

ОТБОР МОЛЕКУЛЯРНЫХ МАРКЕРОВ ДЛЯ ГЕНЕТИЧЕСКОЙ ПАСПОРТИЗАЦИИ *TRITICUM AESTIVUM*

М. А. Самарина^{1,2}, младший научный сотрудник лаборатории генетических технологий и молекулярного сопровождения селекции зерновых и зернобобовых культур; младший научный сотрудник лаборатории генетики и пребридинга, Samarina.homa@yandex.ru, ORCID: 0000-0001-9102-4208;

Д. С. Ульянов^{1,2}, младший научный сотрудник лаборатории цифрового фенотипирования для селекции растений; младший научный сотрудник лаборатории генетики и пребридинга, uldas1508@gmail.com, ORCID ID: 0000-0002-5880-5931;

Т. Д. Мохов^{1,2}, младший научный сотрудник лаборатории цифрового фенотипирования для селекции растений; младший научный сотрудник лаборатории генетики и пребридинга, Timmikh@yandex.ru, ORCID ID: 0009-0005-8893-7424;

Я. С. Меглицкая^{1,2}, младший научный сотрудник лаборатории цифрового фенотипирования для селекции растений; младший научный сотрудник лаборатории генетики и пребридинга, yanameg20@gmail.com, ORCID ID: 0009-0007-1529-2340;

П. Ю. Крупин^{1,2}, кандидат биологических наук, заведующей лабораторией генетических технологий и молекулярного сопровождения селекции зерновых и зернобобовых культур; младший научный сотрудник лаборатории генетики и пребридинга, iab@iab.ac.ru, ORCID ID: 0000-0001-6858-3941;

Г. И. Карлов¹, доктор биологических наук, академик РАН, профессор, директор, karlov@iab.ac.ru, ORCID ID: 0000-0002-9016-103X;

П. Н. Харченко¹, доктор биологических наук, академик РАН, профессор, заведующий отделом клеточной и геномной инженерии; научный руководитель, kharchenko@iab.ac.ru, ORCID ID: 0009-0008-7054-8801;

С. И. Воронов², доктор биологических наук, член-корреспондент РАН, профессор, директор, vsi08@mail.ru, ORCID ID: 0000-0002-8103-3909;

Н. В. Давыдова², доктор сельскохозяйственных наук, заведующая лабораторией селекции и первичного семеноводства яровой пшеницы, davnat58@yandex.ru;

М. Г. Дивашук^{1,2}, кандидат биологических наук, ведущий научный сотрудник, заведующий лабораторией генетики и пребридинга; руководитель Курчатовского геномного центра, divashuk@gmail.com, ORCID ID: 0000-0001-6221-3659

¹Федеральное государственное бюджетное научное учреждение «Всероссийский научно-исследовательский институт сельскохозяйственной биотехнологии» (ФГБНУ ВНИИСБ), 127550, г. Москва, ул. Тимирязевская, д. 42; e-mail: iab@iab.ac.ru

²Федеральное государственное бюджетное научное учреждение «Федеральный Исследовательский Центр «Немчиновка», 143026, Московская область, г.п. Одинцово, р. п. Новоивановское, ул. Агрохимиков, д. 6; e-mail: mosniish@yandex.ru

Согласно закону «О семеноводстве» для всех сортов или гибридов, включаемых в Госреестр, предусматривается выдача генетических паспортов, а также создание перечня родов и видов растений, производство которых направлено на обеспечение продовольственной безопасности. Однако использование морфологических признаков не позволяет различать генетически близкие образцы, выявлять скрытую генетическую изменчивость и обеспечивать контроль однородности исходного материала. В связи с этими ограничениями целью данной работы являлась разработка подходов для отбора молекулярно-генетических микросателлитных маркеров (SSR-маркеров) для дифференциации сортов мягкой пшеницы. В соответствии с заявленной целью нами был проведен обзор литературы в части методов генетической паспортизации и оценки разнообразия яровой мягкой пшеницы на основе полиморфизма микросателлитных локусов, сформирован набор сборок генома *Triticum aestivum* из базы данных NCBI Gen Bank. На основе литературных данных были выбраны наиболее полиморфные SSR-маркеры с использованием созданного нами алгоритма. Проведенный с помощью биоинформатических методов анализ баз данных SSR-маркеров в геноме мягкой пшеницы позволил создать минимальный дискриминирующий набор из 20 маркеров, которые способны детектировать 419 различных аллелей у *Triticum aestivum*. Полученные результаты закладывают необходимую теоретическую основу для проведения дальнейших практических исследований.

Ключевые слова. *Triticum aestivum* L., SSR-маркеры, микросателлитные маркеры, мягкая пшеница, генотипирование.

Для цитирования: Самарина М. А., Ульянов Д. С., Мохов Т. Д., Меглицкая Я. С., Крупин П. Ю., Карлов Г. И., Харченко П. Н., Воронов С. И., Давыдова Н. В., Дивашук М. Г. Отбор молекулярных маркеров для генетической паспортизации *Triticum aestivum* // Зерновое хозяйство России. 2024. Т. 16, № 4. С. 33–40. DOI: 10.31367/2079-8725-2024-93-4-33-40.



