreen. Clin Chem 49:853-860.

HRM может использоваться для:

- детектирования известных полиморфизмов (генотипирования);
- поиска новых полиморфизмов (gene-scanning).

С помощью HRM можно генотипировать:

- SNP полиморфизмы;
- полиморфизм длин фрагментов - SSR;
- инсерционно-делеционаные полиморфизмы.

Анализ кривых плавления позволяет с высокой точностью дифференцировать гомозиготы разных типов и гетерозиготы.

Преимущества HRM

- Быстрая и возможность частичной автоматизации;
- Высокая чувствительность и точность генотипирования;
- Простота реализации;
- Относительная дешевизна;
- Так называемый closed-tube метод



HR-1



LightCycler

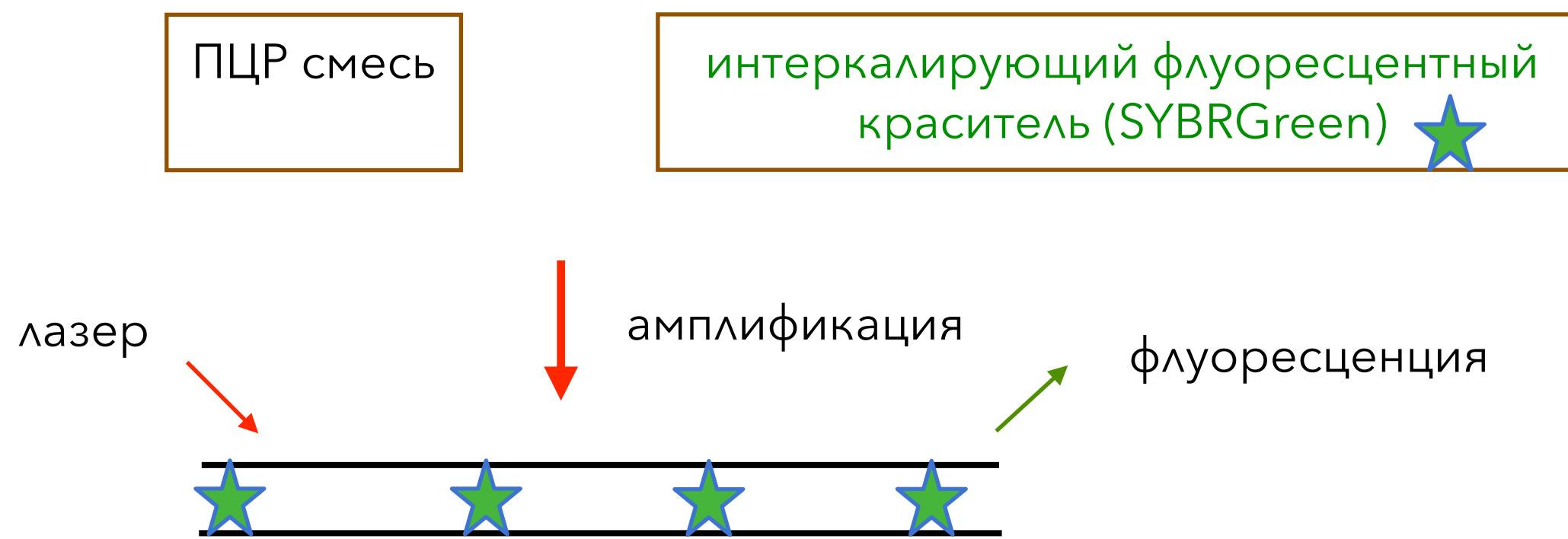


SYBR/LCGreen dye

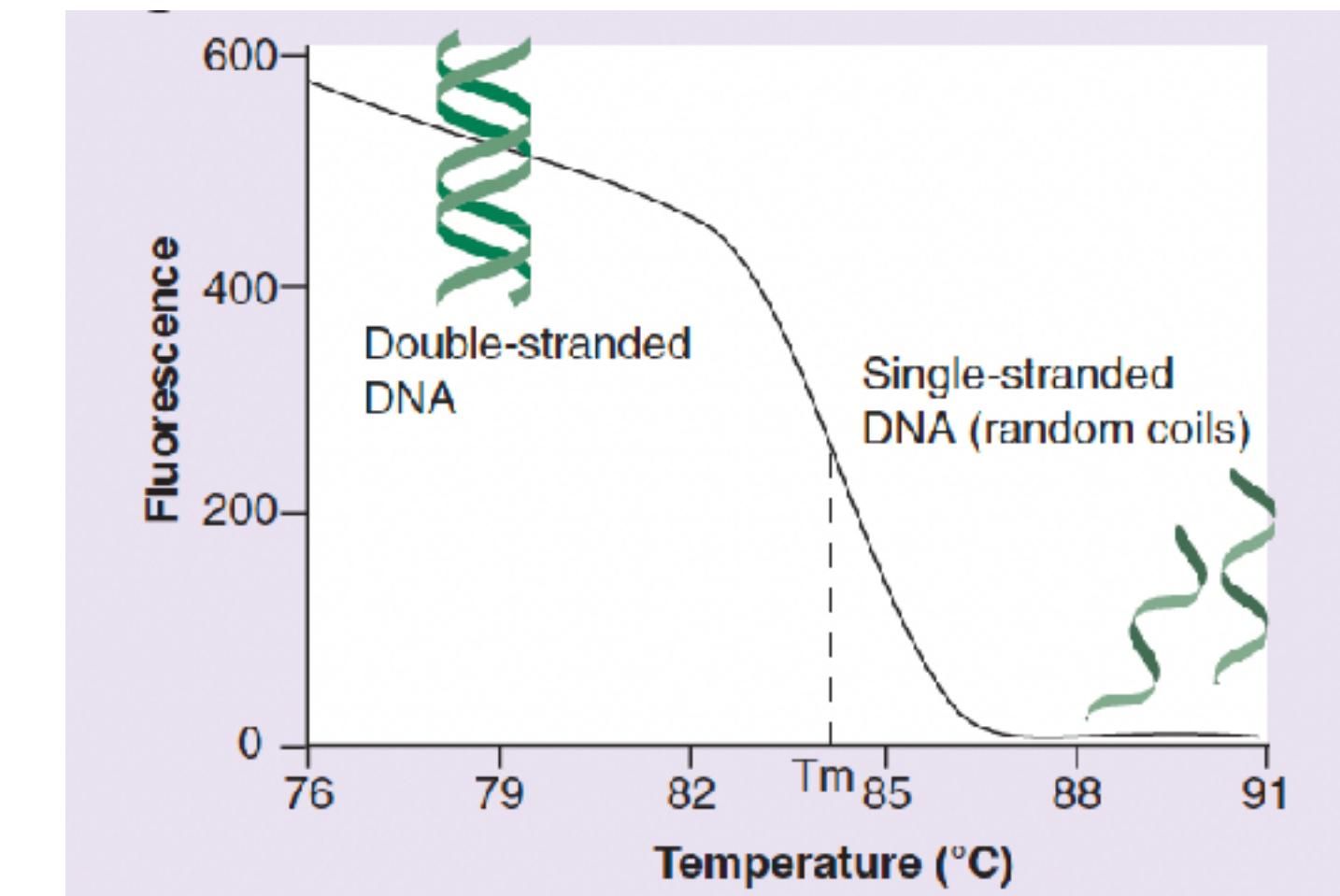
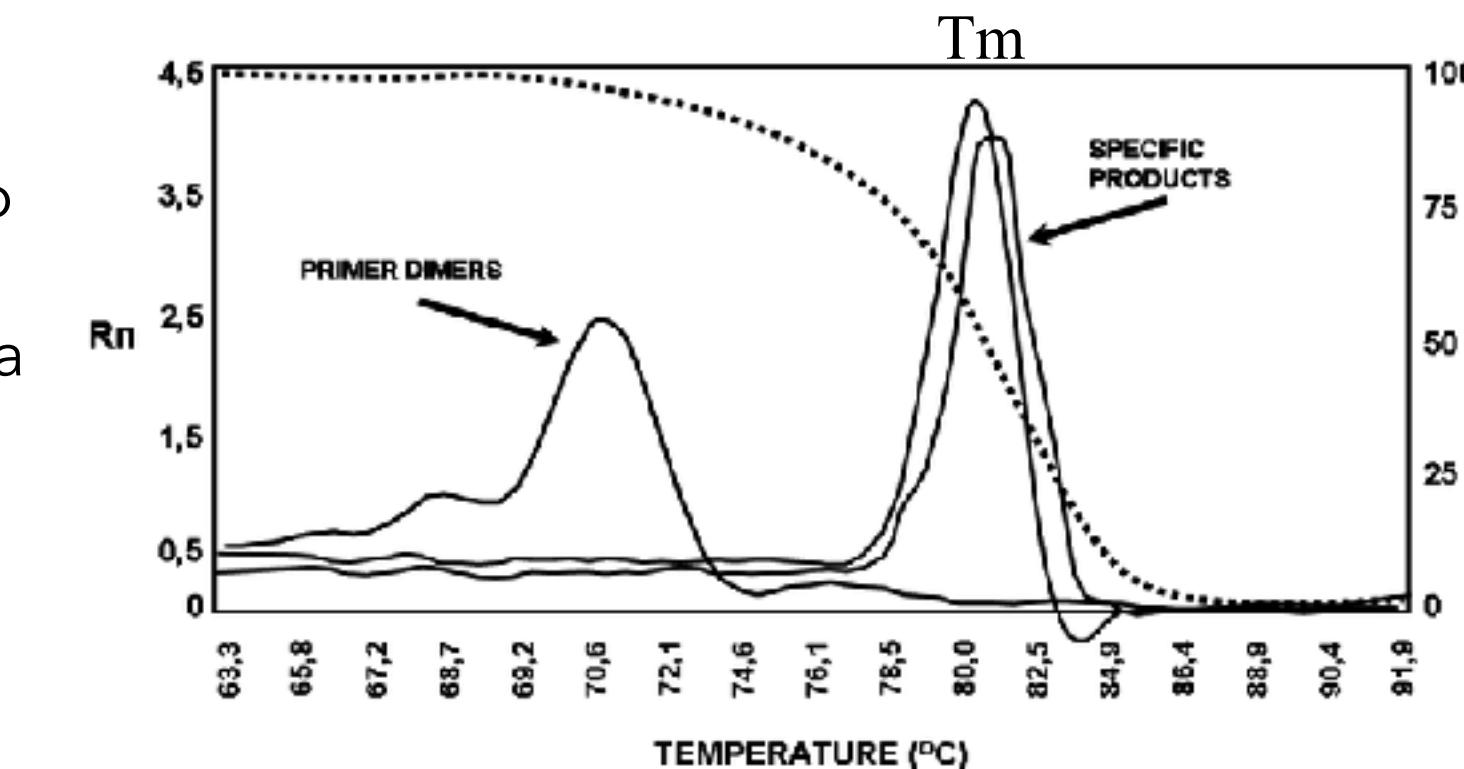
HRM - High Resolution Melting (анализ кривых плавления)

