

УДК 579.262:57.083.18

ШТАММ-СПЕЦИФИЧЕСКИЕ МАРКЕРЫ РИЗОБИЙ ПО ДАННЫМ ПОЛНОГЕНОМНОГО СЕКВЕНИРОВАНИЯ¹

© 2025 г. Е. Е. Андронов^а, *, Т. С. Аксенова^а, О. П. Онищук^а, О. Н. Курчак^а,
В. И. Сафронова^а, А. Г. Пинаев^а, И. В. Евсюков^а, Н. А. Проворов^а

^аВсероссийский научно-исследовательский институт сельскохозяйственной микробиологии,
Санкт-Петербург, 196608, Россия
*e-mail: eeandr@gmail.com

Поступила в редакцию 06.08.2024 г.

После доработки 09.09.2024 г.

Принята к публикации 10.09.2024 г.

Полевые испытания ризобияльных инокулянтов требуют использования простых и надежных методов идентификации используемых штаммов для выяснения того, какой именно штамм образовал азотфиксирующий клубенек. Эта задача возникает при испытании конкурентоспособности штаммов-инокулянтов по отношению к местным штаммам ризобий, для отслеживания судьбы штаммов-инокулянтов в длительные периоды после внесения штаммов, наконец, такие методы могут быть востребованы при защите прав собственников и разработчиков штаммов. Суть предлагаемого метода идентификации заключается в поиске штамм-специфичных участков ДНК, отсутствующих в других геномах этого же вида, и конструировании системы праймеров для мультиплексной ПЦР, позволяющей простую, надежную и быструю идентификацию штамма. Преимущества данного подхода по отношению к другим способам идентификации состоят, во-первых, в высокой воспроизводимости, а во-вторых, в том, что способ основан на детекции структурных вариантов, вклад которых в эволюцию геномов ризобий весьма высок, в то время как большинство методов геномного фингерпринтинга (AFLP, RAPD, REP, ERIC и др.) основаны на детекции нуклеотидных полиморфизмов в коротких фрагментах генома, но упускают из поля зрения множество событий, связанных с геномными перестройками и горизонтальным переносом генов. Использование предложенного метода также может послужить для мониторинга эволюционной динамики ризобияльных штаммов-инокулянтов, в особенности в уникальных фрагментах генома, что весьма важно для *Rhizobium leguminosarum*, где доля уникальных последовательностей намного выше, чем у других ризобий.

Ключевые слова: ризобии, геном, структурные варианты генома, полногеномное секвенирование, молекулярные маркеры, идентификация производственных штаммов

DOI: 10.31857/S0026365625010049

Проблема идентификации штаммов микроорганизмов — одна из наиболее старых и лежащих на поверхности проблем, к выполнению которой так или иначе привлекались почти все методы характеристики штаммов. Особенно данная проблема важна для большого числа практических приложений, где весьма важно выполнить элементарную задачу идентификации — установление идентичности двух штаммов. Так, например, при использовании ризобияльных препаратов для бобовых культур, весьма большой проблемой являются местные штаммы ризобий, зачастую обладающие более высокой конкурентоспособностью, но меньшей симбиотической

эффективностью, чем производственные штаммы-инокулянты (Онищук и соавт., 2017). Поэтому при создании таких препаратов весьма важной частью исследований является разработка методов идентификации штаммов ризобий, образовавших азотфиксирующие клубеньки в условиях полевого эксперимента. Другой проблемой является мониторинг судьбы ризобияльных штаммов-инокулянтов после однократного внесения под бобовые культуры в течение нескольких последовательных лет вегетации, в особенности когда на повестке дня находятся вопросы обеспечения генетической безопасности (Андронов и соавт., 2009). Наконец, немаловажной

¹Дополнительные материалы размещены в электронном виде по DOI статьи: <https://doi.org/10.31857/S0026365625010049>

проблемой, особенно в последнее время, является защита прав разработчиков и производителей на коммерческие штаммы (Сафронова и соавт., 2012). Разрешение всех этих проблем так или иначе связано с наличием надежных методов быстрой и недорогой идентификации штамма.

Для решения этой проблемы в разное время были предложены различные методы геномного фингерпринтинга, такие как RAPD (Harrison et al., 1992), REP- (ERIC-, BOX- и др.) ПЦР (De Bruijn, 1992), AFLP (Vos et al., 1995) или рестрикционный анализ амплификатов участков генома с высоким уровнем нуклеотидного полиморфизма, таких как, например, межгенный участок ITS (IGS) рибосомального оперона (Laguerre et al., 1994). Важная проблема, связанная с использованием таких методов, как RAPD и REP-ПЦР, заключается в отсутствии их надежной воспроизводимости, так как на особенности геномного фингерпринта могут влиять тип термостабильной полимеразы, концентрация матричной ДНК и способ ее выделения, особенности термоцикла и другие факторы (Vos et al., 1995). Отчасти эти проблемы разрешаются в методе AFLP, обладающем гораздо более надежной воспроизводимостью и стабильностью фингерпринтов, но и требующем при этом проведения существенно более сложных процедур, как при генерации фингерпринтов, так и при их интерпретации. Однако недостатки указанных методов не ограничиваются указанными проблемами, так как существуют проблемы более глубокого плана, связанные со спецификой геномных фингерпринтов, состоящей в том, что в основе всех этих методов лежат особенности нуклеотидных полиморфизмов в районах отжига RAPD или REP-праймеров, либо в полиморфизмах, приходящихся на сайты рестрикции и их окрестности (AFLP), в то время как структурные варианты ризобияльных геномов, являющиеся следствием геномных перестроек, инсерций, делеций, горизонтального переноса генов могут не слишком сильно сказываться на небольших наборах коротких фрагментов амплификации, генерируемых указанными выше методами геномного фингерпринтинга.