SELECTION OF MOLECULAR MARKERS FOR GENETIC CERTIFICATION OF *TRITICUM AESTIVUM*

M. A. Samarina^{1,2}, junior researcher of the laboratory for genetic technologies and molecular support of breeding of grain crops and legumes; junior researcher of the laboratory for genetics and pre-breeding, Samarina.homa@yandex.ru, ORCID ID: 0000-0001-9102-4208;

D. S. Uliyanov^{1,2}, junior researcher of the laboratory for digital phenotyping for plant breeding; junior researcher of the laboratory for genetics and pre-breeding, uldas1508@gmail.com, ORCID ID: 0000-0002-5880-5931;

T. D. Mokhov^{1,2}, junior researcher of the laboratory for digital phenotyping for plant breeding; junior researcher of the laboratory for genetics and pre-breeding, Timmokh@yandex.ru, ORCID ID: 0009-0005-8893-7424;

Ya. S. Metlitskaya^{1,2}, junior researcher of the laboratory for digital phenotyping for plant breeding; junior researcher of the laboratory for genetics and pre-breeding, yanameg20@gmail.com, ORCID ID: 0009-0007-1529-2340;

P. Yu. Krupin^{1,2}, Candidate of Biological Sciences, head of the laboratory for genetic technologies and molecular support of breeding of grain crops and legumes; junior researcher of the laboratory for genetics and pre-breeding, iab@iab.ac.ru, ORCID ID: 0000-0001-6858-3941;

G. I. Karlov¹, Doctor of Biological Sciences, academician of RAS, professor, head, karlov@iab.ac.ru, ORCID ID: 0000-0002-9016-103X;

P. N. Kharchenko¹, Doctor of Biological Sciences, academician of RAS, professor, head of the department of cell and gene engineering; scientific supervisor, kharchenko@iab.ac.ru;

S. I. Voronov², Doctor of Biological Sciences, correspondent member of the RAS, professor, head, vsi08@mail.ru;

N. V. Davydova², Doctor of Agricultural Sciences, head of the laboratory for breeding and primary seed production of spring wheat, davnat58@yandex.ru;

M. G. Divashuk^{1,2}, Candidate of Biological Sciences, leading researcher, head of the laboratory for genetics and pre-breeding; head of the Kurchatov Genome Center, e-mail: divashuk@gmail.com, ORCID ID: 0000-0001-6221-3659

¹FSBSI "All-Russian Research Institute of Agricultural Biotechnology" (FSBSI ARRIAB), 127550, Moscow, Timiryazevskaya Str., 42; e-mail: iab@iab.ac.ru;

²FSBSI "Federal Research Center "Nemchinovka" (FRC "Nemchinovka"), 143026, Moscow region, Odintsovsky region, Odintsovsky district, v. of Novoivanovskoe, Agrokhimikov Str., 6; e-mail: mosniish@yandex.ru

According to the Law "On Seed Production", all varieties or hybrids included in the state register shall be provided for genetic passports, as well as making a list of plants' species, the production of which is aimed at ensuring food security. However, the use of morphological characteristics does not allow distinguishing genetically similar samples, identifying hidden genetic variability and ensuring control of the homogeneity of the initial material. In connection with these limitations, the purpose of the current work was to develop approaches for the selection of molecular genetic microsatellite markers (SSR markers) to differentiate common wheat varieties. In accordance with the purpose, there has been conducted a literature review regarding methods for genetic certification and evaluation of the diversity of spring common wheat based on polymorphism of microsatellite loci, and there has been generated a set of *Triticum aestivum* genome assemblies from the NCBI Gen Bank. Based on the literature data, there have been selected the most polymorphic SSR markers using the invented algorithm. The analysis of databases of SSR markers in the genome of common wheat using bioinformatics methods allowed establishing a minimal discriminatory set of 20 markers that can detect 419 different alleles in *Triticum aestivum*. The results obtained can become the necessary theoretical foundation for further practical research.

Keywords. *Triticum aestivum* L., SSR-маркеры, microsatellite markers, common wheat, genotyping.

Введение. Пшеница играет важную роль в обеспечении продовольствием населения планеты, а также является одной из самых распространенных сельскохозяйственных культур в мире. В питании человека мягкая пшеница выступает в роли источника крахмала, белка, витаминов и минеральных элементов.

Снижение роста урожая продовольственных культур в мире приводит к необходимости развития производства пшеницы, в первую очередь создания новых сверхпродуктивных и устойчивых к различным стрессам сортов. Для этого необходимо производить поиск и включение в селекционный процесс новых перспективных источников хозяйственно ценных признаков из мирового генофонда.

Паспортизация и идентификация сортов и гибридов является актуальной задачей в современной селекционной работе. Их иденти-

фикация позволяет предотвратить нарушение авторских прав при создании и реализации сортов и гибридов различных сельскохозяйственных культур.

Однако зачастую идентификация сортов, линий и гибридов с использованием морфологических признаков имеет низкую эффективность из-за невозможности дифференцировать близкие по происхождению организмы и выявить скрытую генетическую изменчивость. Кроме того, число морфологических признаков ограничено, и они не всегда стабильны.

Создание базы сортов с известной генетической структурой и возможность их идентификации позволяют использовать в селекции родительские формы с лучшей урожайностью и качеством продукции, а также устойчивостью к болезням и абиотическим стрессам.

Наличие генетического полиморфизма является основой для классификации и идентификации сортов, что, в свою очередь, позволяет не только повысить эффективность селекции и семеноводства, но и улучшить технологичность производства, связанного с переработкой сельскохозяйственной продукции (Митрофанова и др., 2009).

