

# ИДЕНТИФИКАЦИЯ ГЕНЕТИЧЕСКОГО РАЗНООБРАЗИЯ КУЛЬТУРНЫХ РАСТЕНИЙ И ИХ ДИКИХ РОДИЧЕЙ ДЛЯ РЕШЕНИЯ ФУНДАМЕНТАЛЬНЫХ И ПРИКЛАДНЫХ ПРОБЛЕМ

Научная статья

УДК 631.523:633.854.78

DOI: 10.30901/2227-8834-2025-4-142-154



## Разработка высокоинформативных микросателлитных маркеров для генетического анализа подсолнечника

С. З. Гучетль, А. В. Головатская, Д. Л. Савиченко, Е. Д. Логинова, Э. И. Жудина

Всероссийский научно-исследовательский институт масличных культур им. В.С. Пустовойта, Краснодар, Россия

Автор, ответственный за переписку: Саида Заурбиеевна Гучетль, saida.guchetl@mail.ru

**Актуальность.** Маркеры микросателлитных локусов, используемые для изучения генетического разнообразия подсолнечника и его идентификации, обладают рядом ограничений. Необходимо создание новых маркеров с три- и более длинным нуклеотидным мотивом, строгой специфичностью к целевому локусу и высоким дискриминационным потенциалом.

**Материалы и методы.** Поиск микросателлитных локусов проводили с использованием программы GMATA по последовательности эталонной сборки генома подсолнечника HanXRQr2.0-SUNRISE. Локусы с фланкирующими областями проверили на наличие копий в других хромосомах с помощью Nucleotide BLAST и провели подбор пар праймеров с помощью Primer-BLAST (NCBI). Предполагаемый полиморфизм микросателлитного локуса оценили путем выравнивания ожидаемого продукта ПЦР на другие сборки геномов подсолнечника из базы GenBank с помощью Nucleotide BLAST. Апробация разработанных маркеров осуществлена с помощью ПЦР и разделения продуктов амплификации методом капиллярного электрофореза в денатурирующих условиях.

**Результаты.** Всего разработано 186 маркеров. Из них 16 характеризовались неспецифичной амплификацией, 31 – полным отсутствием амплификации, у 26 выявлены артефакты SSR-анализа, а 19 не выявили полиморфизм. Остальные 94 маркера оказались информативными и выявили от 2 до 6 аллелей. Диапазон длин амплифицированных фрагментов ДНК составил от 125 до 469 пн. Выявлена статистически значимая взаимосвязь между количеством повторов в микросателлитном локусе и числом аллелей.

**Заключение.** На основании полученных данных возможен отбор наиболее информативных маркеров для формирования мультиплексных ПЦР и создания эффективной системы генотипирования подсолнечника.

**Ключевые слова:** *Helianthus annuus*, SSR-локусы, идентификация, генотипирование

**Благодарности:** работа выполнена в рамках реализации государственных заданий Всероссийского научно-исследовательского института масличных культур им. В.С. Пустовойта № 082-3-2023-0002 «Разработка технологии генетической идентификации и паспортизации линий и гибридов подсолнечника с использованием системы маркеров микросателлитных локусов ДНК».

Авторы благодарят рецензентов за их вклад в экспертную оценку этой работы.

**Для цитирования:** Гучетль С.З., Головатская А.В., Савиченко Д.Л., Логинова Е.Д., Жудина Э.И. Разработка высокоинформативных микросателлитных маркеров для генетического анализа подсолнечника. *Труды по прикладной ботанике, генетике и селекции*. 2025;186(4):142-154. DOI: 10.30901/2227-8834-2025-4-142-154

## IDENTIFICATION OF THE DIVERSITY OF CULTIVATED PLANTS AND THEIR WILD RELATIVES FOR SOLVING FUNDAMENTAL AND APPLIED PROBLEMS

Original article

DOI: 10.30901/2227-8834-2025-4-142-154

### Development of highly informative microsatellite markers for the genetic analysis of sunflower

Saida Z. Guchetl, Anna V. Golovatskaya, Dmitrii L. Savichenko, Elizaveta D. Loginova, Elvina I. Zhudina

V.S. Pustovoit All-Russia Research Institute of Oil Crops, Krasnodar, Russia

**Corresponding author:** Saida Z. Guchetl, [saida.guchetl@mail.ru](mailto:saida.guchetl@mail.ru)

**Background.** The existing microsatellite loci markers, used for studying and identifying sunflower genetic diversity, have certain limitations. There is a need for new markers that will meet the following criteria: three or more nucleotides in the microsatellite motif, specificity to the target locus, and high discriminatory power.

**Materials and methods.** Microsatellite loci were mined using the GMATA software based on the reference genome assembly of sunflower HanXRQr2.0-SUNRISE. Loci with flanking regions were screened for the presence of copies on other chromosomes using Nucleotide BLAST, and primer pairs were designed using Primer-BLAST (NCBI). The expected polymorphism of each microsatellite locus was evaluated by aligning the predicted PCR product to other sunflower genome assemblies from the GenBank database using Nucleotide BLAST. The performance of the developed markers was tested by PCR and separation of amplification products by capillary electrophoresis under denaturing conditions.

**Results.** A total of 186 markers were developed. Of these, 16 were characterized by nonspecific amplification, 31 showed complete absence of amplification, 26 revealed SSR-specific artifacts, and 19 were monomorphic. The remaining 94 markers were informative and detected from 2 to 6 alleles. Among the selected sequences, simple microsatellites with trinucleotide motifs were the most prevalent. The number of markers per chromosome ranged from 2 to 19. The size range of the amplified DNA fragments was from 125 to 469 bp. A statistically significant correlation was found between the number of repeats in the microsatellite locus and the number of detected alleles.

**Conclusion.** The obtained data make it possible to select the most informative markers for the development of multiplex SSR assays and the establishment of an effective genotyping system for sunflower.

**Keywords:** *Helianthus annuus*, microsatellites, identification, genotyping

**Acknowledgments:** the research was carried out within the framework of the state tasks assigned to the V.S. Pustovoit All-Russia Research Institute of Oil Crops, Project No. 082-3-2023-0002 "Development of technology for genetic identification and certification of sunflower lines and hybrids using a system of DNA microsatellite loci markers".

The authors thank the reviewers for their contribution to the peer review of this work.

**For citation:** Guchetl S.Z., Golovatskaya A.V., Savichenko D.L., Loginova E.D., Zhudina E.I. Development of highly informative microsatellite markers for the genetic analysis of sunflower. *Proceedings on Applied Botany, Genetics and Breeding*. 2025;186(4):142-154. (In Russ.). DOI: 10.30901/2227-8834-2025-4-142-154

## Введение

Подсолнечник (*Helianthus annuus* L.) – важная масличная культура как в Российской Федерации, так и во всем мире. Ежегодно растут посевные площади и востребованность семян подсолнечника, что предполагает проведение селекционной работы по созданию новых, высокопродуктивных, устойчивых к биотическим и абиотическим стрессорам сортов и гибридов, а также поддержание сортовой чистоты в процессе семеноводства. Кроме того, в ситуации высокой конкуренции на рынке важно иметь возможность защищать авторские права селекционеров. Немаловажную проблему представляет и появление на рынке фальсифицированных семян. В связи с этим необходимо идентифицировать селекционные достижения. Один из способов – описание отличительных, чаще всего морфологических, признаков. Однако только ограниченный набор признаков удовлетворяет критериям маркеров для идентификации, а их вариативность в пределах вида обычно незначительна. Так, для этой цели Государственной комиссией Российской Федерации по испытанию и охране селекционных достижений принятая методика «Подсолнечник (*Helianthus annuus* L.)», представленная в документе RTG/01/3 «Общее введение по испытанию на отличимость, однородность и стабильность и составлению описаний» от 22.07.2002 г. № 1206/52 (Официальный бюллетень Госкомиссии № 6, 2002 г.). Согласно данной методике, для подсолнечника описывают всего 42 признака. Кроме ограничения количества, морфологические признаки отличаются органоспецифичностью, субъективностью оценки, временным интервалом оценки и т. д. Поэтому актуальна идентификация селекционных достижений с помощью высокотехнологичных инструментов. Такими инструментами в настоящее время служат молекулярно-генетические маркеры ДНК. Методы оценки полиморфизма ДНК, в отличие от морфологических, обладают универсальностью, объективностью и высокой производительностью. В последние несколько десятилетий молекулярные маркеры используются в различных областях, включая исследование генетического разнообразия, картирование генов селекционно ценных признаков, построение карт сцепления, паспортизацию и идентификацию селекционных достижений, маркер-ассоциированную селекцию, популяционную генетику и филогенетические исследования (Liu et al., 2023; Sheller et al., 2023; Fallah et al., 2024).