Сегодня, в эпоху геномного секвенирования, становится понятно, что генерация структурных вариантов, в особенности у ризобий, геномы которых насыщены повторами (Freiberg et al., 1997; Guo et al., 2003; Young et al., 2006), а горизонтальный перенос генов является одним из главных механизмов их эволюции (Young et al., 2006; Epsten, Tiffin, 2021; Provorov et al., 2022), является весьма важным и очень частым типом изменчивости, вносящим существенный вклад в геномное разнообразие микроорганизмов. Поэтому структурные вариации геномов не должны упускаться при разработке методов генетической идентификации штаммов. Реализации этого направления весьма способствуют уже ставшие

рутинными методы полногеномного секвенирования, в особенности способные производить длинные прочтения, такие как Oxford Nanopore и PacBio, в геномных сборках которых структурные варианты могут быть идентифицированы с гораздо большей полнотой, чем, например, в технологиях Illumina, по причине относительно небольших длин прочтений последней.

В настоящем исследовании предлагается несложный способ конструирования набора олигонуклеотидов для мультиплексной ПЦР на основе полных геномов ризобияльных штаммов, доступный исследователям с самыми общими навыками в биоинформатике и геномике, позволяющий создавать простой и эффективный инструмент для идентификации штамма в условиях простой реакции ПЦР. Основой предлагаемого подхода является поиск нуклеотидных последовательностей, уникальных для конкретных геномов ризобий.

МАТЕРИАЛЫ И МЕТОДЫ ИССЛЕДОВАНИЯ

Культуры микроорганизмов. В ходе настоящего исследования геномные маркеры были разработаны для двух штаммов ризобий из Сетевой биоресурсной коллекции в области генетических технологий для сельского хозяйства (ВКСМ) ФГБНУ ВНИИСХМ — штамм RCAM0626 *Rhizobium leguminosarum* bv. *viciae*, выделенный из клубеньков вики посевной на Павловской опытной станции ВИРА (Пушкинский район), и штамм RCAM1365 *Rhizobium leguminosarum* bv. *trifolii*, выделенный из клубеньков клевера белого в Бельничском районе Могилевской области, полные геномы которых были нами просеквенированы и депонированы в GenBank ранее: RCAM0626 (CP050553—CP050557), RCAM1365 (CP050514—CP050519) (Safronova et al., 2023).

В качестве группы контрольных штаммов вида *R. leguminosarum* было выбрано 11 штаммов коллекции лаборатории микробиологического мониторинга и биоремедиации почв ФГБНУ ВНИИСХМ различного происхождения. В табл. 1 представлены данные об использованных штаммах.

Все штаммы выращивали в жидкой среде ТУ при 28°C на термостатируемой качалке. Для выделения ДНК из бактерий был использован модифицированный СТАВ метод (Rahimah et al., 2006). Разрушение микроорганизмов осуществляли активным встряхиванием биомассы в присутствии гранатового песка.

Конструирование геномных маркеров. Для обеспечения стабильности маркеров выбран наиболее консервативный репликон — хромосома (RCAM0626 — CP050555.1; RCAM1365 — CP050514.1). Разработка маркеров выполнена в три этапа:

Таблица 1. Штаммы *R. leguminosarum*, использованные в работе

Штамм <i>R. leguminosarum</i>	Биовар	Растение-хозяин	Объект выделения	Регион выделения	Год выделения	Примечание
RCAM1365	<i>trifolii</i>	<i>Trifolium pratense</i>	Клубенек	Могилевская обл.	1981	Коллекция ВНИИСХМ
RCAM0626	<i>viciae</i>	<i>Vicia sativa</i>	Почва	СПб, Павловск	1986	Коллекция ВНИИСХМ
11Ta	<i>trifolii</i>	<i>T. ambiguum</i>	Почва	Северный Кавказ	2019	Коллекция лаб. 2 ВНИИСХМ
Ta3k	<i>trifolii</i>	<i>T. ambiguum</i>	Почва	Северный Кавказ	2019	Коллекция лаб. 2 ВНИИСХМ
Ta6	<i>trifolii</i>	<i>T. ambiguum</i>	Почва	Северный Кавказ	2019	Коллекция лаб. 2 ВНИИСХМ
Tm2	<i>trifolii</i>	<i>T. montanum</i>	Почва	Северный Кавказ	2019	Коллекция лаб. 2 ВНИИСХМ
TP65	<i>trifolii</i>	<i>T. pratense</i>	Клубенек	СПб, Пушкин	2022	Коллекция лаб. 2 ВНИИСХМ
Vs35	<i>viciae</i>	<i>V. sativa</i>	Почва	Северный Кавказ	2019	Коллекция лаб. 2 ВНИИСХМ
Vs37-3	<i>viciae</i>	<i>V. sativa</i>	Почва	Северный Кавказ	2019	Коллекция лаб. 2 ВНИИСХМ
Vsa1	<i>viciae</i>	<i>V. subrotunda</i>	Клубенек	Приморский край	2016	Коллекция лаб. 2 ВНИИСХМ
Vf45	<i>viciae</i>	<i>Vavilovia formosa</i>	Клубенек	Армения	2016	Коллекция лаб. 2 ВНИИСХМ
Vf12	<i>viciae</i>	<i>V. formosa</i>	Клубенек	Северная Осетия	2016	Коллекция лаб. 2 ВНИИСХМ
Vf08	<i>viciae</i>	<i>V. formosa</i>	Клубенек	Северная Осетия	2016	Коллекция лаб. 2 ВНИИСХМ