В качестве инструмента могут выступать микросателлитные маркеры, которые широко используются для исследования генетического разнообразия. Микросателлиты (простые повторяющиеся последовательности – SSR) – короткие некодирующие последовательности ДНК длиной 1–6 нуклеотидов, повторяющиеся до нескольких десятков раз и расположенные в разных частях генома. Они получили широкое распространение из-за их простоты, высокого уровня полиморфизма (Wang et al., 2014), высокой воспроизводимости, кодоминантной наследуемости, равномерного распределения по хромосомам при правильном подборе. В последние 40 лет микросателлитный анализ является популярной темой научных работ, так как находит применение не только в паспортизации сельскохозяйственных животных, но и является общепринятым методом генетической идентификации человека.

Важным преимуществом использования SSR-маркеров является возможность различать внутри полиморфного сорта отдельные биотипы. Различные биотипы обеспечивают экологическую пластичность сорта, однако при этом по морфологическим признакам они могут не отличаться, что усложняет задачу по их идентификации. В связи с этим использование микросателлитных маркеров может упростить оценку селекционных достижений на ООС, особенно это важно для культур, идентифицировать которые традиционными методами затруднительно, например, винограда (Фомина и др., 2014). Известно, что SSR подходят для анализа генетического разнообразия и идентификации генотипов у самоопыляемых видов, таких как пшеница, поэтому SSR-маркеры ранее уже были использованы для оценки разнообразия в генетическом материале пшеницы элитных линий, сортов.

Кроме того, SSR-маркеры представляют собой ценный инструмент для анализа генетиче-

ского разнообразия растений, включая сою, ячмень, пшеницу и др.

SSR-маркеры зарекомендовали себя как надежный и эффективный способ характеристики генетических ресурсов мягкой пшеницы (Пискарев и Бойко, 2015).

Согласно базе данных «Triticum aestivum SSR» (TaSSRDb) для гексаплоидного генома пшеницы сегодня известно более 450 тысяч SSR-маркеров. Этот геномный ресурс можно использовать для картирования сцепления, анализа разнообразия, идентификации сортов, тестирования на отличимость, однородность и стабильность (ООС) и контроля чистоты семян. Такой большой ассортимент требует применения биоинформатического анализа при подборе минимального дискриминирующего набора.

Схема работы с микросателлитными маркерами, описанная в большинстве работ, включая Фомина и др. (2014), подразумевает многоэтапное пошаговое исследование образцов с использованием специфических праймеров и дальнейшим биоинформатическим анализом. Однако такая схема может быть неэффективна из-за трат времени на постановку реакций и средств, требующихся в том числе и на синтез специфичных праймеров.

Поэтому для повышения эффективности генотипирования является актуальной и практически востребованной задачей создание минимального дискриминирующего набора маркеров с высокой степенью полиморфизма.

Таким образом, целью нашей работы является разработка алгоритма, который позволит эффективно анализировать большие объемы данных по микросателлитным маркерам и отбирать наиболее информативные.

Материалы и методы исследований. Исходные данные для отбора минимального дискриминирующего набора SSR-маркеров. В литературную подборку вошли как отечественные статьи (Митрофанова, 2012), так и зарубежные исследования (Wang et al., 2014) по данной теме, которые стали основными источниками анализируемых локусов. Их информативность оценивалась на 17 предварительно найденных аннотированных сборках генома *Triticum aestivum* из базы NCBI GenBank (National Center for Biotechnology Information) (табл. 1).

Таблица 1. Геномные сборки *Triticum aestivum*
Table 1. Genome assemblies of *Triticum aestivum*

Код сборки в NCBI	Название сборки	Культурный сорт	Общая длина последовательности	Уровень сборки	Дата публикации сборки
GCA_018294505.1	IWGSC CS RefSeq v2.1	Chinese Spring	14566502436	Chromosome	2021-05-06
GCA_910594105.1	Tae_Kariega_v1		14677204660	Chromosome	2021-12-19
GCA_937894285.1	Triticum_aestivum_Renan_v2.1		14195643615	Chromosome	2022-05-03
GCA_907166925.1	wheat_cv_fielder_v1_assembly	Fielder	14702880414	Chromosome	2021-07-14
GCA_918797515.1	wheat_cv_attraktion_v1	Attraktion	14679377615	Chromosome	2022-06-24
GCA_025895885.1	ASM2589588v1	Aikang58	14752023501	Chromosome	2022-10-31
GCA_903994165.1	10wheat_assembly_spelt	PI190962 (Spelt)	14439018734	Chromosome	2020-08-15
GCA_903993985.1	10wheat_assembly_arinaLrFor	ArinaLrFor	14645464107	Chromosome	2020-08-06