Для изучения генетического разнообразия подсолнечника и его идентификации применимы различные виды молекулярных маркеров. Чаще всего для этих целей используются маркеры микросателлитных последовательностей. Микросателлиты – простые повторяющиеся последовательности (Simple Sequence Repeats, SSR), обнаруженные у прокариот и эукариот. Они широко распространены по всему геному, особенно в эухроматине, а также в кодирующей и некодирующей ядерной ДНК и ДНК органелл (Brunel, 1994; Yadav et al., 2021). Их преимущества заключаются в кодоминантном типе наследования, мультиаллельности, значительной информативности и простоте применения.

В предыдущие годы рядом авторов с использованием обогащенных и необогащенных геномных библиотек ДНК разработано более 500 SSR-маркеров подсолнечника (Brunel, 1994; Paniego et al., 2002). Из них менее половины демонстрировали полиморфизм среди инбредных линий. Позже, в результате анализа 2033 клонов из обогащенных (AC)n/(AG)n геномных библиотек, было

выявлено 1093 уникальных SSR-последовательностей, на основе которых разработано 879 SSR-маркеров. Тестирование показало, что 579 из них (65,9%) являлись полиморфными (Tang et al., 2002). Для своего времени это были прорывные исследования. С помощью этих SSR-маркеров выполнено множество исследований по паспортизации родительских линий и гибридов подсолнечника селекции разных стран (Ramya et al., 2019; Ahmed et al., 2022), изучению генетического разнообразия и структуры популяций (Filippi et al., 2015), картированию генома подсолнечника (Tang et al., 2002), маркированию селекционно ценных признаков (Ramazanova et al., 2021). Установлено, что SSR-маркеры, разработанные на основе микросателлитов из некодирующих областей генома, выявляли большее разнообразие, чем EST-SSR (Expressed Sequence Tag) (Filippi et al., 2015).

Авторами данного исследования на основе SSR-маркеров, разработанных S. Tang с соавторами (Tang et al., 2002) и N. Paniego с соавторами (Paniego et al., 2002), были созданы системы маркеров для идентификации и паспортизации селекционного материала подсолнечника Всероссийского научно-исследовательского института масличных культур им. В.С. Пустовойта (ВНИИМК), определения генетической чистоты партий семян и исследования генетического разнообразия линий с использованием электрофоретического разделения фрагментов ДНК в агарозном или полиакриламидном геле (Guchetl et al., 2023). Однако разрешающая способность классических методов гель-электрофореза не позволяет достичь высокой точности, сходимости и воспроизводимости результатов анализа. Также снижается информативность маркерных систем за счет невозможности надежной регистрации нуклеотидного полиморфизма в 1-5 пн. Большая часть применяемых микросателлитов характеризуется динуклеотидными мотивами, которые могут вызывать образование статтеров (stutters) – артефактов ПЦР-амплификации, возникающих из-за «проскальзывания» ДНК-полимеразы на микросателлитных повторах. Во время капиллярного электрофореза это приводит к появлению дополнительных фрагментов разной длины, что затрудняет точное определение истинного размера аллеля микросателлитного локуса. С развитием методов полногеномного секвенирования и сборки генома появилась возможность оценить качество маркеров с помощью биоинформационных подходов. В ходе оценки было установлено, что ранее разработанные маркеры не всегда обладают строгой монолокусной специфичностью (Guchetl et al., 2023). Их достаточно сложно комбинировать для проведения мультиплексной ПЦР (Shilov et al., 2023). Одно из требований к маркерам – это их высокий дискриминационный потенциал. Поскольку генетическое разнообразие культурного подсолнечника снижается в процессе селекции, ДНК-маркеры должны иметь достаточный дискриминационный потенциал для идентификации растений в рамках локальных коллекций, адаптированных к местным условиям выращивания.

В последнее время доступность технологий высокопроизводительного секвенирования позволила исследователям идентифицировать значительное количество микросателлитов с меньшими затратами и усилиями, чем традиционные подходы. Применение NGS (Next-Generation Sequencing) значительно упростило разработку SSR-маркеров, распределенных по всему геному. Созданы и различные базы данных микросателлитных последовательностей для разных культур, которые поз-

воляют исследователям отбирать требуемые маркеры (Avvaru et al., 2020; Topu et al., 2023). Также современный высокопроизводительный подход к генотипированию живых организмов требует автоматизации на всех этапах анализа. И одним из наиболее эффективных на сегодняшний день способов автоматической детекции результатов ПЦР является метод капиллярного гель-электрофореза. Данный подход позволяет проводить генотипирование подсолнечника сразу по набору маркеров, благодаря проведению мультиплексных ПЦР с использованием нескольких флуоресцентных красителей. Однако создание таких наборов предъявляет требования к маркерам, такие как схожая температура отжига, отсутствие неспецифичной амплификации, отсутствие артефактов амплификации.

Ранее нами с целью поиска перспективных микросателлитных ДНК-маркеров из литературных источников с применением инструмента Primer-BLAST и эталонного генома подсолнечника *in silico* проверены более 200 SSR-локусов. Предпочтение отдавалось локусам с высокой степенью полиморфизма, а также микросателлитам с три- и более нуклеотидными повторами, локализацией в некодирующих последовательностях ДНК, отжигом flankирующих праймеров только на целевой последовательности. По результатам поиска отобрали и апробировали на 25 генотипах подсолнечника селекции ВНИИМК 21 SSR-маркер. Только три маркера оказались приемлемыми для генотипирования подсолнечника (Golovatskaya, Guchetl, 2023). Эти результаты показывают, что разработанные ранее маркеры не отвечают требованиям для включения их в наборы для мультиплексных ПЦР. В свою очередь, создание новых SSR-маркеров, заранее отвечающих заданным требованиям при их оценке *in silico* и проведении апробации, является перспективным подходом для построения систем генетической идентификации и паспортизации подсолнечника.

Цель исследования – разработать новые маркеры микросателлитных локусов подсолнечника на основе эталонной сборки генома подсолнечника HanXRQr2.0-SUNRISE (GCA\_002127325.2), размещенной в базе данных NCBI (National Center for Biotechnology Information), для паспортизации и идентификации линий, гибридов и сортов подсолнечника и апробировать их методом капиллярного гель-электрофореза.

## Материалы и методы

В качестве растительного материала использовали 5 линий (ВК1 сур, ВК21 сур, ЭД73, ВК678, ВК585) и 2 гибрида (Сурус и Паритет) подсолнечника селекции ВНИИМК.

Для выделения ДНК растительный материал предварительно измельчали с помощью гомогенезатора KZ-III-FP (Servicebio, КНР). Экстракция ДНК осуществлялась из разных органов подсолнечника с помощью наборов реагентов «МагноПрайм ФИТО» («НекстБио», РФ) с применением станции автоматического выделения и очистки нуклеиновых кислот Auto-Pure 96 (Allsheng, КНР). Определение количества и качества выделенной ДНК проводилось на спектрофотометре Nano-300 (Allsheng, КНР).

Для проведения ПЦР использовались 25 мкл реакционной смеси следующего состава: 2,5 мкл 10 × ПЦР-Буфер-Б («Синтол», РФ); 2,5 мкл 25 mM MgCl<sub>2</sub>; 0,5 мкл 100 mM dNTP; по 0,1 мкл 100 пМ/мкл каждого праймера; 50 нг матричной ДНК и 0,2 мкл 5 е. а. SynTaq ДНК-поли-

меразы («Синтол», РФ). Амплификация выполнялась в термоцикльере MiniAmp™ plus (Thermo Scientific, США) при следующих температурно-временных режимах: начальная денатурация при 96°C в течение 2 мин, 35 циклов при температурно-временном режиме: денатурация при 94°C – 30 сек, отжиг при 58–60°C – 40 сек, элонгация при 72°C – 1 мин, финальная элонгация при 72°C – 2 мин.

Разделение продуктов амплификации осуществляли методом капиллярного электрофореза в денатурирующих условиях на генетическом анализаторе «Нанофор-05» (ИАП РАН, РФ). Для повышения пропускной способности использовали мультиплексный гель-электрофорез: смешивали по 5 мкл ПЦР-продуктов четырех разных реакций. Подбор реакций для приготовления смеси основывался на использовании праймеров, меченых разными флуоресцентными красителями. Для анализа длин фрагментов 1 мкл смеси ПЦР-продуктов смешивали с 0,5 мкл размерного стандарта СД-600 и 8,5 мкл деионизированного формамида, затем проводили денатурацию фрагментов в течение 5 мин при 95°C. Размер фрагментов определяли относительно размерного стандарта СД-600 меченым флуоресцентным красителем (Dy-632) с помощью компьютерного программного обеспечения GeneMarker software version 3.0.1.