Как работает технология

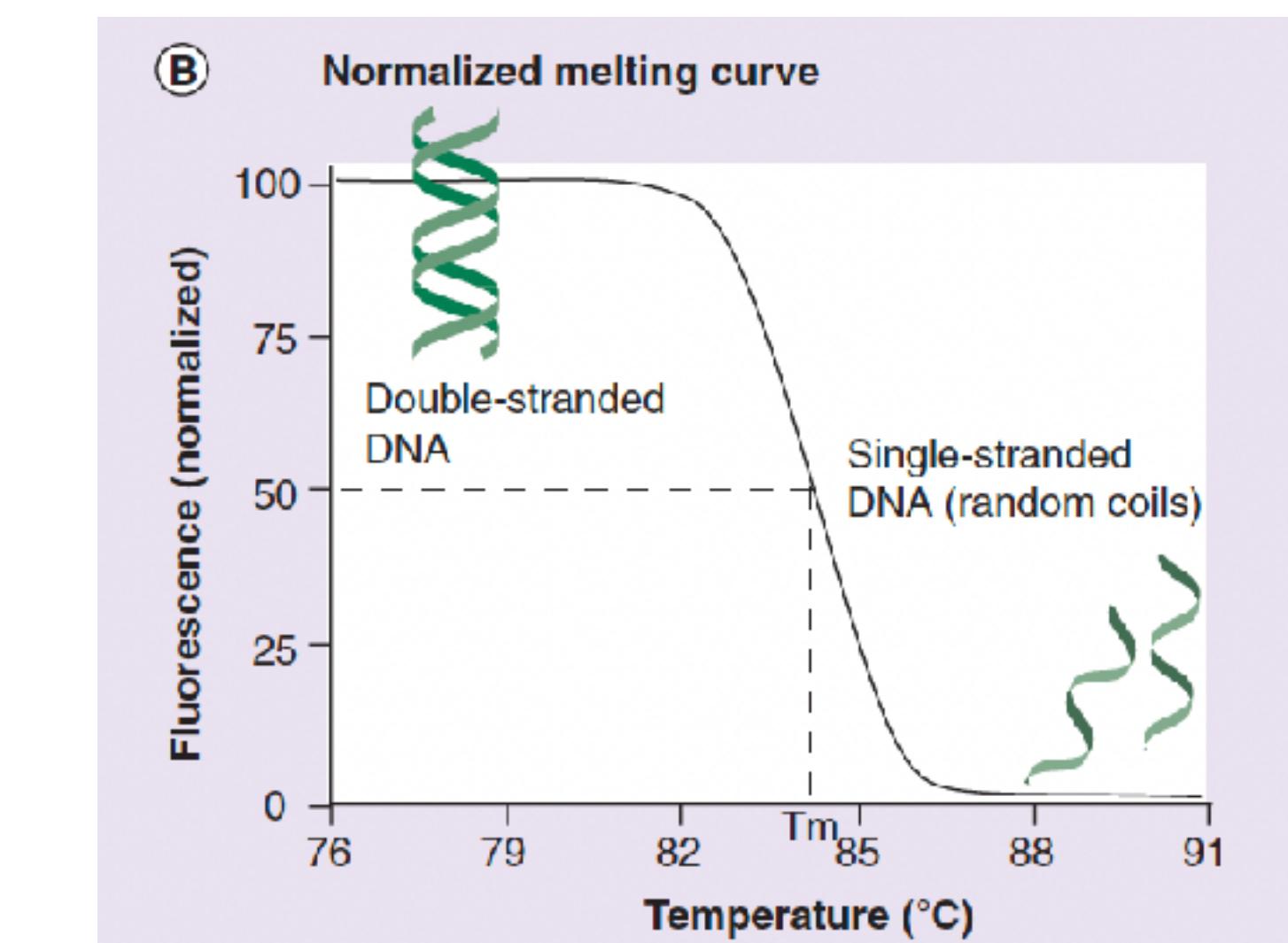
Каждый индивидуальный фрагмент ДНК имеет уникальную T_m (температуру плавления). Она определяется **GC-составом фрагмента, длиной фрагмента и последовательностью нуклеотидов.**



Можно построить кривую плавления для любого ПЦР фрагмента и определить специфическую T_m и идентифицировать его



T_m - точка, когда 50% двуцепочечных фрагментов становятся одноцепочечными.



HRM - High Resolution Melting (анализ кривых плавления)