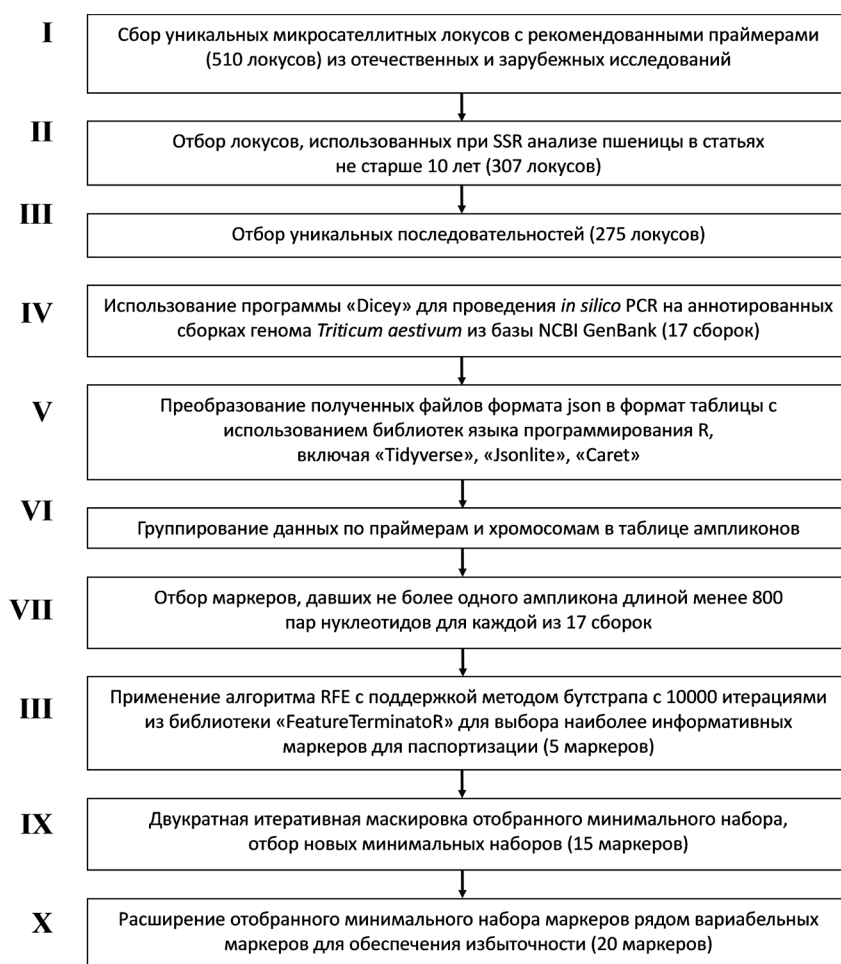
Продолжение табл. 1

Код сборки в NCBI	Название сборки	Культурный сорт	Общая длина последовательности	Уровень сборки	Дата публикации сборки
GCA_903994185.1	10wheat_assembly_sy_mattis	SY Mattis	14884612126	Chromosome	2020-08-20
GCA_903993975.1	10wheat_assembly_lancer	LongReach Lancer	14281318419	Chromosome	2020-08-06
GCA_903994175.1	10wheat_assembly_mace	Mace	14350847906	Chromosome	2020-08-14
GCA_904066035.1	10wheat_assembly_norin61	Norin 61	14910390598	Chromosome	2020-09-03
GCA_903994155.1	10wheat_assembly_stanley	CDC Stanley	14463524036	Chromosome	2020-08-20
GCA_903994195.1	10wheat_assembly_julius	Julius	14385030226	Chromosome	2020-08-15
GCA_903995565.1	10wheat_assembly_landmark1	CDC Landmark	14433190783	Chromosome	2020-08-22
GCA_903993795.1	10wheat_assembly_jagger	Jagger	14538313898	Chromosome	2020-08-06
GCA_951799155.1	NIAB Elite MAGIC Alchemy	Alchemy	15334867051	Chromosome	2023-06-29

Программное обеспечение для биоинформатического анализа маркеров. Собранная нами из литературы (Wang et al., 2014; Митрофанова, 2012) база маркеров содержала 275 уникальных локусов с праймерами и использовалась для оценки информативности маркеров посредством виртуальной симуляции ПЦР (*in silico* ПЦР) с использованием программы «Dicey» (Rausch et al., 2020) на найденных сборках (17 сборок) генома *Triticum aestivum*. Полученные данные были обработаны с использованием библиотек R, включая «Tidyverse» (Wickham et al., 2019), «Jsonlite», «Caret» (Khun et al., 2020) и «FeatureTerminator» (Hutson, 2023). Отбор маркеров проводили

методами Recursive Feature Elimination (RFE) с бутстрапом с 10000 итерациями в реализации библиотеки «FeatureTerminator». Программно отбор был реализован с использованием языков программирования Bash и R для автоматизации отбора. Для визуализации результатов использовали библиотеку ggplot.

Результаты и их обсуждение. Для упрощения биоинформатического анализа больших массивов существующих на сегодняшний день SSR-маркеров мягкой пшеницы и отбора минимального дискриминирующего набора нами был разработан собственный алгоритм анализа, реализованный на языках программирования Bash и R (см. рисунок).



Алгоритм отбора дополненного минимального набора SSR маркеров
Algorithm for selecting an augmented minimum set of SSR markers

Он представляет собой набор этапов и инструментов, расположенных последовательно, ускоряющих и автоматизирующих процесс анализа, обрабатывающих объемы данных и представляющих собой быстрый и воспроизводимый способ получения результатов. Применение алгоритма в анализе молекулярных маркеров значительно упрощает и ускоряет исследование и позволяет сосредоточиться на интерпретации результатов. Результатом выполнения алгоритма является набор праймеров для отобранных маркеров, оцененных с помощью виртуальной симуляции ПЦР. Полный код алгоритма с комментариями, содержащими подробности про различные его этапы, содержится на сайте <https://github.com/Stathmin/minimal-marker>.

Этап I разработки набора дискриминирующих маркеров заключался в сборе и анализе существующих SSR-маркеров из литературных источников. Последующие этапы осуществляются с использованием командной строки с помощью авторского кода. Нами было отобрано 510 маркеров с рекомендованными праймерами.