Для разработки новых SSR-маркеров с применением биоинформационных методов использовали базы данных GenBank, RefSeq (Reference Sequence) (NCBI) (<https://www.ncbi.nlm.nih.gov>). С помощью эталонной сборки генома подсолнечника HanXRQr2.0-SUNRISE и программы GMATA проводили поиск микросателлитных локусов с три- и более нуклеотидными мотивами (Wang X., Wang L., 2016). Для фильтрации микросателлитных последовательностей по количеству повторов использовали программную среду R. Проверку нуклеотидной последовательности, содержащей искомый микросателлит и обе flankирующие последовательности, на наличие копий с высокой степенью идентичности в других хромосомах проводили *in silico* с помощью программы Nucleotide BLAST (NCBI) (Camacho et al., 2009). Подбор праймеров, flankирующих микросателлитный локус, осуществляли с помощью инструмента Primer-BLAST (NCBI) (<https://www.ncbi.nlm.nih.gov/tools/primer-blast>). Также с помощью Nucleotide BLAST оценивали предполагаемый полиморфизм микросателлитного локуса выравниванием полученных последовательностей на другие сборки геномов подсолнечника (HanHA89r1.0-20210811, ASM2653828v1, HanOQP8r0.9-20211025), опубликованных в базе данных GenBank. Веб-версией инструмента Multiple Primer Analyzer (Thermo Fisher Scientific) провели предварительную оценку возможности образования само- и кроссдимеров последовательностей прямого и обратного праймеров.

Предварительную обработку и структурирование данных осуществляли в программной среде R версии 4.2.3 (<https://www.r-project.org>) с использованием пакета dplyr. Графическую визуализацию данных выполняли с помощью пакета ggplot2 (Wickham, 2016). Предварительное распределение локусов по группам в зависимости от сумм количества повторов осуществляли методом кластеризации k-средних, с числом кластеров равным трем и фиксированным начальным значением генератора случайных чисел set.seed(123) для обеспечения воспроизводимости.

Коэффициент размера эффекта *h* рассчитывали по формуле согласно J. Cohen (1988):

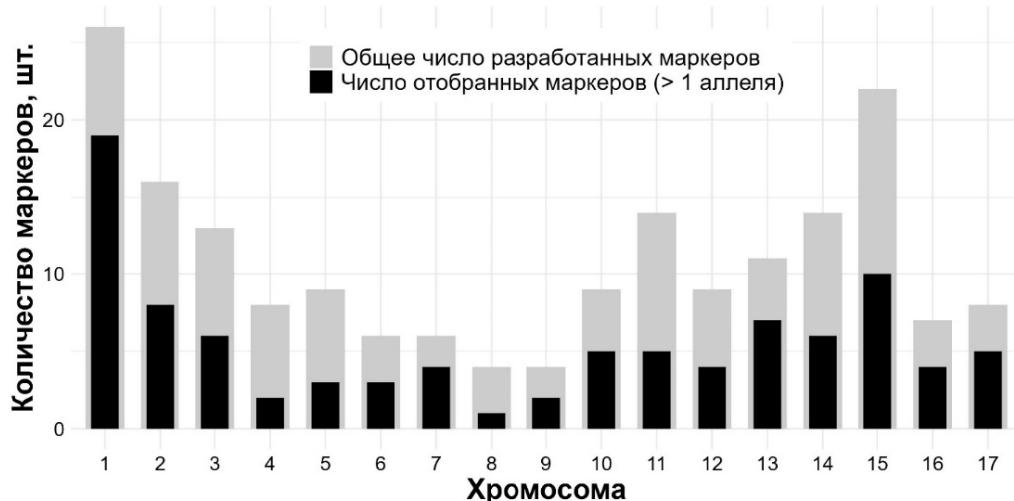
$$h = 2 \arcsin(\sqrt{p_1}) - 2 \arcsin(\sqrt{p_2})$$

## Результаты

Проверку значений на нормальность распределения осуществляли с использованием критерия Шапиро – Уилка. При отклонении от нормальности использовали непараметрические методы: для оценки взаимосвязи между переменными – коэффициент ранговой корреляции Спирмена; для сравнения значений между несколькими группами – критерий Краскела – Уоллиса с апостериорными попарными сравнениями по методу Данна с поправкой Холма – Бонферрони, реализованными в пакете FSA. Для анализа категориальных данных применяли критерий  $\chi^2$  Пирсона. Распределение категориальных признаков оценивали с помощью теста попарного сравнения долей с поправкой Холма – Бонферрони, реализованного в функции pairwise.prop.test(). Статистические тесты, для которых не указано иное, выполняли с помощью базового пакета stats.

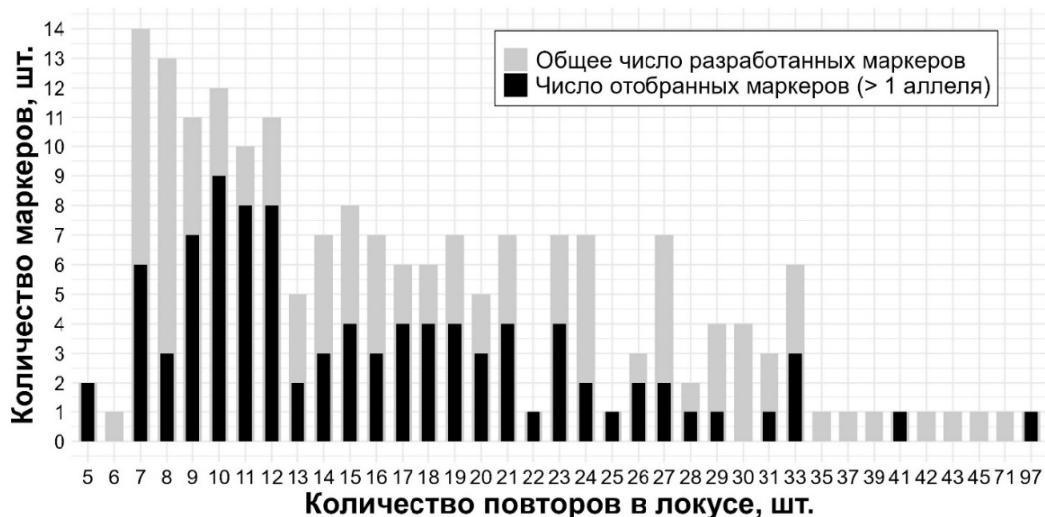
В результате работы нами было разработано 186 маркеров. Маркеры покрывали все хромосомы, однако их распределение оказалось неравномерным, поскольку хромосомы отличались по доле локусов, удовлетворяющих установленным критериям. Количество маркеров на хромосому составило от 4 (для 8 и 9 хромосом) до 26 (для 1 хромосомы) (рис. 1).

33 микросателлита были составными, то есть содержали два или более разных мотива, разделенных короткой уникальной последовательностью, и оставшиеся 153 – простыми (содержали один мотив). Наибольшим было количество простых микросателлитов с тринуклеотидным мотивом – 126, 25 – с тетрануклеотидным и 2 – с пентануклеотидным. Количество повторов тоже различалось и варьировалось от 3 до 97 (рис. 2).



**Рис. 1. Распределение по хромосомам разработанных и отобранных после аprobации SSR-маркеров подсолнечника**

**Fig. 1. Chromosomal distribution of the developed and selected sunflower SSR markers after testing**



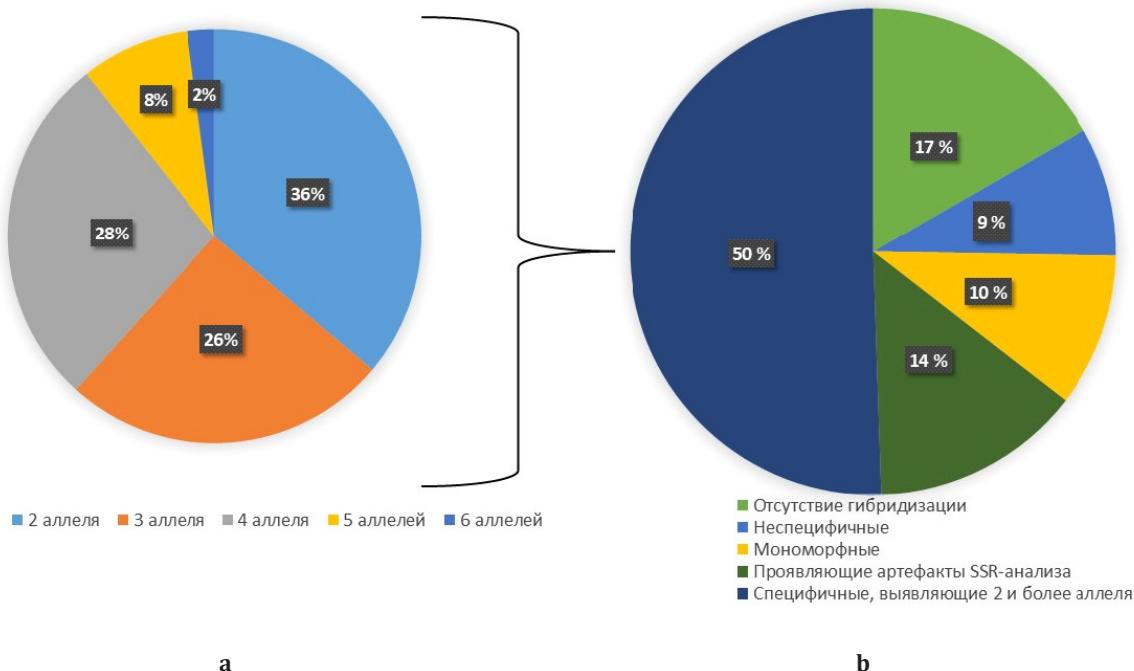
**Рис. 2. Число мотивов в микросателлитных локусах, для которых разработаны маркеры, и количество маркеров до и после аprobации**