(1) Выравнивание хромосомы штамма с хромосомой любого другого штамма этого же вида в программе Mauve ver. 20150226 (Darling et al., 2004). Для первого этапа разработки маркеров достаточно выравнивания с хромосомой одного штамма. Так как в настоящей работе были использованы два штамма одного вида, мы ограничились выравниванием их хромосом друг с другом. После выравнивания хромосом для каждого штамма были идентифицированы уникальные фрагменты ДНК длиной не менее 1000 п.н., из которых для каждого из штаммов были сконструированы конкатенаты, объединяющие по 15–20 уникальных для штамма фрагментов ДНК суммарной длиной около 300 тыс. п.н. Смысл данного этапа – уменьшение числа последовательностей-кандидатов для ускорения дальнейшего поиска фрагментов генома с минимальным количеством гомологов в GenBank;

(2) С использованием алгоритма BLASTn был проведен поиск в базе GenBank для каждого из конкатенатов, после чего были выбраны по 12 фрагментов конкатената с минимальным числом гомологов в базе данных;

(3) С использованием опции Multiplex PCR в программе Vector NTI 10.1.1 для выбранных 10 фрагментов каждого из геномов, для удобства конструирования разделенными линкерами (N)₁₀, были сконструированы по 10 пар праймеров для мультиплексной ПЦР, фланкирующие фрагменты в диапазоне от 100 до 1100 п.н.

Мультиплексная ПЦР. Каждый из разработанных 20 пар праймеров (по 10 для каждого из штаммов) был проверен в реакции ПЦР с использованием набора с полимеразой Encyclo (“Евроген”, Россия). Температурный профиль был одинаковым для всех праймеров в условиях обычной и мультиплексной ПЦР (95°C – 30 с, 55°C – 30 с, 72°C – 1 мин; всего 35 циклов). Для постановки мультиплексной ПЦР использовали наборы всех праймеров (по 2.5 пмоль каждого из праймеров). При постановке мультиплексной ПЦР в 2 раза увеличивали количество смеси dNTP. Результаты амплификации анализировали в условиях гель-электрофореза в 2% агарозе.

Работа проводилась с использованием оборудования ЦКП “Геномные технологии, протеомика и клеточная биология” ФГБНУ ВНИИСХМ.

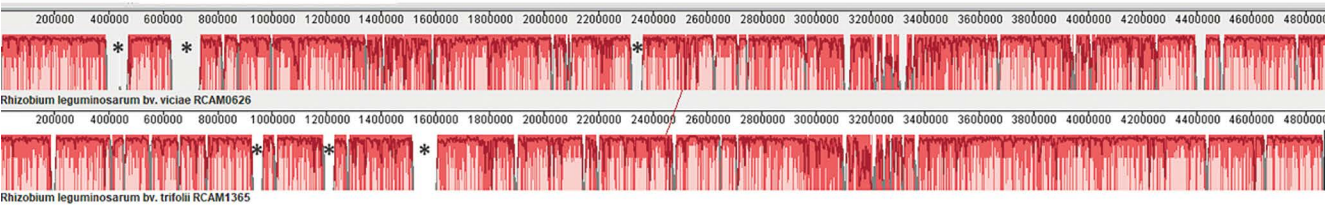


Рис. 1. Выравнивание хромосом геномов штаммов RCAM1365 и RCAM0626. Звездочкой отмечены три самых крупных уникальных фрагмента для каждой из хромосом.

РЕЗУЛЬТАТЫ

На рис. 1 представлены данные выравнивания хромосом штаммов RCAM1365 и RCAM0626 в программе Mauve. Из приведенных данных очевидно, что обе хромосомы имеют достаточно большое количество уникальных фрагментов, присутствующих только в геноме одного штамма. Общее количество таких фрагментов составляло 15–20, а их размер – до 100 тыс. п.н.

После проведения первого предварительного этапа были получены два конкатената, уникальных для каждого из геномов последовательностей, содержащих по 20–30 фрагментов. Суммарный размер каждого из конкатенатов – 348 т.п.н. для штамма RCAM0626 и 288 т.п.н. для штамма RCAM1365. Второй этап конструирования геномных маркеров состоял в анализе каждого из конкатенатов с использованием алгоритма BLASTn в базе данных

Таблица 2. Праймеры для мультиплексной ПЦР штамма RCAM0626, размеры и локализация фрагментов в геномах, функциональная аннотация амплифицируемых фрагментов