Наборы SSR-маркеров могут терять свою актуальность из-за изменения их структуры, поэтом на этапе II нами были отобраны маркеры, использованные для SSR анализа пшеницы в статьях не старше 10 лет. В итоге был получен набор из 307 маркеров.

Так как маркеры в различных источниках могут дублироваться, на этапе III для дальнейшего анализа мы отобрали только уникальные последовательности. В результате после отсеивания повторяющихся маркеров осталось 275 наиболее полиморфных SSR-маркеров из различных литературных источников.

На этапе IV для оценки работы маркеров был сформирован набор из 17 аннотированных сборок генома *Triticum aestivum* из базы данных NCBI Gen Bank (табл. 1). В связи с тем, что лабораторная оценка большого количества маркеров на наборе из множества сортов требует больших затрат труда, была проведена компьютерная симуляция ПЦР (*in silico* ПЦР). Результаты были получены в виде файлов формата json.

Формат json требует установки дополнительных библиотек для чтения программы, написанными на языке R. Поэтому для обработки полученных файлов на этапе V мы использо-

вали библиотеки языка программирования R. С помощью этих библиотек собранные данные были преобразованы в удобный формат csv.

На VI этапе для систематизации данных маркеры были сгруппированы по праймерам и хромосомам в полученных таблицах ампликонов. Это позволило упростить дальнейший анализ данных.

Разработанный набор SSR-маркеров в будущем может быть использован для проведения фрагментного анализа. Размер ампликона для фрагментного анализа не должен превышать 1200 нуклеотидов. Поэтому на этапе VII отбирались только те маркеры, которые дали не более одного ампликона длиной менее 800 пар нуклеотидов для каждой из 17 сборок.

Для того чтобы дискриминирующий набор праймеров содержал как можно меньшее количество маркеров, требуется отобрать наиболее информативные из них. Для этого на этапе VIII применялся алгоритм RFE (recursive feature elimination) с поддержкой методом бутстрапа с 10000 итерациями из библиотеки «FeatureTerminator». На данном этапе был получен минимальный набор маркеров для паспортизации (5 маркеров), который успешно различал все 17 сортов.

Чтобы обеспечить возможность паспортизации теоретических генотипов, которые не представлены в открытом доступе, на этапе IX проводился отбор новых минимальных наборов. Уже отобранные маркеры подверглись двукратной итеративной маскировке для исключения из анализа. Было дополнительно отобрано 10 информативных маркеров.

Так как зачастую при проведении ПЦР могут происходить ошибки, связанные с несрабатыванием маркеров, необходимо расширять минимальный дискриминирующий набор рядом высокополиморфных маркеров для обеспечения избыточности. Для этого на этапе X отобранный набор маркеров был дополнен 5 маркерами.

Разработанный нами набор маркеров отличается высоким уровнем полиморфизма каждого маркера, а также высокой воспроизводимостью и равномерным расположением по хромосомам. 20 маркеров расположены на 15 хромосомах и представлены во всех субгеномах *Triticum aestivum*, они могут детектировать 419 различных аллелей (табл. 2).

Таблица 2. Микросателлитные локусы, используемые для формирования минимального дискриминирующего набора SSR-маркеров
Table 2. Microsatellite loci used to form a minimal discriminatory set of SSR markers

Локус	Хромосома	Координата в сборке генома, bp	Структура повтора	Размер ампликона, bp	Количество аллелей	PIC
SSR1	1B	78879608-78879626	(GA)18	96	16	0.85
SSR2	4B	455690490-455690513	(CA)28	110	21	0.8
SSR3	1A	466753490-466753506	(CA)21	117	18	0.8
SSR4	3A	455690490-455690513	(CA)18	127	11	0.78
SSR5	5A	472158367-472158388	(GA)26	129	23	0.87
SSR6	7B	266463-266485	(CA)17	130	11	0.7
SSR7	6D	101438006-101438028	(CT)16	131	13	0.8

Продолжение табл. 2

Локус	Хромосома	Координата в сборке генома, bp	Структура повтора	Размер ампликона, bp	Количество аллелей	PIС
SSR8	1A	478529771-478529791	(GA)20	138	25	0.78
SSR9	2D	515210168-515210190	(GA)27	140	40	0.94
SSR10	2B	112251859-112251879	(CT)11(CA)18	152	17	0.87
SSR11	3D	110040859-110040878	(CT)26	153	25	0.92
SSR12	7B	171733011-171733032	(GA)2GC(GA)33	181	24	0.9
SSR13	2D	22183434-22183455	(CT)21	184	21	0.75
SSR14	3A	721870736-721870756	(CT)16(CA)13	191	14	0.66
SSR15	2B	49239665-49239682	(GT)30	193	17	0.87
SSR16	5D	8864559-8864579	(CT)22	204	17	0.9
SSR17	2A	712796006-712796026	(GA)37	226-230	34	0.9
SSR18	5B	8249313-8249332	(CT)16(CA)20	241	19	0.9
SSR19	7D	52079448-52079626	(CT)24	256	20	0.92
SSR20	3B	70079328-70079626	(GA)27	376	33	0.9

За последние годы по всему миру проведено значительное количество исследований разнообразия мягкой пшеницы с использованием SSR-маркеров, затрагивающих сотни различных сортов (Клыков и др., 2017; Адылова А. Т., 2018; Пискарев и Бойко, 2015; Портников И. В. и др., 2020). Результаты исследований показывают стабильную высокую разрешающую способность SSR-маркеров в оценке генетического разнообразия.