**Fig. 2. Number of motifs in the microsatellite loci for which markers were developed and the number of these markers before and after testing**

Разработанные SSR-маркеры были апробированы с помощью полимеразной цепной реакции ПЦР на 7 генотипах подсолнечника. При данных условиях проведения ПЦР 31 пары праймеров не давала целевого фрагмента, 16 амплифицировали неспецифические продукты, а у 26 были выявлены артефакты SSR-анализа (стартеры, побочные пики, неполное аденилирование). У 19 SSR-локусов был выявлен только 1 аллель. 94 маркера (около 50% из общего числа разработанных) оказались информативными и выявляли от 2 до 6 аллелей у изученных образцов (рис. 3).

## Обсуждение результатов

По литературным данным, разработка SSR-маркеров, как правило, включает следующие этапы: (I) наличие предварительных данных о нуклеотидных последовательностях, содержащих SSR; (II) подбор олигонуклеотидов (праймеров), комплементарных участкам, flankирующим SSR; (III) проверку праймеров методом ПЦР с последующим электрофоретическим анализом продуктов реакции; (IV) выявление полиморфизма между исследуемыми образцами (Mason, 2015). В данном исследовании



**Рис. 3. Диаграмма результатов апробации экспериментальных SSR-маркеров на семи генотипах подсолнечника: а – количество обнаруженных аллелей у специфичных и полиморфных SSR-маркеров; б – процентное отношение числа информативных маркеров к общему числу по результатам апробации**

**Fig. 3. Diagram of the results of testing experimental SSR markers on seven sunflower genotypes:**

**а** – the number of detected alleles in specific and polymorphic SSR markers; **б** – the percentage of informative markers relative to the total number of markers based on the testing results

Из 94 отобранных микросателлитных последовательностей 14 были составными и 80 – простыми. Наибольшим было количество простых микросателлитов с тринуклеотидным мотивом – 69, 10 – с тетрануклеотидным и 1 – с пентануклеотидным. Количество маркеров на хромосому варьировало от 2 (хромосомы 4 и 8) до 19 (хромосома 1) (см. рис. 1).

Количество повторов у апробируемых маркеров варьировалось от 5 до 97 (см. рис. 2).

Диапазон длин амплифицированных фрагментов ДНК находился в пределах от 125 до 469 пн, что согласуется с ожидаемыми *in silico* и укладывается в рабочий диапазон стандарта длин СД-600 (до 600 пн), применяемого для капиллярного электрофореза. Также 59 маркеров характеризовались разницей в длинах аллелей  $\geq 10$  пн, что делает их перспективными для визуализации в агарозном геле. Однако это значительно снижает их информативность: 45 маркеров имели 2 аллеля, 13 – 3 аллеля и 1 маркер – 4 аллеля. При этом следует отметить, что эти значения не проверялись напрямую в агарозном геле, а являются аппроксимацией, основанной на данных, полученных с помощью капиллярного электрофореза.

при разработке новых маркеров были определены этапы, близкие по структуре к описанным рекомендациям. Также для этапов *in silico* были выделены следующие критерии при разработке набора маркеров: три и более нуклеотидов в мотиве искомой микросателлитной последовательности, охват всех хромосом, специфичность к целевому локусу (отсутствие неспецифичной мультилокусной амплификации), внегенная локализация, теоретически ожидаемый размер амплифицируемого фрагмента от 100 до 500 пн. Использование этих этапов и критерии позволило разработать ряд пригодных для дальнейшей апробации маркеров.

Однако перечисленные этапы отбора относятся к теоретическому моделированию и в основном позволяют только снизить долю отбракованных в будущем маркеров. Самым важным критерием качества ДНК-маркера является результат этапа его оценки посредством проведения ПЦР на выборке из разных генотипов. На данном этапе ключевое значение имеет способ детекции продуктов ПЦР – аллельных вариантов SSR-локусов. Полиморфизм длин амплифицированных фрагментов обычно выявляют с помощью метода гель-электрофореза, используя агарозный или поликарбамидный гель

(ПААГ). Наибольшей разрешающей способностью в 1 пн обладает метод капиллярного гель-электрофореза в N,N-полидиметилакриламиде (ПДМА) с использованием генетического анализатора. Первые два метода являются экономически менее затратными и потому могут использоваться в качестве предварительных. Однако в процессе работы нами была выявлена низкая эффективность поэтапной апробации маркеров с использованием гелей перед капиллярным электрофорезом. Это связано с тем, что высокая чувствительность последнего метода позволяет выявлять такие артефакты ПЦР, как «заикание» (stuttering), неполное аденилирование, которые незаметны при анализе с помощью классического гель-электрофореза как в агарозе, так и в нативном и денатурирующем ПААГ. Кроме того, в нашем исследовании отсутствие амплификации наблюдалось лишь у 20% маркеров, а оценка по уровню полиморфизма сопряжена с высокой вероятностью ошибочной отбраковки информативных маркеров, ошибочно классифицирующихся как мономорфные или низкополиморфные по причине низкой разрешающей способности агарозного геля. Главным образом такой подход приводит к большим временным затратам, что говорит о нецелесообразности предварительного скрининга для отбраковки маркеров. Исключение предварительных этапов позволило в несколько раз повысить скорость разработки и проверки новых маркеров и в конечном счете уменьшило экономические затраты, несмотря на повышенную стоимость праймеров с флуоресцентными метками. Апробация маркеров была также усовершенствована мультиплексным гель-электрофорезом. Амплификация каждого маркера выполнялась отдельно, но детекция результатов для каждого образца проводилась путем смешивания продуктов амплификации нескольких маркеров и проведением фрагментного анализа этой смеси.

В контексте выбора лучших локусов для создания новых SSR-маркеров большой интерес представляет вопрос об оптимальной длине микросателлита. Другими словами, связь между количеством повторов в локусе и количеством обнаруженных аллелей. В литературе описаны разные типы зависимости дискриминационного потенциала маркера от числа нуклеотидов мотива и длины микросателлита. Считается, что локусы SSR с большим количеством повторов имеют более высокую частоту мутаций (Verbiest et al., 2023) и выбор локусов с достаточным количеством повторов необходим для обеспечения полиморфизма. Однако SSR с многочисленными повторами имеют и некоторые недостатки, такие как повышенное выпадение аллелей (Buchan et al., 2005) и усиление статтеринга (Hoffman, Amos, 2005). Промежуточное количество повторов может представлять собой хороший компромисс, сохраняя большинство преимуществ SSR (высокое разнообразие), избегая при этом некоторых их недостатков, вызванных очень высокой частотой мутаций (Guichoux et al., 2011). В связи с необходимостью учета всех этих факторов нами были отобраны локусы с различным количеством повторяющихся мотивов. Несмотря на ограниченный объем выборки генотипов, который потенциально привел к заниженной оценке количества аллелей, все рассматриваемые локусы находились в равных условиях, что потенциально позволяет провести поиск закономерностей.

Для проведения анализа из исходного набора данных была исключена информация по отбракованным маркерам, а также локус, состоящий из 97 повторов и характеризующийся тремя аллелями. Для составных микроса-

теллитов учитывалась сумма повторов всех мотивов. Это обосновано тем, что полиморфизм длин фрагментов может быть обусловлен изменением количества каждого из мотивов.

Проверка данных на нормальность распределения показала значимое отклонение от нормальности по обеим переменным, что не позволило применить метод Пирсона для оценки корреляции. В связи с этим была проведена непараметрическая оценка – коэффициент ранговой корреляции Спирмена (рис. 4).