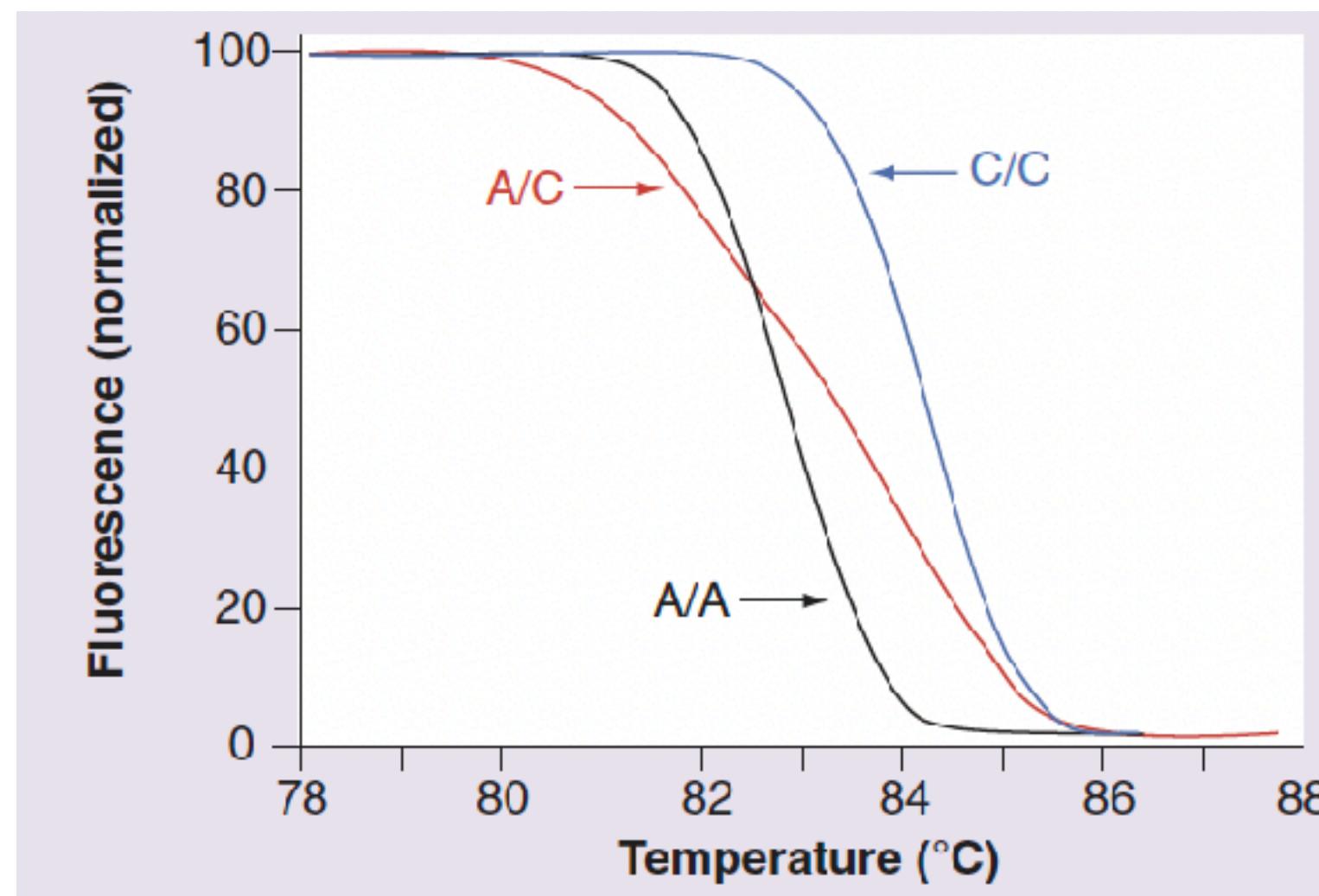
Как работает технология

Результат идентификации будет более точным если анализировать не одну точку - Т_м, а **форму** всей кривой плавления, которая является функцией нуклеотидной последовательности.