На сегодняшний день доступны как сотни хорошо изученных SSR-маркеров, так и тысячи потенциально новых маркеров, которые могут быть использованы для различных целей, включая картирование сцепления, анализ разнообразия, идентификацию сорта. Множество проведенных исследований (Клыков и др., 2017; Адылова, 2018; Пискарев и Бойко, 2015; Портников и др., 2020; Гучетль и Челюстникова, 2020; Колобова и др., 2017) подтверждает высокую разрешающую способность и стабильность SSR-маркеров, а также показывает большой научный интерес и возможность дальнейшего использования полученных с помощью SSR-маркеров данных в селекции.

В настоящее время онлайн доступны данные по более чем 1700 последовательностям SSR-маркеров, которые были охарактеризованы и локализованы на хромосомах для 27 доступных сборок пшеницы из базы данных GrainGens (<https://wheat.pw.usda.gov>). Их высокая информативность, равномерное распределение по всем геномам пшеницы и удобство использования показывают, что SSR-маркеры являются полезным инструментом, что подтверждает ценность проведенного исследования. Полученные нами результаты позволяют не только актуализировать данные по микросателлитным маркерам, но и предлагают минимальный набор маркеров для генетической дифференциации сортов. Кроме того, предложенный нами биоинформатический алгоритм должен упростить анализ данных по SSR-маркерам для важных сельскохозяйственных культур.

Развитие технологии и использование SSR-маркеров для изучения мягкой пшеницы яв-

ляется актуальной задачей и темой многих исследований в мировом научном сообществе. В статье Адылова и др. (2018) описана разработка набора из 36 SSR-маркеров, которые позволили различить 141 аллель у сортов *Triticum aestivum*. Оценка генетического разнообразия белорусских и зарубежных сортов гексаплоидной пшеницы, описанная в работе Фомина и др. (2014), потребовала использования 16 различных SSR-маркеров, которые располагаются на разных хромосомах пшеницы. В результате проведенного исследования был идентифицирован 91 аллель и обнаружено различие в уровнях полиморфизма между сортами зарубежной и белорусской селекции, а также между яровыми и озимыми сортами пшеницы. Наше исследование предлагает более широкий спектр маркеров, однако при этом нам удалось идентифицировать гораздо большее количество аллелей, что говорит о повышенном уровне полиморфизма, разработанного нами набора маркеров.

Другие исследования (Митрофанова и др., 2009) сообщают, что для оценки генетического разнообразия 32 сортов мягкой пшеницы из Узбекистана было необходимо провести разработку 36 SSR-маркеров. Такой набор маркеров может быть затруднителен в обработке, поэтому предложенная нами схема отбора микросателлитных маркеров может помочь получить более полиморфные маркеры, которые смогут получать минимальные дискриминирующие наборы с меньшим количеством маркеров, что потенциально упростит их практическое использование при генотипировании большого количества сортов, а также повысит воспроизводимость результатов исследований. Анализ исследований заставляет задуматься о различиях в достаточном количестве SSR-маркеров для дифференциации сортов, вызванных степенью полиморфизма локусов и зависящих от ряда эволюционно-географических признаков.

Важным вопросом также является сохранение актуальности разработанных наборов SSR-маркеров для паспортизации в связи с активными выпадами либо вставками повто-

ров. Разработанный в статье Wang et al. (2014) метод оценки стабильности сортов пшеницы с использованием генотипирования SSR-маркерами дает возможность прогнозировать стабильность маркеров из оценки гомозиготности. Кроме того, исследования (Митрофанова и др., 2009; Wang et al., 2014) подтверждают надежность разработанных наборов маркеров для паспортизации сортов. Приведенные данные позволяют сделать вывод о стабильности паспортизации с использованием SSR-маркеров.

Как показывает практика, SSR-маркеры часто применяются для оценки популяций и сортов различных сельскохозяйственных растений на наличие генов хозяйственно ценных признаков. В статье Haque et al. (2020) описана разработка 13 молекулярных маркеров на гены засухоустойчивости пшеницы, их использовали для определения 26 генотипов. Также была подтверждена связь SSR-маркеров с генами хозяйственно ценных признаков у культурных растений, таких как *Linum usitatissimum* (Гучетль и Челюстникова, 2020) и *Solanum tuberosum* (Колобова и др., 2017).

Применение разработанного нами алгоритма дает возможность с высокой эффективностью разрабатывать наборы маркеров на гены различных признаков. Кроме того, с помощью нашего алгоритма можно создавать наборы SSR-маркеров для оценки биоразнообразия,

а также для подбора родительских пар удавленных форм для скрещивания. Вместе с тем применение биоинформатической оценки праймеров на эти маркеры позволит оценить характеристики их работы, что в свою очередь упростит поиск новых ассоциаций в популяциях и сортах растений. Разработанные нами методы биоинформатического анализа данных и подобранный с использованием этих методов набор SSR-маркеров соответствуют современным требованиям генотипирования растений и имеют высокую практическую ценность.

Выводы. Полученные в данной статье результаты могут быть использованы для дальнейшей разработки подходов к паспортизации мягкой пшеницы и других ценных культур на основе SSR-маркеров. Полученный набор молекулярных маркеров может послужить основой для разработки систем защиты результатов интеллектуальной собственности селекционеров. С использованием биоинформатических инструментов и алгоритмов, разработанных в данном исследовании, возможно проведение аналогичных разработок для других ценных сельскохозяйственных культур.