Номер фрагмента	Последовательности праймеров	Старт	Конец	Размер фрагмента	Аннотация
1*	GATTTGGTGGGAATTGGCAGC/ GTTTCCGCAAAGCAGAAGCT	392946	393124	179	HB771_07790, hypotetical protein
2*	TTCTTGTGTCCTCGAAGG/ TGCGACAAGAATGGCATTCT	412000	412200	201	HB771_07890, aspartate aminotransferase family protein
3*	CCTTCTTTTGTGCAAACCAA/ TAATAGGGTTGCTGCGTTGC	637996	638307	312	HB771_08910, hypotetical protein
4	CTTGGAACAACATGGCAAACC/ TATTTTCCCCTCAACAAGC	659962	660398	437	HB771_09030, adenylate/guanylate cyclase domain-containing protein
5*	GAGTTCGAACAGCATCACCC/ CGCCGGACACTTCTGATTGT	681998	682498	501	HB771_09120, AAA family ATPase
6*	AGTCTTTTTTCCGACGCTGG/ GAACATGTGATCGAGACGGC	698960	699619	660	HB771_09185, damage-inducible mutagenesis protein
7*	GTAAACCGTCCCGTCAGCAG/ TCTGTCACCAATGCCAGTT	2341932	2342693	762	HB771_09190, DNA polymerase Y family protein
8	CCATCTTCTGACAAACCGCC/ GTCAGTGCCGCTCTCGAAAC	817996	818848	853	HB771_17330, hypotetical protein
9	GCCTTTAAAGTTGAT- GTTTTCTTGG/ AAATGAAAACCTCGGCTAC	2328986	2329913	928	HB771_09780, glycosyltransferase
10	AAGTTGTTGTCGGCGACGAC/ CTTCATTCCGACCCAGCAA	2345982	2346994	1013	HB771_17250, hypotetical protein

Примечание. Звездочкой отмечены фрагменты, представленные на электрофореze на рис. 3а. Старт и конец фрагмента амплификации даны по позициям праймеров в хромосоме штамма.

GenBank для поиска фрагментов конкатената, максимально специфичных не только при сравнении двух геномов, но и всех геномов *R. leguminosarum*, депонированных в GenBank (более 400) и, таким образом, оптимальных для конструирования штамм-специфичных маркеров. По результатам анализа BLASTn конкатената штамма RCAM0626 было показано, что в масштабах всей базы данных участков генома, специфичных только для генома штамма RCAM0626, мало, однако создание набора фрагментов, которые встречались бы только в одном геноме, является вполне выполнимой задачей. В результате было выбрано 10 фрагментов с минимальным числом гомологов в GenBank, для которых с использованием опции Multiplex PCR из программы Vector NTI был сконструирован набор олигонуклеотидов для амплификации 10 фрагментов различного размера в диапазоне от 100 до 1100 п.н. Аналогичные процедуры были проведены для генома штамма RCAM1365. Разработанные олигонуклеотиды, размеры и локализация фрагментов в геномах, функциональная аннотация амплифицируемых фрагментов приведены в табл. 2 и 3.

На рис. 2а, 2б приведены результаты амплификации каждого из фрагментов на матрице ДНК того штамма, для генома которого они были

сконструированы. Все сконструированные праймеры позволяли амплифицировать специфичный фрагмент ДНК ожидаемого размера.

Наконец, была проведена амплификация с использованием сконструированных праймеров в условиях мультиплексной ПЦР с использованием в качестве контролей ДНК штаммов RCAM0626 и RCAM1365, а также 11 штаммов *R. leguminosarum* из лабораторной коллекции. Для амплификации использовали 10 пар праймеров. Результаты мультиплексной ПЦР приведены на рис. 3а, 3б.

Как видно из представленных данных, амплификация всех 10 фрагментов удалась не полностью: для обоих штаммов в условиях мультиплексной ПЦР амплифицировалось 6 из 10 фрагментов. Это неудивительно, так как уже при конструировании набора праймеров для мультиплексной ПЦР было показано, что весь набор фрагментов амплифицировать одновременно невозможно. Однако, как видно из представленных данных, для данной группы штаммов 6 фрагментов вполне достаточно для однозначной идентификации штаммов RCAM0626 и RCAM1365 в пределах группы штаммов, выбранной для контроля. Остальные разработанные праймеры могут быть использованы в индивидуальном порядке в реакции ПЦР, если в этом будет необходимость.

Таблица 3. Праймеры для мультиплексной ПЦР штамма RCAM1365, размеры и локализация фрагментов в геномах, функциональная аннотация амплифицируемых фрагментов

Номер фрагмента	Последовательности праймеров	Старт	Конец	Размер фрагмента	Аннотация
1*	TTCTATGCAATTGCCGGATG/ TTCCTTCGAAAAGACCTGCG	657943	658117	175	HB775_03050, ATP-binding protein
2*	AGCTTACGGAAGTCAATGCC/ GATCTGGACCGGTCTGAACG	552955	553186	232	HB775_02525, response regulator transcription factor
3	AAGACGACCGAAAGGCTCTG/ ACGATACAGCTGGTAGGCAG	928971	929305	335	HB775_04355, CHAT domain-containing protein
4*	CATTCTTCAGGCGATCGAGC/ AGGTATCGCCTTTCGGTGTG	4435973	4436416	444	HB775_21920, hypothetical protein
5*	CTGTTCTCCTGGCCGTATCC/ GACGGCAAGGTCTTGGACTA	3246977	3247488	512	HB775_16080, DUF262 domain-containing protein
6*	TAGATCACATCGGCTGCATC/ GAAACAGAAACGAGCCCACG	1517258	1517907	650	HB775_07375, methyltransferase domain-containing protein
7	GTCAGAAAGTTGCCTCGGTG/ GTAGCCTTCGTTTCGCATTGC	2195947	2196751	805	HB775_10625, nitronate monooxygenase
8*	AGACGTCTCGACGGTGTGTTG/ ATCAAGCAGATCGTTGTCGG	1197972	1198827	856	HB775_05680, phage portal protein
9	GGTTCACGGTCCCTCCTAT/ CGGTAATAGAGGTACAGAATGC	1208998	1209902	905	HB775_05750, hypothetical protein
10	GTTTCTGTCATCGCTGAAGT/ GTTGACTATCCCCTGCTTC	1009399	1010431	1033	HB775_04795, carbohydrate ABC transporter permease

Примечание. Звездочкой отмечены фрагменты, представленные на электрофореze на рис. 3б. Старт и конец фрагмента амплификации даны по позициям праймеров в хромосоме штамма.