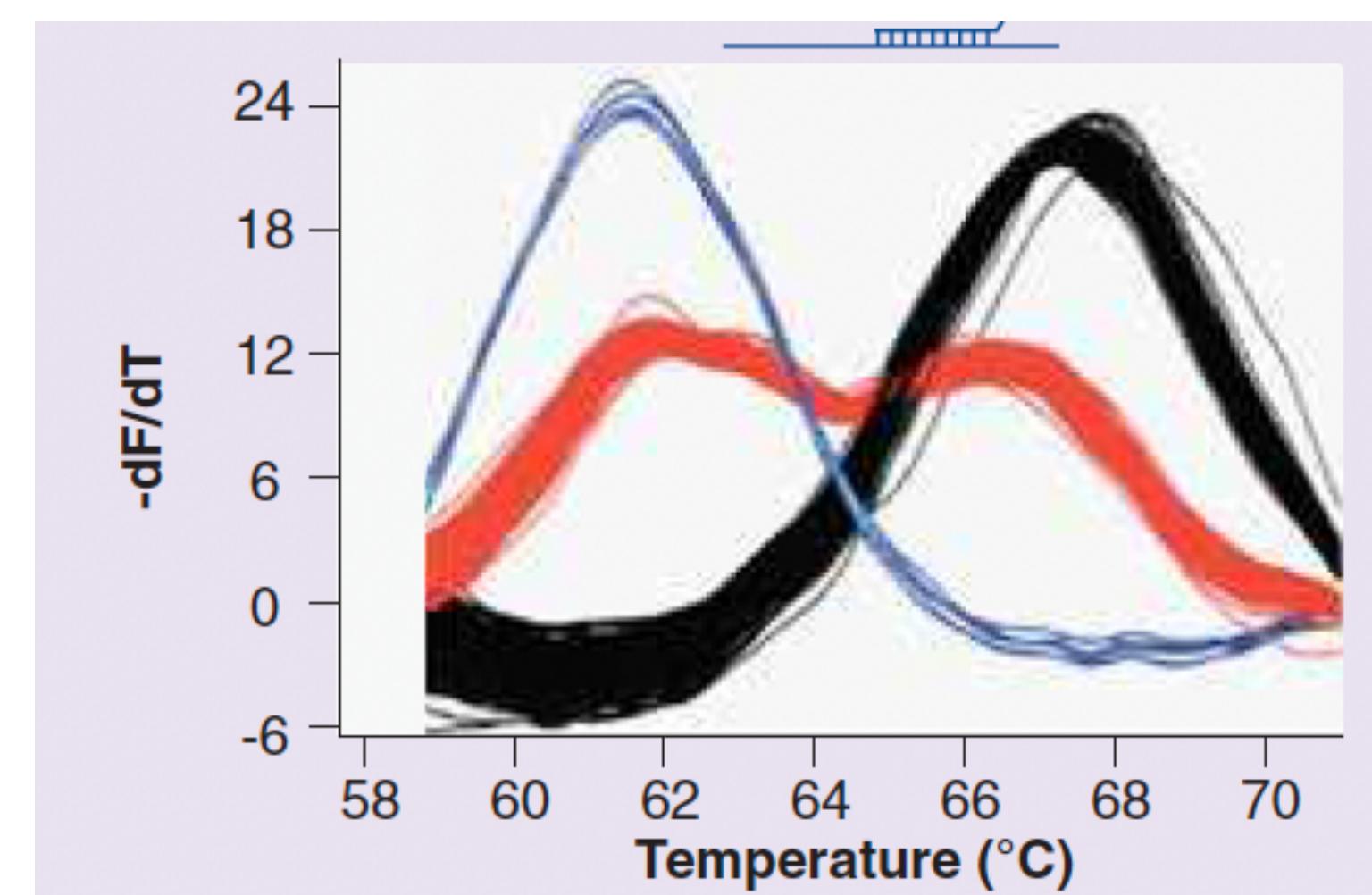
Для этого нужно снять показатели флуоресценции в как можно большем количестве температурных точек в ходе плавления двуцепочечного продукта - это и есть HRM. В ходе HRM плавление дуплекса ДНК происходит со скоростью 0.1-1°C/с, на один градус увеличения температуры регистрация флуоресценции происходит до 65 раз.

Данные по флуоресценции нормализуются - переводятся в % от исходной флуоресценции, которая принимается за 100%. Строится график зависимости флуоресценции от температуры.

Кривая плавления для трех различных генотипов одного SNP локуса
В координатах «нормализованная флуоресценция»-«температура».



То же самое в координатах «производная от флуоресценции»-«температура».