Финансирование. Работа выполнена в рамках Государственного задания FGGE-2023-001 «Разработка подходов для генетической паспортизации яровой мягкой пшеницы на уровне геномного разнообразия».

Библиографические ссылки

1. Адылова А. Т., Норбеков Ж. К., Хуршут Э. Э., Никитина Е. В., Кушанов Ф. Н. SSR-анализ геномной ДНК перспективных сортов мягкой озимой пшеницы узбекстанской селекции // Вавиловский журнал генетики и селекции. 2018. Т. 22, № 6. С. 634–639. DOI: 10.18699/VJ18.404
2. Гучетль С. З., Челюстникова Т. А. Генотипирование сортов масличного льна с использованием системы микросателлитных ДНК-маркеров // Аграрная наука Евро-Северо-Востока. 2020. Т. 21, № 5. С. 531–539. DOI: 10.30766/2072-9081.2020.21.5.531-539
3. Клыков А. Г., Коновалова И. В., Богдан П. М., Шадрин Д. М., Цзюймэй Ч., Чжан Х., Ма Ш., Чжан Ж. Анализ сортов яровой мягкой пшеницы (*Triticum aestivum* L.) с использованием микросателлитных маркеров // Российская сельскохозяйственная наука. 2017. № 5. С. 3–6.
4. Колобова О. С., Малюченко О. П., Шалаева Т. В., Шанина Е. П., Шилов И. А., Алексеев Я. И., Велишаева Н. С. Генетическая паспортизация картофеля на основе мультиплексного анализа 10 микросателлитных маркеров // Вавиловский журнал генетики и селекции. 2017. Т. 21, № 1. С. 124–127. DOI: 10.18699/VJ17.230
5. Митрофанова О. П., Стрельченко П. П., Конарев А. В., Балфорьер Ф. Генетическая дифференциация гексаплоидной пшеницы по данным анализа микросателлитных локусов // Генетика. 2009. Т. 45, № 11. С. 1530–1539.
6. Митрофанова, О. П. Генетические ресурсы пшеницы в России: состояние и предселекционное изучение // Вавиловский журнал генетики и селекции. 2012. Т. 16, № 1. С. 10–20.
7. Пискарев В. В., Бойко Н. И. Полиморфизм глиадинкодирующих локусов сортообразцов пшеницы мягкой яровой Сибирского генофонда // Сибирский вестник сельскохозяйственной науки. 2015. № 6. С. 19–24.
8. Портников И. В., Антонова О. Ю., Митрофанова О. П. Молекулярные маркеры в генетическом анализе скрещиваемости мягкой пшеницы с рожью // Вавиловский журнал генетики и селекции. 2020. Т. 24, № 6. С. 557. DOI: 10.18699/VJ20.649
9. Фомина Е. А., Картель Н. А., Гриб С. И., Кулинкович С. Н., Малышев С. В. Использование микросателлитных маркеров для анализа коллекции сортов пшеницы (*Triticum aestivum*) // Известия Национальной академии наук Беларуси. Серия биологических наук. 2014. № 3. С. 31–37.
10. Haque M. S., Saha N. R., Islam M. T., Islam M. M., Kwon S.-J., Roy S. K., Woo S.-H. Screening for drought tolerance in wheat genotypes by morphological and SSR markers // Journal of Crop Science and Biotechnology. 2020. Vol. 24, P. 27–39. DOI: 10.1007/s12892-020-00036-7
11. Hutson, G. FeatureTerminator: A package to perform feature removal for statistical and machine learning models [Electronic resource]. URL: <https://github.com/StatsGary/FeatureTerminator> (accessed: 15.10.2023).
12. Kuhn M., Wing J., Weston S., Williams A., Keefer C., Engelhardt A. Package 'caret' // The R Journal. 2020. Vol. 223, № 7. Article number: 48.

13. Rausch T., Fritz M.H.Y., Untergasser A., Bennes V. Tracy: basecalling, alignment, assembly and deconvolution of sanger chromatogram trace files // BMC genomics. 2020. Vol. 21, P. 1–9. DOI: 10.1186/s12864-020-6635-8
14. Wang L. X., Li H. B., Gu T. C., Liu L. H., Pang B. S., Qui J., Zhao C. P. Assessment of wheat variety stability using SSR markers // Euphytica. 2014. Vol. 195, P. 435–452. DOI: 10.1007/s10681-013-1006-z
15. Wickham H., Averick M., Bryan J., Chang W., McGowan L. D. A., François R., Yutani H. Welcome to the Tidyverse // Journal of open source software. 2019. Vol. 4, № 43. Article number: 1686. DOI: 10.21105/joss.01686