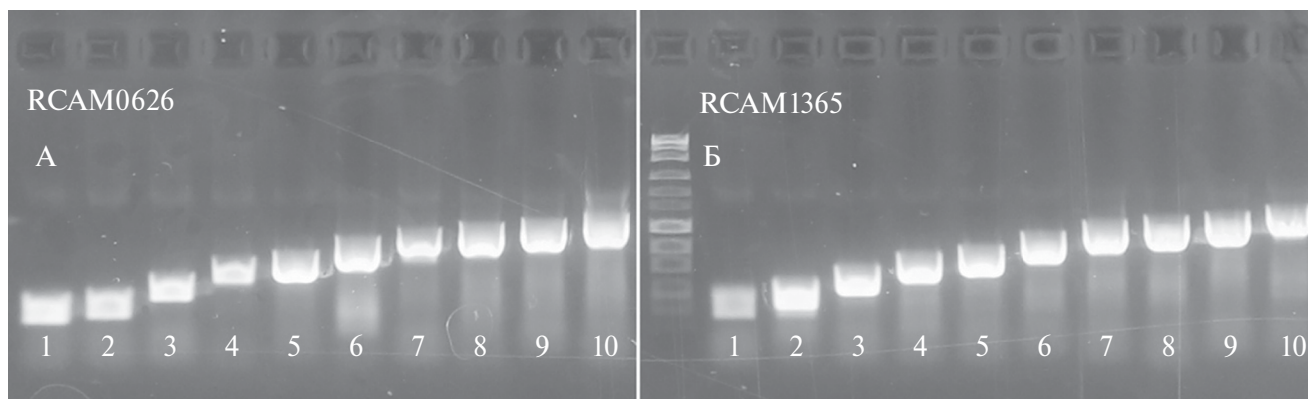


Рис. 2. Амплификация маркерных фрагментов для штамма RCAM0626 (а) и RCAM1365 (б). Номера соответствуют праймерам в табл. 2 и 3.

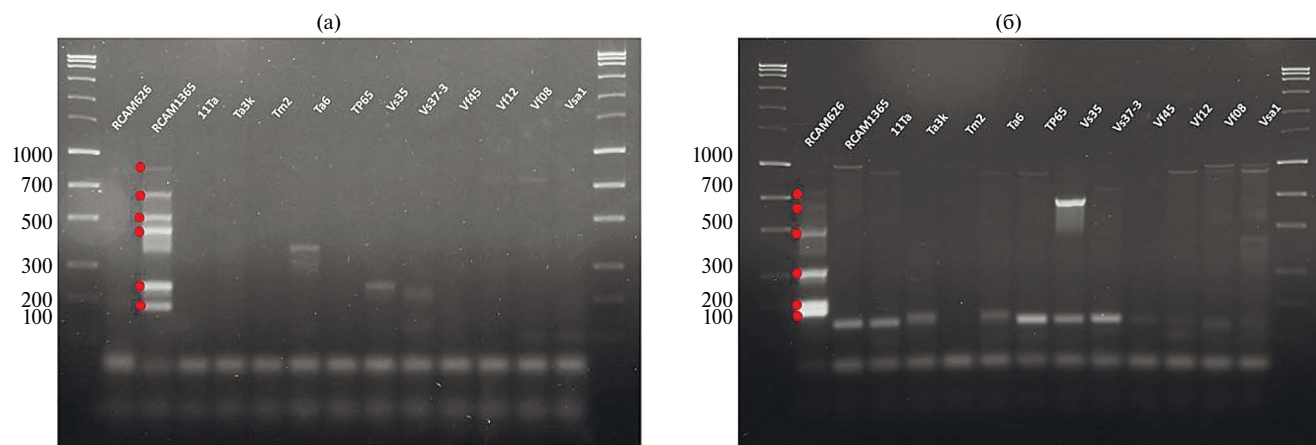


Рис. 3. Мультиплексная ПЦР с 10-ю парами праймеров для маркирования штаммов: а – штамм RCAM1365 (фрагменты 1, 2, 3, 5, 6, 7 в соответствии с табл. 2 отмечены красными точками); б – штамм RCAM0626 (фрагменты 1, 2, 4, 5, 6, 8 в соответствии с табл. 3 отмечены красными точками). В качестве контролей использованы штаммы RCAM0626, RCAM1365 и 11 штаммов *R. leguminosarum* различного происхождения.