Возможно достоверно отличить две гомозиготы и гетеродуплекс - гомозиготу

Можно использовать для генотипирования полиплоидов

HRM - High Resolution Melting (анализ кривых плавления)

Как работает технология

Этапы HRM анализа

- (1) ПЦР-амплификация целевого региона до высокого числа копий в присутствии флуоресцентного двухцепочечного (ds) ДНК-связывающего красителя;
- (2) плавление флуоресцентно меченного продукта, получение данных о флуоресценции;
- (3) анализ кривых плавления с помощью встроенного в прибор программного обеспечения.

Недостатки HRM анализа

1. Необходим прибор, позволяющий проводить точное считывание флуоресценции и анализировать данные. Ряд современных real-time амплификаторов способны это делать.
2. Необходимо, но необязательно использовать «высоконасыщающий» интеркалирующий краситель типа LCGreen.
3. Зависимость результата от качества подбора праймеров (специфичность ПЦР), полимеразы, условий проведения ПЦР (специфичность и количество продукта ПЦР).
4. Не все мутации возможно отличить по кривой плавления.

Class	Base change	Tm melt curve shift
1	C>T, T>C, G>A, A>G	Large
2	G>A, A>C, G>T, T>G	
3	C>G, G>C	
4	A>T, T>A	Very small

HRM - High Resolution Melting (анализ кривых плавления)

Пример генотипирования SSR-локусов у вишни

Исследование проводилось для тестирования возможности использовать HRM для проведения идентификации Греческих сортов вишни с помощью SSR маркеров.

В исследование были включены 11 сортов вишни:

- 2 международных стандартных сорта Burlat и Var
- 9 редких греческих сортов, включая сорт с Защищенным наименованием (PDO) - Tragana Edessis;

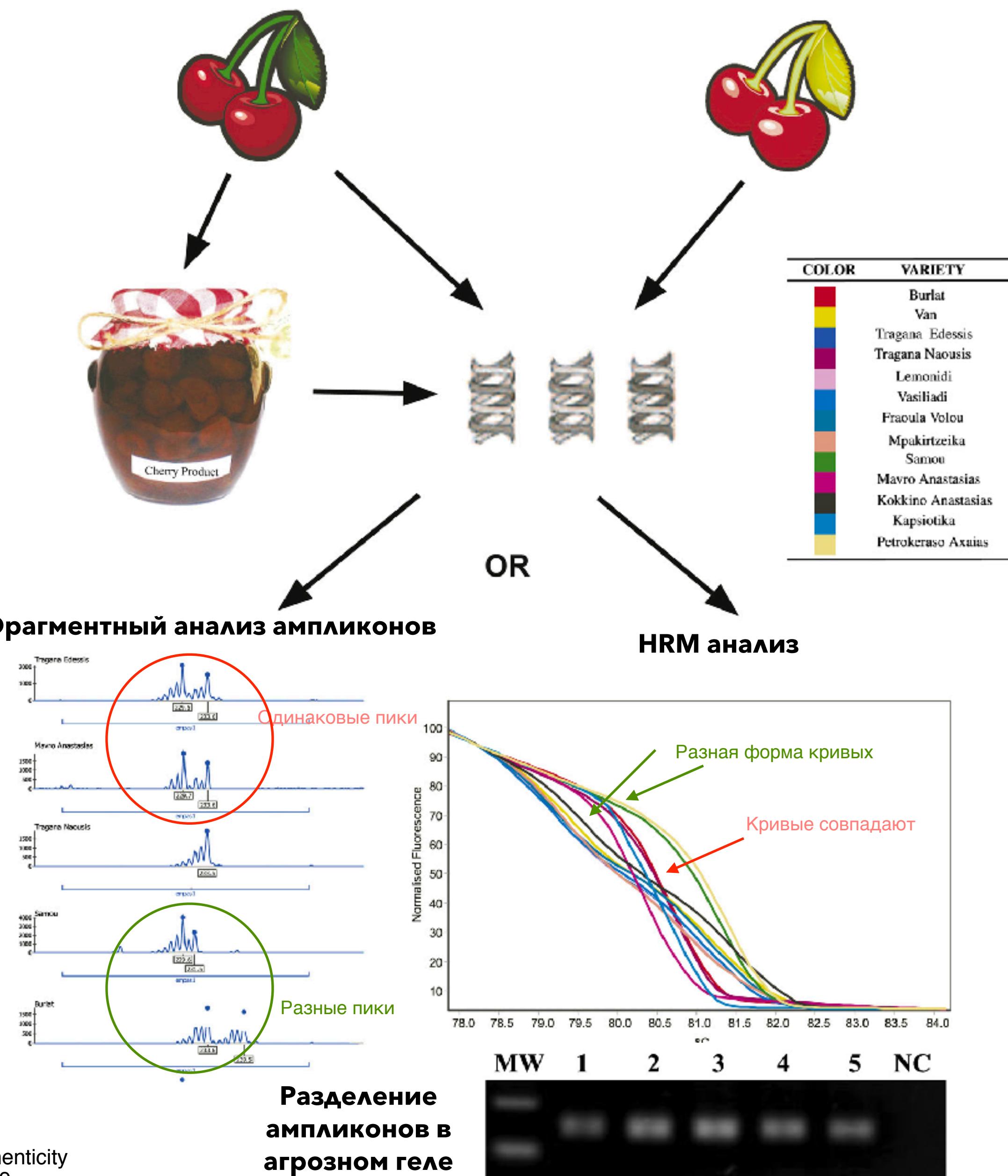
Проводилась оценка состояния SSR локусов непосредственно у растений и у продуктов, полученных из плодов вишни (варенье).