References

1. Adylova A. T., Norbekov Zh. K., Khurshut E. E., Nikitina E. V., Kushanov F. N. SSR-analiz genomnoi DNK perspektivnykh sortov myagkoi ozimoi pshenitsy uzbekistanskoi selektsii [SSR-analysis of genomic DNA of promising winter common wheat varieties of Uzbekistan breeding] // Vavilovskii zhurnal genetiki i selektsii. 2018. T. 22, № 6. S. 634–639. DOI: 10.18699/VJ18.404
2. Guchetl' S. Z., CHelyustnikova T. A. Genotipirovanie sortov maslichnogo l'na s ispol'zovaniem sistemy mikrosatelitnykh DNK-markeroV [Genotyping of oilseed flax varieties using a system of microsatellite DNA markers] // Agrarnaya nauka Evro-Severo-Vostoka. 2020. T. 21, № 5. S. 531–539. DOI: 10.30766/2072-9081.2020.21.5.531-539
3. Kolobova O. S., Malyuchenko O. P., SHalaeva T. V., SHanina E. P., SHilov I. A., Alekseev YA. I., Velishaeva N. S. Geneticheskaya pasportizaciya kartofelya na osnove multipleksnogo analiza 10 mikrosatelitnykh markerov [Genetic certification of potatoes based on multiplex analysis of 10 microsatellite markers] // Vavilovskii zhurnal genetiki i selektsii. 2017. T. 21, № 1. S. 124–127. DOI: 10.18699/VJ17.230
4. Mitrofanova O.P., Strel'chenko P.P., Konarev A.V., Balfor'er F. Geneticheskaya differentsiatsiya geksaploidnoi pshenitsy po dannym analiza mikrosatelitnykh lokusov [Genetic differentiation of hexaploid wheat according to the analysis of microsatellite loci] // Genetika. 2009. T. 45, № 11. S. 1530–1539.
5. Mitrofanova, O. P. Geneticheskie resursy pshenitsy v Rossii: sostoyanie i predselektsionnoe izucheniye [Genetic resources of wheat in Russia: state and pre-breeding study] // Vavilovskii zhurnal genetiki i selektsii. 2012. T. 16, № 1. S. 10–20.
6. Piskarev V. V., Boiko N. I. Polimorfizm gliadinkodiruyushchikh lokusov sortoobraztsov pshenitsy myagkoi yarovoi Sibirskogo genofonda [Polymorphism of gliadin-coding loci of spring common wheat varieties from the Siberian gene pool] // Sibirskii vestnik sel'skokhozyaistvennoi nauki. 2015. № 6. S. 19–24.
7. Portnikov I.V., Antonova O.YU., Mitrofanova O.P. Molekulyarnye markery v geneticheskom analize skreshchivaemosti myagkoj pshenicy s rozh'yu [Molecular markers in the genetic analysis of the crossbreeding of soft wheat with rye] // Vavilovskii zhurnal genetiki i selektsii. 2020. T. 24, № 6. S. 557. DOI: 10.18699/VJ20.649
8. Fomina E.A., Kartel' N.A., Grib S. I., Kulinkovich S. N., Malyshev S. V. Ispol'zovanie mikrosatelitnykh markerov dlya analiza kollektzii sortov pshenitsy (Triticum aestivum) [Use of microsatellite markers to analyze a collection of wheat varieties (Triticum aestivum)] // Izvestiya Natsional'noi akademii nauk Belarusi. Seriya biologicheskikh nauk. 2014. № 3. S. 31–37.
9. Haque M. S., Saha N. R., Islam M. T., Islam M. M., Kwon S.-J., Roy S. K., Woo S.-H. Screening for drought tolerance in wheat genotypes by morphological and SSR markers // Journal of Crop Science and Biotechnology. 2020. Vol. 24, P. 27–39. DOI: 10.1007/s12892-020-00036-7
10. Hutson, G. FeatureTerminator: A package to perform feature removal for statistical and machine learning models [Electronic resource]. URL: <https://github.com/StatsGary/FeatureTerminator> (accessed: 15.10.2023).
11. Kuhn M., Wing J., Weston S., Williams A., Keefer C., Engelhardt A. Package 'caret' // The R Journal. 2020. Vol. 223, № 7. Article number: 48.
12. Rausch T., Fritz M. H. Y., Untergasser A., Bennes V. Tracy: basecalling, alignment, assembly and deconvolution of sanger chromatogram trace files // BMC genomics. 2020. Vol. 21, P. 1–9. DOI: 10.1186/s12864-020-6635-8
13. Wang L. X., Li H. B., Gu T. C., Liu L. H., Pang B. S., Qui J., Zhao C. P. Assessment of wheat variety stability using SSR markers // Euphytica. 2014. Vol. 195, P. 435–452. DOI: 10.1007/s10681-013-1006-z
14. Wickham H., Averick M., Bryan J., Chang W., McGowan L. D. A., François R., Yutani H. Welcome to the Tidyverse // Journal of open source software. 2019. Vol. 4, № 43. Article number: 1686. DOI: 10.21105/joss.01686

Поступила: 10.06.24; доработана после рецензирования: 167.07.24; принята к публикации: 16.07.24.

Критерии авторства. Авторы статьи подтверждают, что имеют на статью равные права и несут равную ответственность за плагиат.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Авторский вклад. Ульянов Д. С. – анализ данных, их интерпретация и сбор данных; Самарина М. А. – анализ данных, их интерпретация и сбор данных; Мохов Т. Д. – сбор данных и подготовка рукописи; Меглицкая Я. С. – сбор данных и подготовка рукописи; Крупин П. Ю. – интерпретация данных и подготовка рукописи; Карлов Г. И. – концептуализация исследования; Харченко П.Н. – концептуализация исследования; Воронов С. И. – подготовка опыта и концептуализация исследования; Давыдова Н. Ю. – подготовка опыта и концептуализация исследования; Дивашук М. Г. – подготовка опыта и концептуализация исследования.

Все авторы прочитали и одобрили окончательный вариант рукописи.