Для дополнительного контроля специфичности системы разработанных праймеров, для каждого из двух штаммов были сконструированы конкатенаты всех 10 фрагментов, праймеры для амплификации которых представлены в табл. 2 и 3. Данные конкатенаты были использованы для поиска гомологов в GenBank, с использованием алгоритма BLASTn. Было показано, что из всей базы данных, содержащей около 400 полных геномов *R. leguminosarum*, максимальное сходство с набором фрагментов штамма RCAM0626 выявлено у двух геномов, содержащих 6 фрагментов из 10, для остальных геномов, депонированных в GenBank, количество гомологов 5 (2 генома) или меньше. При этом 6 фрагментов, амплифицирующихся в условиях мультиплексной ПЦР (рис. 3а, 3б), будут достаточны для выявления минимум двух различий с любым штаммом, геном которого представлен в GenBank.

Важно отметить, что топ-лист результатов BLASTn с максимальным сходством по наличию маркерных последовательностей генома RCAM0626 возглавляют четыре генома штаммов-симбионтов клевера *Rhizobium ruizarguesonis*: TP15 (CP140868.1), TP33 (CP140840.1), TP42 (CP140830.1), TP53 (CP140814.1), которые были выделены на территории ФГБНУ ВНИИСХМ в 2022 году, просеквенированы и депонированы в GenBank. Графическое представление результатов BLASTn приведено в дополнительных материалах (рис. S1a). Некоторые предположения о причинах такого неожиданного сходства геномов штаммов, выделенных с разницей в 40 лет в одном районе, будут приведены в обсуждении. Для набора маркерных последовательностей штамма RCAM1365 в GenBank выявлены только 2 генома, содержащих 3 маркерных последовательности из 10, у остальных геномов

таких последовательностей выявлено 2 или меньше. Графическое представление результатов BLASTn приведено в дополнительных материалах (рис. S16).

Таким образом, показано, что разработанный метод позволит уверенно отличить штаммы RCAM0626 и RCAM1365 как от любого из 11 штаммов *R. leguminosarum*, использованных в качестве контрольной группы при постановке мультиплексной ПЦР, так и от любого из штаммов этого и родственных видов, геномы которых депонированы в GenBank.

ОБСУЖДЕНИЕ

Необходимость наличия геномных маркеров, позволяющих однозначно идентифицировать штаммы ризобий в условиях полевого эксперимента, — это давняя проблема практических приложений генетики этих бактерий. На протяжении 20–30 лет в этой области господствовали методы ПЦР-фингерпринтинга, основанные на амплификации либо случайных фрагментов генома (RAPD), либо диспергированных повторов (ERIC, REP, BOX), либо селективной амплификацией небольшой фракции рестрикционных фрагментов генома (AFLP), либо рестрикционным анализом отдельных участков генома с высоким уровнем нуклеотидного полиморфизма (ITS). Недостатки этих методов обсуждались кратко во введении. Ставшее к настоящему времени рутинным секвенирование полных геномов микроорганизмов позволяет по-новому подойти к проблеме маркирования ризобий штаммов. И одним из первых феноменов, ставшим очевидным в эпоху геномного анализа, является сложная структура видовой пангенома, в котором можно выделить кор и уникальную часть, размер которой у ризобий может достигать 50% генома (Rosselli et al., 2021).

Типичными представителями акцессорного компонента являются симбиотические гены ризобий, локализованные на отдельной плазмиде, однако по причине того, что в настоящей работе для разработки маркеров была выбрана хромосома, как наиболее стабильный репликон ризобийного генома, здесь уместно говорить об акцессорном компоненте ризобийной хромосомы, который также может иметь весьма внушительные размеры. Так, уже при сопоставлении хромосом штаммов RCAM0626 и RCAM1365, относящихся к одному и тому же виду *R. leguminosarum*, для конструирования геномных маркеров нами были собраны конкатенаты, состоящие из уникальных фрагментов, присутствующих в одном геноме и отсутствующих в другом. Их размеры составили 348 и 288 т.п.н. для каждого из геномов, что составляет более 5% каждой из хромосом, причем, как следует из данных, приведенных в табл. 2 и 3, функции многих генов,

представленных в них, могут оказаться существенными для штаммов. На основе этого хромосомного акцессорного компонента ризобийной хромосомы в рамках настоящего исследования были разработаны геномные маркеры штаммов. И, как нам представляется, весьма важно, что данные маркеры представляют именно акцессорный компонент ризобийного генома, по всей видимости, играющий существенную роль в определении фенотипической индивидуальности штаммов.

Так как целью настоящей работы было представление метода разработки геномных маркеров для идентификации ризобий для использования в полевых экспериментах, в обсуждении мы ограничимся кратким упоминанием проблем, связанных с детальным анализом функций акцессорного генома ризобий, оставив эту проблему для других публикаций. Однако мы не можем не коснуться одного из неожиданных феноменов, выявленных в данном исследовании: в результате поиска гомологов конкатената 10 фрагментов, выбранных в качестве маркеров штамма RCAM0626 (симбионт вики), в GenBank, топ-лист гомологов возглавили 4 генома ризобий клевера, относящиеся к виду со сложно произносимым названием *R. ruizarguesonis*. Проблема видовой принадлежности этих штаммов объясняется недавней расклассификацией нескольких геновидов *R. leguminosarum* (Young et al., 2021) в отдельные виды, в том числе и в вышеупомянутый вид (Jorin et al., 2020). Проблема того, что сам штамм RCAM0626 относится к биовару *viceae*, а 4 упомянутых родственных генома к биовару *trifolii*, объясняется свободным переносом симбиотических плазмид между хромосомами различных биоваров *R. leguminosarum* (Laguette et al., 1992). Наконец, третья проблема требует особого внимания. Дело в том, что все 4 штамма *R. ruizarguesonis*, о которых речь шла выше, выделены в 2022 году на территории ФГБНУ ВНИИСХМ Пушкинского района Санкт-Петербурга, географически близкого к Павловскому району, где в 1986 году был выделен штамм RCAM0626. И это при том, что количество геномов *R. leguminosarum*, депонированных в GenBank, составляет около 400. Очевидно, что сходство акцессорных компонентов у геномов, происходящих из одного региона, хотя и выделенных с разницей в 40 лет, вряд ли является случайностью. Полученные результаты могут быть связаны либо с географической специфичностью акцессорного компонента ризобий, либо с горизонтальным переносом некоторых хромосомных генов штамма RCAM0626 в местные ризобии. Последнее весьма вероятно, так как этот штамм активно использовался в ходе полевых и микроделяночных экспериментах на территории ФГБНУ ВНИИСХМ многие годы. Мы предполагаем провести более детальный анализ этого феномена в недалеком будущем.