Проведенный анализ выявил статистически значимую, но слабую положительную взаимосвязь между числом аллелей и количеством повторов в микросателлитном локусе ( $p = 0,34$ ;  $p < 0,001$ ). Несмотря на это, визуализация наблюдений показала, что данная связь носит приближенный характер. Наблюдаемая высокая дисперсия значений, в особенности при 1–4 аллелях, указывает на то, что один и тот же уровень аллельного разнообразия может соответствовать разному количеству повторов в широком диапазоне.

С целью определения уровней количества аллелей, между которыми наблюдаются статистически значимые различия в числе повторов в микросателлитном локусе, провели попарное сравнение с использованием теста Манна – Уитни, за исключением группы с 6 аллелями, которая была представлена двумя наблюдениями (табл. 1).

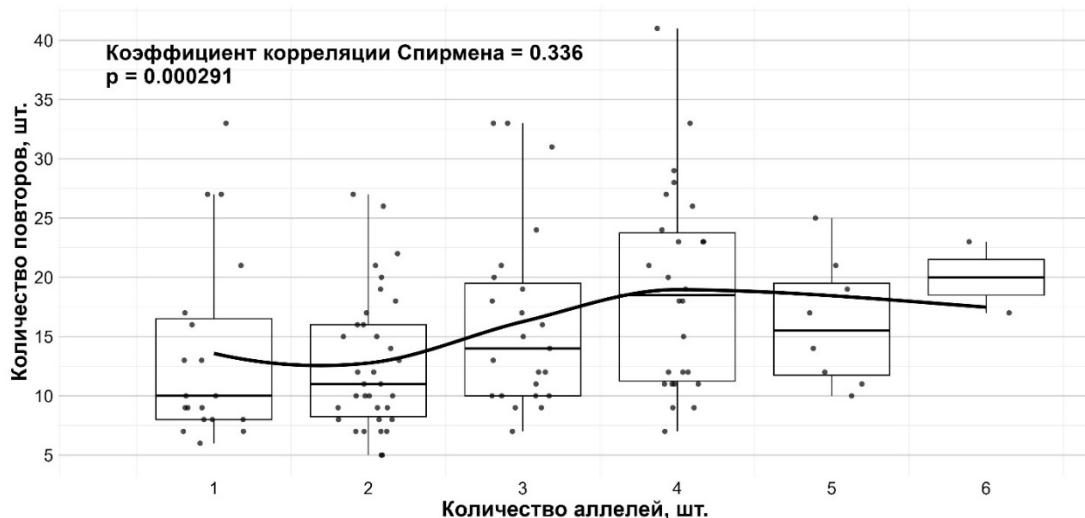
Результаты проведенного анализа согласуются с коэффициентом корреляции: после поправки на множественные сравнения статистически значимым оказалось только сравнение 2 и 4 аллелей, в остальных парах различия не были значимы.

Несмотря на слабую корреляцию и ограниченные различия между группами аллелей, наблюдаемая тенденция в распределении локусов с разным числом повторов в зависимости от аллельного разнообразия указывает на перспективность определения ориентировочного диапазона числа повторов, при котором вероятность обнаружения более полиморфных локусов оказывается выше. Данная информация может быть полезна для повышения эффективности разработки новых высокополиморфных SSR-маркеров.

Для определения диапазонов числа повторов была применена кластеризация локусов методом k-средних. В качестве группирующей переменной также использовалось суммарное количество повторов всех мотивов в локусе (рис. 5).

Было определено три кластера, включающие 58 локусов с суммарным количеством повторов от 5 до 13, 33 локуса с 14–22 повторами и 21 локус с 23–41 повтором соответственно. Далее проводился анализ распределения локусов по числу аллелей в каждом кластере (рис. 6).

Из рисунка 6 видно, что в первом кластере преобладали локусы с двумя аллелями, тогда как количество локусов с одним, тремя и четырьмя аллелями постепенно уменьшалось с резким снижением на пятом аллеле. Во втором кластере распределение было более равномерным – максимум также приходился на локусы с двумя аллелями, но с более равномерным уменьшением количества локусов по мере увеличения числа аллелей. В третьем кластере наблюдалось ярко выраженное преобладание локусов с четырьмя аллелями, количество которых более чем в два раза превышало количество локусов с любым другим числом аллелей. Эти данные согласуются с результатами ранее проведенного теста Манна – Уитни, показавшего статистически значимые различия



**Рис. 4.** Распределение суммарного количества повторов в микросателлитных локусах в зависимости от числа выявленных аллелей (прямоугольник соответствует межквартильному размаху (IQR), линия внутри прямоугольника – медиана, «усы» показывают минимальное и максимальное значения; точки отражают отдельные наблюдения, линия – локально сглаженный тренд, отражающий общую тенденцию)

**Fig. 4. Distribution of the total number of repeats in microsatellite loci depending on the number of detected alleles** (the rectangle represents the interquartile range (IQR), the line inside the rectangle indicates the median, and the “whiskers” show the minimum and maximum values; dots represent individual observations, and the line shows a locally smoothed trend reflecting the overall pattern)

**Таблица 1.** Результаты попарного сравнения групп маркеров, различающихся по числу аллелей, в зависимости от количества повторов в микросателлитном локусе

**Table 1. Results of pairwise comparison of the groups of markers differing in the number of alleles, depending on the number of repeats in the microsatellite locus**

Группы маркеров по количеству аллелей		Статистика (W)	P	P (скоррект.)
Группа 1	Группа 2			
2	4	237,5	0,002	0,023
1	4	141,0	0,015	0,136
2	3	279,5	0,043	0,342
1	3	149,5	0,056	0,390
2	5	83,0	0,092	0,552
1	5	46,0	0,116	0,581
1	2	316,0	0,904	1,000
3	4	265,0	0,366	1,000
3	5	89,5	0,794	1,000
4	5	121,5	0,489	1,000

по количеству повторов между локусами с 2 и 4 аллелями.

Для проверки статистической значимости различий в распределении числа аллелей между кластерами провели непараметрический тест Краскела – Уоллиса и апостериорные попарные сравнения с помощью теста Данна с поправкой Холма – Бонферрони (табл. 2).

Проведенный тест Краскела – Уоллиса выявил статистически значимые различия между тремя кластерами по числу аллелей ( $p = 0,016$ ). Апостериорный критерий

доказал наличие статистически значимых различий только между первым и третьим кластерами. Отличия между остальными кластерами не достигли порога статистической значимости. Однако стоит отметить, что несмотря на значимость различий первого и третьего кластеров ( $p = 0,022$ ),  $p$ -значение не преодолевает более строгий порог в 0,01, что, в совокупности с умеренной корреляцией между числом повторов и уровнем полиморфизма ( $p = 0,34$ ), указывает на ограниченную силу взаимосвязи. Поскольку проведенный

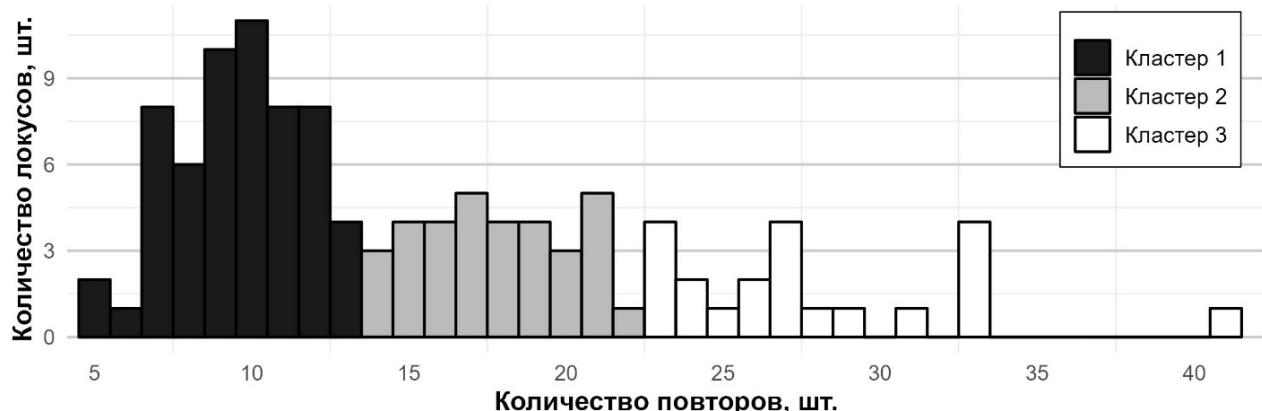


Рис. 5. Кластеризация локусов по суммарному количеству повторов методом k-средних

Fig. 5. Clustering of loci according to the total number of repeats using the k-means method