В исследовании использовали 10 SSR локусов

SSR	Applied material	Repeat motif	Forward primer	Reverse primer	Tm (°C)	Allele size range (bp)
EMPA004	C	[GA] ₄ AA[GA] ₄ AA[GA] ₁₅	TACGGTAGGCTTCTGCAAGG	TTGGCAGGTCTGTTCACAT	55	185–212
EMPA005	C	[CT] ₃ CAT[CT] ₁₂ T[AC] ₂₃	TGGGTTTGAGCAATATGCAA	CACCAATAACATGCACACGG	55	246–258
EMPA015	C	[GA] ₃₂	TTTTGGTCATCTGCTGCTG	CTCTCATCTTCCCCCTCCCTC	55	218–255
EMPA018	C	[GA] ₁₈	TCCAAGAACAAAGCCAAAATC	AATTCAATGCATTCTGGATAG	55	97–117
EMPAS01	C	[GA] ₉ [GA] ₁₁	CAAAATCAACAAAATCTAAACC	CAAGAATCTCTAGCTCAAACC	48	224–239
EMPAS02	C, F	[TTG] ₇ CTGC [TG] ₁₀ [AG] ₈	CTACTTCCATGATTGCCCTCAC	AACATCCAGAACATCAACACAC	48	128–149
EMPAS10	C, F	[GA] ₂₈	GCTAATATCAAATCCCAGCTCTC	TGAAGAACTATGGCTTCTGTGG	48	147–169
EMPAS11	C, F	[TC] ₂₅	ACCACTTGAGGAACCTGGG	CTGCCTGGAAGAGCAATAAC	48	68–110
EMPAS12	C, F	[TG] ₁₀ A/GA] ₁₀ AA[GA] ₁₃	TGTGCTAATGCCAAAATACC	ACATGCATTCAACCCACTC	48	123–155
UDP98-412	C, F	[AG]	AGGGAAAGTTCTGCTGCAC	GCTGAAGACGACGATGATGA	60	102–140

Исследование проводили на приборе Rotor-Gene 6500 real-time 5P HRM PCR Thermocycler (Corbett Research).

SSR	Alleles on sweet cherry cultivars	Alleles on sweet cherry products
EMPAS02	132,138,140,144,145,147	128,132,134,136,140
EMPAS10	156,159,161,169	147,149,151
EMPAS11	68,70,72,78,88,106,110	70,72,78,80
EMPAS12	123,139,140,142,147,155	137,139,143
UDP98-412	102,114,120,122,126,128	108,120,132



Спасибо за внимание!