Таким образом, в настоящей работе представлен весьма простой способ разработки геномных маркеров, позволяющих надежную и быструю идентификацию ризобийных штаммов в условиях полевого эксперимента для исследований конкурентоспособности штаммов, их динамики в условиях многолетних полевых экспериментов, а также имеющий некоторые перспективы для защиты прав правообладателей. Разработка данного метода поднимает целый ряд вопросов эволюционной динамики ризобийного генома, для анализа которых требуются дополнительные исследования.

ФИНАНСИРОВАНИЕ РАБОТЫ

Работа выполнена при поддержке Минобрнауки России в рамках Соглашения № 075-15-2021-1055 от 28 сентября 2021 г. о предоставлении гранта в форме субсидии из федерального бюджета на реализацию проекта “Мобилизация генетических ресурсов микроорганизмов на базе Ведомственной коллекции полезных микроорганизмов сельскохозяйственного назначения (ВКСМ) при ФГБНУ ВНИИСХМ с использованием сетевого принципа организации”.

СОБЛЮДЕНИЕ ЭТИЧЕСКИХ СТАНДАРТОВ

Настоящая статья не содержит результатов исследований с использованием животных в качестве объекта.

КОНФЛИКТ ИНТЕРЕСОВ

Авторы заявляют, что у них нет конфликта интересов.

СПИСОК ЛИТЕРАТУРЫ

- Андронов Е.Е., Петрова С.Н., Чижевская Е.П., Коростик Е.В., Ахтемова Г.А., Пинаев А.Г. Влияние внесения генетически модифицированного штамма *Sinorhizobium meliloti* Ach-5 на структуру почвенного сообщества микроорганизмов // Микробиология. 2009. Т. 78. С. 525–534.
- Andronov E.E., Chizhevskaya E.P., Korostik E.V., Akhtemova G.A., Pinaev A.G., Petrova S.N. Influence of introducing the genetically modified strain *Sinorhizobium meliloti* ACH-5 on the structure of the soil microbial community // Microbiology (Moscow). 2009. V. 78. P. 474–482.
- Онищук О.П., Воробьев Н.И., Проворов Н.А. Нодуляционная конкурентоспособность клубеньковых бактерий: генетический контроль и адаптивное значение // Прикл. Биохимия и микробиология. 2017. Т. 53. С. 127–135.
- Onishchuk O.P., Vorobyov N.I., Provorov N.A. Nodulation competitiveness of nodule bacteria: Genetic control and adaptive significance // Appl. Biochem. Microbiol. 2017. V. 53. P. 131–139.
- Сафронова В.И., Чижевская Е.П., Андронов Е.Е. Разработка методики молекулярно-генетической паспортизации штаммов сельскохозяйственных микроорганизмов с помощью AFLP-фингерпринтинга // Сельскохозяйственная биология. 2012. Т. 47. № 6. С. 116–121.
- Darling A.C., Mau B., Blattner F.R., Perna N.T. Mauve: multiple alignment of conserved genomic sequence with rearrangements // Genome Res. 2004. V. 14. P. 1394–1403.
- De Bruijn F.J. Use of repetitive (repetitive extragenic palindromic and enterobacterial repetitive intergeneric consensus) sequences and polymerase chain reaction to fingerprint the genomes of *Rhizobium meliloti* isolates and other soil bacteria // Appl. Environ. Microb. 1992. V. 58. P. 2180–2187.
- Epstein B., Tiffin P. Comparative genomics reveals high rates of horizontal transfer and strong purifying selection on rhizobial symbiosis genes // Proc. Biol. Sci. 2021. V. 288. Art. 20201804.
- Freiberg C., Fellay R., Bairoch A., Broughton W.J., Rosenthal A., Perret X. Molecular basis of symbiosis between *Rhizobium* and legumes // Nature. 1997. V. 387. P. 394–401.
- Guo X., Flores M., Mavingui P., Fuentes S.I., Hernandez G., Davila G., Palacios R. Natural genomic design in *Sinorhizobium meliloti*: novel genomic architecture // Genome Res. 2003. V. 13. P. 1810–1817.
- Harrison S.P., Mytton L.R., Scot L., Dye M., Cresswell A. Characterization of *Rhizobium* isolates by amplification of DNA polymorphisms using random primers // Can. J. Microbiol. 1992. V. 38. P. 1009–1015.
- Jorriin B., Palacios J.M., Peix Á., Imperial J. *Rhizobium ruizarguesonis* sp. nov., isolated from nodules of *Pisum sativum* L. // Syst. Appl. Microbiol. 2020. V. 43. Art.126090.
- Laguerre G., Geniaux E., Mazurier S.I., Rodríguez Casartelli R., Amarger N. Conformity and diversity among field isolates of *Rhizobium leguminosarum* bv. *viciae*, bv. *trifolii*, and bv. *phaseoli* revealed by DNA hybridization using chromosome and plasmid probes // Can. J. Microbiol. 1992. V. 39. P. 412–419.
- Laguerre G., Allard M.R., Revoy F., Amarger N. Rapid identification of *Rhizobia* by restriction fragment length polymorphism analysis of PCR-amplified 16S rRNA genes // Appl. Environ. Microb. 1994. V. 60. P. 56–63.
- Provorov N.A., Andronov E.E., Kimeklis A.K., Onishchuk O.P., Iolkina A.A., Karasev E.S. Microevolution, speciation and macroevolution in rhizobia: genomic mechanisms and selective patterns // Front. Plant. Sci. 2022. V. 13. Art. 1026943.