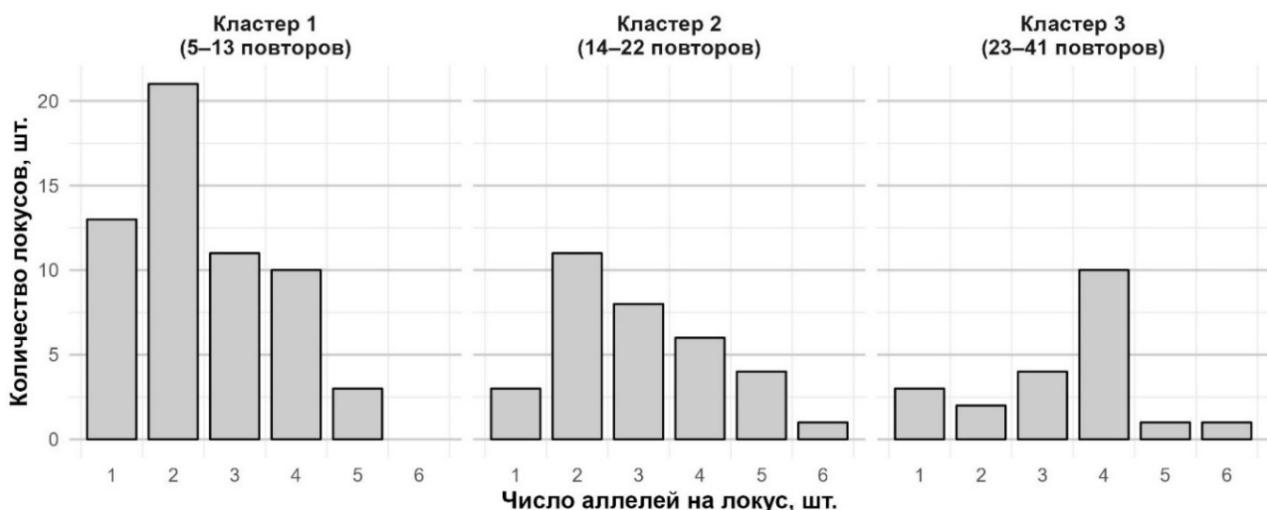


Рис. 6. Распределение локусов с различным числом аллелей в кластерах, сформированных по суммарному количеству повторов

Fig. 6. Distribution of loci with different numbers of alleles in the clusters arranged by the total number of repeats

Таблица 2. Результаты статистических тестов

Table 2. Results of statistical tests

Название теста	Сравнение кластеров	Статистика	P
Краскела – Уоллиса	Общее	$H = 8,31$	0,016*
Данна	Кластеры 1–2	$Z = -1,84$	0,132
	Кластеры 1–3	$Z = -2,68$	0,022*
	Кластеры 2–3	$Z = -1,00$	0,313

Примечание: \* – статистически значимые различия при 0,05 уровне значимости

Note: \* – statistically significant differences at the 0.05 level

анализ оценивал различия по всем значениям числа аллелей, это не позволило сделать выводы о различиях в частоте именно высокополиморфных локусов. Для решения этой задачи проведена категоризация локусов в зависимости от количества повторов: к низкополиморфным отнесены локусы с числом аллелей  $\leq 2$ , а к высокополиморфным –  $> 2$ . Полученные категориальные данные проанализировали с использованием  $\chi^2$ -теста на равенство долей между кластерами. Этот анализ также позволил учесть неравенство объема выборок. Результаты теста показали статистически значимые различия частот высокополиморфных локусов в зависимости от принадлежности к кластеру ( $\chi^2 = 7,95$ ;  $p = 0,019$ ). Множественные попарные сравнения с поправкой Холма – Бонферрони выявили значимое различие между первым и третьим кластерами ( $p = 0,039$ ), что согласуется с результатами теста Данна и уточняет, что различия между этими кластерами обусловлены разницей в частотах высоко- и низкополиморфных локусов.

В дополнение для первого и третьего кластера рассчитали размер эффекта (Cohen's  $h$ ) на основе долей высокополиморфных локусов. Полученное значение ( $h = 0,75$ ) соответствовало умеренному эффекту, близкому к большому (порог  $> 0,80$ ) (Cohen, 1988). Учитывали различия в размерах выборок кластеров ( $n$ ) посредством расчета 95-процентного доверительного интервала для разности пропорций (0,14–0,57). Поскольку интервал не включал ноль, выявленные различия являются статистически значимыми, а размер эффекта отражает их практическую применимость.

Перед тем как рекомендовать локусы с количеством повторов  $> 23$ , необходимо проанализировать баланс между частотой высокополиморфных локусов и частотой отбракованных маркеров. В данной работе маркеры исключали из анализа по трем причинам: отсутствие амплификации, неспецифичная амплификация и артефакты амплификации (статтеры, побочные пики, ярко выраженное неполное аденилирование). Первые две причины, вероятнее всего, не имеют причинно-следственной связи с длиной повтора, а являются следствием некорректной работы подобранный пары праймеров. В свою очередь, вероятность образования статтеров может расти в зависимости от количества повторов, как

случаев возможного проскальзывания ДНК-полимеразы (рис. 7).

Многие праймеры, фланкирующие микросателлитные последовательности с большим количеством повторяющихся мотивов (30, 35, 37, 39, 42, 43, 45 и 71), в данных условиях проведения ПЦР показали либо отсутствие амплификации, либо наличие артефактов SSR-анализа. На рисунке 7 наглядно продемонстрировано увеличение доли отбракованных маркеров с ростом количества повторов в локусе. Однако статистическая проверка с использованием попарных сравнений с поправкой Холма – Бонферрони не выявила значимых различий между кластерами, в том числе между первым и третьим кластером ( $p = 0,081$ ). Таким образом, не наблюдалось значимого увеличения доли отбракованных маркеров при возрастании количества повторов.

Выявленные закономерности позволяют рекомендовать в будущих исследованиях отдавать предпочтение микросателлитным локусам с количеством повторов больше 23, что потенциально может способствовать увеличению доли высокополиморфных маркеров. Поскольку данное исследование не было направлено на установление этой зависимости, не представлялось возможным определить верхний порог количества повторов, после которого негативные эффекты, такие как вероятность статтеров, могут превысить эффект повышенной информативности. Следует также учитывать, что выявленные закономерности могут варьироваться в зависимости от изучаемой культуры. Несмотря на это, было установлено, что использование количества повторов в локусе в качестве одного из критериев при отборе *in silico* может повысить среднюю информативность разрабатываемых SSR-маркеров.

Согласно литературным данным, общее количество микросателлитов в геноме подсолнечника составляет 3 543 083 и покрывает 1,72% всего генома (Avvari et al., 2020). Однако множество из них не подходят для разработки SSR-маркеров по таким причинам, как низкая сложность фланкирующих последовательностей, внутригенная локализация, моно- и динуклеотидный мотив и т. д. Мы провели отбраковку таких локусов и разработали 186 микросателлитных маркеров подсолнечника. И из них только 94 проявили себя эффективными для

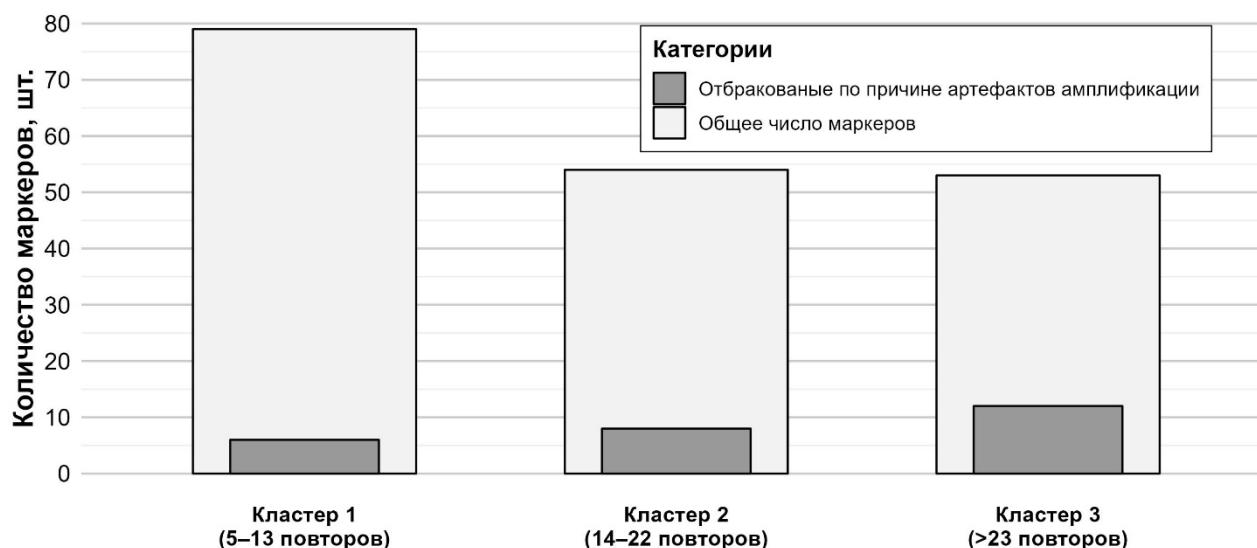


Рис. 7. Доля отбракованных маркеров от общего количества в зависимости от кластера

Fig. 7. Proportion of rejected markers from the total number depending on the cluster

дальнейшего использования в исследованиях по изучению генетического разнообразия генетических ресурсов подсолнечника и их идентификации. Согласно литературным источникам, среди заново разработанных маркеров около 70% могут быть неэффективными (Ascunce et al., 2013). Полученные нами результаты превышают этот показатель. Для части этих маркеров в предыдущих исследованиях уже подтверждена возможность объединения их в системы для мультиплексной ПЦР и показан высокий дискриминационный потенциал системы маркеров (Golovatskaya, Guchetl, 2024). Последовательности праймеров и характеристики маркеров представлены в созданной нами базе данных микросателлитных ДНК-маркеров подсолнечника (Golovatskaya et al., 2025). Дальнейшие наши исследования направлены на отбор маркеров, обладающих наибольшим дискриминационным потенциалом, с возможностью объединения в наборы для мультиплексной ПЦР и разработку технологии генотипирования подсолнечника.

### Заключение

С использованием биоинформационных методов на эталонной сборке генома подсолнечника HanXRQr2.0-SUNRISE разработаны 186 новых микросателлитных маркеров с три- и более нуклеотидами мотива с числом повторов от 3 до 97. Из них, по результатам апробации, 94 оказались полиморфными и обладали специфичностью к целевому локусу. Установлена статистически значимая, но слабая положительная корреляция между числом аллелей и количеством повторов в микросателлитном локусе. Повышение числа аллелей выражалось в смещении распределения в сторону более полиморфных локусов при числе повторов выше 23, что позволяет рассматривать этот порог как ориентир при разработке SSR-маркеров. На основании полученных данных возможен отбор наиболее информативных маркеров, формирование наборов для мультиплексной ПЦР и создание эффективной системы для генотипирования подсолнечника.

### References / Литература

- Ahmed H.G.M.D., Rizwan M., Naeem M., Khan M.A., Baloch F.S., Sun S. et al. Molecular characterization and validation of sunflower (*Helianthus annuus* L.) hybrids through SSR markers. *PLoS One*. 2022;17(5):e0267383. DOI: 10.1371/journal.pone.0267383
- Ascunce M.S., Toups M.A., Kassu G., Fane J., Scholl K., Reed D.L. Nuclear genetic diversity in human lice (*Pediculus humanus*) reveals continental differences and high inbreeding among worldwide populations. *PLoS One*. 2013;8(2):e57619. DOI: 10.1371/journal.pone.0057619
- Avvaru A.K., Sharma D., Verma A., Mishra R.K., Sowpati D.T. MSDB: a comprehensive, annotated database of microsatellites. *Nucleic Acids Research*. 2020;48(D1):D155-D159. DOI: 10.1093/nar/gkz886
- Brunel D. A microsatellite marker in *Helianthus annuus* L. *Plant Molecular Biology*. 1994;24(2):397-400. DOI: 10.1007/BF00020177
- Buchan J.C., Archie E.A., Van Horn R.C., Moss C.J., Alberts S.C. Locus effects and sources of error in noninvasive genotyping. *Molecular Ecology Notes*. 2005;5(3):680-683. DOI: 10.1111/j.1471-8286.2005.01002.x
- Camacho C., Coulouris G., Avagyan V., Ma N., Papadopoulos J., Bealer K. et al. BLAST+: architecture and applica-
- tions. *BMC Bioinformatics*. 2009;10:421. DOI: 10.1186/1471-2105-10-421
- Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. New York, NY: Routledge; 1988. DOI: 10.4324/9780203771587
- Fallah M., Paizila A., Karci H., Arab M.M., Sarikhani S., Suprun I. et al. Validation and implementation of marker-assisted selection (MAS) for the leafing date trait in Persian walnut populations from Iran. *Euphytica*. 2024;220(2):25. DOI: 10.1007/s10681-023-03281-3
- Filippi C., Aguirre N., Rivas J.G., Zubrzycki J., Puebla A., Corades D. et al. Population structure and genetic diversity characterization of a sunflower association mapping population using SSR and SNP markers. *BMC Plant Biology*. 2015;15:52. DOI: 10.1186/s12870-014-0360-x
- Golovatskaya A.V., Guchetl S.Z. Assessment of the genetic diversity of sunflower lines of VNIIMK breeding based on multiplex microsatellite analysis. *Agrarian Science*. 2024;(11):117-121. [in Russian] (Головатская А.В., Гучетль С.З. Оценка генетического разнообразия линий подсолнечника селекции ВНИИМК на основе мультиплексного микросателлитного анализа. *Аграрная наука*. 2024;(11):117-121). DOI: 10.32634/0869-8155-2024-388-11-117-121
- Golovatskaya A.V., Guchetl S.Z. Screening of microsatellite DNA markers for effective sunflower identification. *Fodder Production*. 2023;(11):48-51. [in Russian] (Головатская А.В., Гучетль С.З. Скрининг микросателлитных ДНК-маркеров для разработки эффективной системы идентификации подсолнечника. *Кормопроизводство*. 2023;(11):48-51). DOI: 10.25685/krm.2023.11.2023.007
- Golovatskaya A.V., Guchetl S.Z., Savichenko D.L., Loginova E.D., Zhudina E.I. Database of microsatellite DNA markers of sunflower (*Helianthus annuus* L.). Russian Federation; database state registration certificate number: 2025622192; 2025. [in Russian] (Головатская А.В., Гучетль С.З., Савиценко Д.Л., Логинова Е.Д., Жудина Э.И. База данных микросателлитных ДНК-маркеров подсолнечника (*Helianthus annuus* L.). Российская Федерация; свидетельство о государственной регистрации базы данных № 2025622192; 2025). URL: <https://vniimk.ru/science/bazy-dannykh-i-programmy-dlya-evm/ljrevtyns-2> [дата обращения: 22.06.2024].
- Guchetl S.Z., Golovatskaya A.V., Ramazanova S.A., Voloshko A.A. Genetic diversity of the Russian sunflower breeding lines revealed by microsatellite loci analysis. *Agricultural Science Euro-North-East*. 2023;24(2):173-186. [in Russian] (Гучетль С.З., Головатская А.В., Рамазанова С.А., Волошко А.А. Генетическое разнообразие линий подсолнечника российской селекции, выявленное с помощью анализа микросателлитных локусов. *Аграрная наука Евро-Северо-Востока*. 2023;24(2):173-186). DOI: 10.30766/2072-9081.2023.24.2.173-186
- Guichoux E., Lagache L., Wagner S., Chaumeil P., Léger P., Lepais O. et al. Current trends in microsatellite genotyping. *Molecular Ecology Resources*. 2011;11(4):591-611. DOI: 10.1111/j.1755-0998.2011.03014.x
- Hoffman J.I., Amos W. Microsatellite genotyping errors: detection approaches, common sources and consequences for paternal exclusion. *Molecular Ecology*. 2005;14(2):599-612. DOI: 10.1111/j.1365-294X.2004.02419.x
- Liu C., Zhao N., Jiang Z.C., Zhang H., Zhai H., He S.Z. et al. Analysis of genetic diversity and population structure in sweetpotato using SSR markers. *Journal of Integrative Agriculture*. 2023;22(11):3408-3415. DOI: 10.1016/j.jia.2023.02.004
- Mason A.S. SSR genotyping. In: J. Batley (ed.). *Methods in Molecular Biology*. Vol. 1245. *Plant Genotyping*. New York,