- Rahimah A.R., Cheah S.C., Rajinder S. Freeze drying of oil palm (*Elaeis guineensis*) leaf and its effect on the quality of extractable DNA // J. Oil Palm Res. 2006. V. 18. P. 296–304.
- Rosselli R., La Porta N., Muresu R., Stevanato P., Concheri G., Squartini A. Pangenomics of the symbiotic *Rhizobiales*. Core and accessory functions across a group endowed with high levels of genomic plasticity // Microorganisms. 2021. V. 9. Art. 407.
- Safronova V., Sazanova A., Belimov A., Guro P., Kuznetsova I., Karlov D., Chirak E., Yuzikhin O., Verkhozina A., Afonin A., Tikhonovich I. Synergy between rhizobial co-microsymbionts leads to an increase in the efficiency of plant–microbe interactions // Microorganisms. 2023. V. 11. Art. 1206.
- Vos P., Hogers R., Bleeker M., Rejans M., Van de Lee T., Hornes M., Frijters A., Pot J., Peleman J., Kuiper M., Zabeau M. AFLP: a new technique for DNA fingerprinting // Nucleic Acids Res. 1995. V. 23. P. 4407–4414.
- Young J.P., Crossman L.C., Johnston A.W., Thomson N.R., Ghazoui Z.F., Hull K.H., Wexler M., Curson A.R., Todd J.D., Poole P.S., Mauchline T.H., East A.K., Quail M.A., Churcher C., Arrowsmith C., Cherevach I., Chillingworth T., Clarke K., Cronin A., Davis P., Fraser A., Hance Z., Hauser H., Jagels K., Moule S., Mungall K., Norbertczak H., Rabbino-witsch E., Sanders M., Simmonds M., Whitehead S., Parkhill J. The genome of *Rhizobium leguminosarum* has recognizable core and accessory components // Genome Biol. 2006. V. 7. R34. <https://doi.org/10.1186/gb-2006-7-4-r34>
- Young J.P.W., Moeskjær S., Afonin A., Rahi P., Maluk M., James E.K., Cavassim M.I.A., Rashid M.H., Aserse A.A., Perry B.J., Wang E.T., Velázquez E., Andronov E.E., Tampakaki A., Flores Félix J.D., Rivas González R., Youseif S.H., Lepetit M., Boivin S., Jorin B., Kenicer G.J., Peix A., Hynes M.F., Ramírez-Bahena M.H., Gulati A., Tian C.F. Defining the *Rhizobium leguminosarum* species complex // Genes. 2021. V. 12. Art. 111. <https://doi.org/10.3390/genes12010111>

EXPERIMENTAL ARTICLES

Strain-Specific Markers of Rhizobia from Whole-Genome Sequencing Data

E. E. Andronov¹*, T. S. Aksenova¹, O. P. Onishchuk¹, O. N. Kurchak¹,
V. I. Safronova¹, A. G. Pinaev¹, I. V. Evsyukov¹, N. A. Provorov¹

¹All-Russian Research Institute of Agricultural Microbiology, St. Petersburg, 196608, Russia

*e-mail: eeandr@gmail.com

Abstract. Field trials of rhizobial inoculants require simple and reliable methods for identifying the strains used to determine which strain has formed a nitrogen-fixing nodule. This task arises when testing the competitiveness of inoculant strains against local rhizobia strains, to track the fate of inoculant strains over long periods after the introduction of strains, and finally, such methods may be in demand when protecting the rights of strain owners and developers. The essence of the proposed identification method is to search for strain-specific DNA regions that are absent in other genomes of the same species and to construct a primer system for multiplex PCR, allowing simple, reliable and rapid identification of the strain. The advantages of this approach over other identification methods are, firstly, high reproducibility, and secondly, that the method is based on the detection of structural variants, the contribution of which to the evolution of rhizobia genomes is very high, while most genomic fingerprinting methods (AFLP, RAPD, REP, ERIC, etc.) are based on the detection of nucleotide polymorphisms in short fragments of the genome, but miss many events associated with genomic rearrangements and horizontal gene transfer. The use of the proposed method can also serve to monitor the evolutionary dynamics of rhizobial inoculant strains, especially in unique fragments of the genome, which is very important for *R. leguminosarum*, where the proportion of unique sequences is much higher than in other rhizobia.

Keywords: rhizobia, genome, structural variants of the genome, whole-genome sequencing, molecular markers, identification of production strains