- NY: Humana Press; 2015. p.77-89. DOI: 10.1007/978-1-4939-1966-6\_6
- NIH. National Library of Medicine. National Center for Biotechnology Information: [website]. Available from: <https://www.ncbi.nlm.nih.gov> [accessed Mar. 25, 2025].
- NIH. National Library of Medicine. National Center for Biotechnology Information. Primer-BLAST: A tool for finding specific primers: [website]. Available from: <https://www.ncbi.nlm.nih.gov/tools/primer-blast> [accessed Mar. 25, 2025].
- Paniego N., Echaide M., Muñoz M., Fernández L., Torales S., Facio P. et al. Microsatellite isolation and characterization in sunflower (*Helianthus annuus* L.). *Genome*. 2002;45(1):34-43. DOI: 10.1139/g01-120
- Ramazanova S.A., Badyanov E.V., Guchetl S.Z. Validation of microsatellite markers to identify *Pl<sub>g</sub>*, *Pl<sub>g</sub>* and *Pl<sub>arg</sub>* genes that control resistance to *Plasmopara halstedii* in sunflower. *Caspian Journal of Environmental Sciences*. 2021;19(5):915-920. DOI: 10.22124/CJES.2021.5266
- Ramya K.T., Vishnuvardhan Reddy A., Sujatha M. Agromorphological and molecular analysis discloses wide genetic variability in sunflower breeding lines from USDA, USA. *Indian Journal of Genetics and Plant Breeding*. 2019;79(2):444-452. DOI: 10.31742/IJGPB.79.2.8
- Sheller M.A., Ciocirlan E., Mikhailov P.V., Kulakov S.S., Kulakova N.N., Ibe A.A. et al. Genetic diversity of Scots pine (*P. sylvestris* L.) in middle Siberia assessed by analysis of variability of chloroplast microsatellite loci. *Biosfera. Interdisciplinary Journal of Basic and Applied Sciences*. 2023;15(4):343-348. [in Russian] (Шеллер М.А. Чиокырлан Е., Михайлов П.В., Кулаков С.С., Кулакова Н.Н., Ибе А.А. и др. Генетическое разнообразие сосны обыкновенной (*P. sylvestris* L.) в средней Сибири по результатам анализа изменчивости хлоропластных микросателлитных локусов. *Биосфера. Междисциплинарный научный и прикладной журнал*. 2023;15(4):343-348). DOI: 10.24855/biosfera.v15i4.868
- Shilov I.A., Velishaeva N.S., Aniskina Yu.V., Kolobova O.S., Shalaeva T.V., Borisenko O.M. et al. Genetic identification of *Helianthus annuus* L. sunflower lines and hybrids based on multiplex microsatellite analysis. *Achievements of Science and Technology of AIC*. 2023;37(1):10-15. [in Russian] (Шилов И.А., Велишаева Н.С., Анискина Ю.В., Колобова О.С., Шалаева Т.В., Борисенко О.М. и др. Генетическая идентификация линий и гибридов подсолнечника *Helianthus annuus* L. на основе мультиплексного микросателлитного анализа. *Достижения науки и техники АПК*. 2023;37(1):10-15). DOI: 10.53859/02352451\_2023\_37\_1\_10
- Tang S., Yu J.K., Slabaugh B., Shintani D.K., Knapp S.J. Simple sequence repeat map of the sunflower genome. *Theoretical and Applied Genetics*. 2002;105(8):1124-1136. DOI: 10.1007/s00122-002-0989-y
- The R Project for Statistical Computing: [website]. Available from: <https://www.r-project.org> [accessed Mar. 25, 2025].
- Topu M., Sesiz U., Bektaş H., Toklu F., Özkan H. Next-generation-sequencing-based simple sequence repeat (SSR) marker development and linkage mapping in lentil (*Lens culinaris* L.). *Life (Basel)*. 2023;13(7):1579. DOI: 10.3390/life13071579
- Verbiest M., Maksimov M., Jin Y., Anisimova M., Gymrek M., Bilgin S.T. Mutation and selection processes regulating short tandem repeats give rise to genetic and phenotypic diversity across species. *Journal of Evolutionary Biology*. 2023;36(2):321-336. DOI: 10.1111/jeb.14106
- Wang X., Wang L. GMATA: an integrated software package for genome-scale SSR mining, marker development and viewing. *Frontiers in Plant Science*. 2016;7:1350. DOI: 10.3389/fpls.2016.01350
- Wickham H. *ggplot2: Elegant graphics for data analysis*. New York, NY: Springer-Verlag; 2016. DOI: 10.1007/978-0-387-98141-3
- Yadav S., Carvalho J., Trujillo I., Prado M. Microsatellite markers in olives (*Olea europaea* L.): Utility in the cataloging of germplasm, food authenticity and traceability studies. *Foods*. 2021;10(8):1907. DOI: 10.3390/foods10081907

### Информация об авторах

**Саида Заурбиеевна Гучетль**, кандидат биологических наук, заведующая лабораторией, Федеральный научный центр «Всероссийский научно-исследовательский институт масличных культур имени В.С. Пустовойта», 350038 Россия, Краснодар, ул. Филатова, 17, saida.guchetl@mail.ru, <https://orcid.org/0000-0002-2193-5230>

**Анна Владимировна Головатская**, младший научный сотрудник, Федеральный научный центр «Всероссийский научно-исследовательский институт масличных культур имени В.С. Пустовойта», 350038 Россия, Краснодар, ул. Филатова, 17, annamoon11@gmail.com, <https://orcid.org/0000-0001-8355-3150>

**Дмитрий Леонидович Савиченко**, научный сотрудник, Федеральный научный центр «Всероссийский научно-исследовательский институт масличных культур имени В.С. Пустовойта», 350038 Россия, Краснодар, ул. Филатова, 17, d\_savichenko@mail.ru, <https://orcid.org/0000-0001-8457-2919>

**Елизавета Дмитриевна Логинова**, младший научный сотрудник, Федеральный научный центр «Всероссийский научно-исследовательский институт масличных культур имени В.С. Пустовойта», 350038 Россия, Краснодар, ул. Филатова, 17, asdfghjklis@mail.ru, <https://orcid.org/0009-0002-5136-1028>

**Эльвина Исмаиловна Жудина**, аналитик, Федеральный научный центр «Всероссийский научно-исследовательский институт масличных культур имени В.С. Пустовойта», 350038 Россия, Краснодар, ул. Филатова, 17, Elvinaalieva98.98@mail.ru, <https://orcid.org/0000-0003-0244-0926>

### Information about the authors

**Saida Z. Guchetl**, Cand. Sci. (Biology), Head of a Laboratory, V.S. Pustovoit All-Russian Research Institute of Oil Crops, 17 Filatova St., Krasnodar 350038, Russia, saida.guchetl@mail.ru, <https://orcid.org/0000-0002-2193-5230>

**Anna V. Golovatskaya**, Associate Researcher, V.S. Pustovoit All-Russian Research Institute of Oil Crops, 17 Filatova St., Krasnodar 350038, Russia, annamoon11@gmail.com, <https://orcid.org/0000-0001-8355-3150>

**Dmitrii L. Savichenko**, Researcher, V.S. Pustovoit All-Russian Research Institute of Oil Crops, 17 Filatova St., Krasnodar 350038, Russia, d\_savichenko@mail.ru, <https://orcid.org/0000-0001-8457-2919>

**Elizaveta D. Loginova**, Associate Researcher, V.S. Pustovoit All-Russian Research Institute of Oil Crops, 17 Filatova St., Krasnodar 350038, Russia, asdfghjklis@mail.ru, <https://orcid.org/0009-0002-5136-1028>

**Elvina I. Zhudina**, Analyst, V.S. Pustovoit All-Russian Research Institute of Oil Crops, 17 Filatova St., Krasnodar 350038, Russia, Elvinaalieva98.98@mail.ru, <https://orcid.org/0000-0003-0244-0926>

**Вклад авторов:** Гучетль С.З. – участие в разработке концепции исследования; написание оригинального текста статьи. Головатская А.В. – участие в разработке концепции исследования; разработка и апробирование маркеров; подготовка и редактирование рукописи. Савиценко Д.Л. – участие в разработке концепции исследования; разработка алгоритма создания молекулярных маркеров; разработка и апробирование маркеров; статистический анализ данных; подготовка и редактирование рукописи. Логинова Е.Д. – разработка и апробирование молекулярных маркеров. Жудина Э.И. – разработка молекулярных маркеров.

**Contribution of the authors:** Guchetl S. Z. – participation in the development of the research concept; writing the original text of the article. Golovatskaya A. V. – participation in the development of the research concept; development and testing of markers; preparation and editing of the manuscript. Savichenko D. L. – participation in the development of the research concept; preparation and editing of the manuscript; development of an algorithm for creating molecular markers; development and testing of molecular markers; statistical data analysis. Loginova E. D. – development and testing of molecular markers. Zhudina E. I. – development of molecular markers.

**Конфликт интересов:** авторы заявляют об отсутствии конфликта интересов.

**Conflict of interests:** the authors declare no conflicts of interests.

Статья поступила в редакцию 22.07.2025; одобрена после рецензирования 08.10.2025; принятая к публикации 22.10.2025. The article was submitted on 22.07.2025; approved after reviewing on 08.10.2025; accepted for publication on 22.10.2025.