

ISSN 0026-8984

Том 57, Номер 3

Май - Июнь 2023



МОЛЕКУЛЯРНАЯ БИОЛОГИЯ

www.sciencejournals.ru



СОДЕРЖАНИЕ

Том 57, номер 3, 2023

ОБЗОРЫ

Улучшение культурных растений при помощи системы CRISPR/Cas:
новые гены-мишени

*Ю. В. Ухатова, М. В. Ерастенкова, Е. С. Коршикова,
Е. А. Крылова, А. С. Михайлова, Т. В. Семилет,
Н. Г. Тихонова, Н. А. Швачко, Е. К. Хлесткина* 387

Строение, экспрессия и неканонические функции рДНК человека:
роль некодирующих регионов

А. А. Садова, Д. Ю. Пантелеев, Г. В. Павлова 411

Как гистондеацетилаза 3 контролирует экспрессию гепсидина и репликацию
вируса гепатита С

А. С. Щербакова, С. Н. Кочетков, М. В. Козлов 427

Методы обогащения ДНК-фрагментов для высокопроизводительного секвенирования

А. Н. Синяков, Е. В. Костина 440

ГЕНОМИКА. ТРАНСКРИПТОМИКА

Транскрипционный анализ гена предполагаемого миристоилированного
мембранного белка (ORF458R) вируса радужной окраски беспозвоночных 6 (PIV6)

С. А. Kuz, E. Ozsahin, R. Nalcacioglu, Z. Demirbag 458

Дифференциальная экспрессия чужеродного гена в митохондриях арабидопсиса
в системе *in organello*

*В. И. Тарасенко, Т. А. Тарасенко, И. В. Горбенко,
Ю. М. Константинов, М. В. Кулинченко* 460

Идентификация функционально значимых полиморфных вариантов
в генах микроРНК при атеросклерозе сонных артерий

*А. А. Зарубин, К. В. Маннанова, Ю. А. Королёва, А. А. Слепцов,
М. С. Кузнецов, Б. Н. Козлов, М. С. Назаренко* 471

Вовлеченность гена *Gagr*, доместичированного гена *gag* ретровирусов,
в пути стрессового ответа у разных видов дрозофилы

А. Н. Гигин, Л. Н. Нефедова 483

Профилирование экспрессии микроРНК в В-крупноклеточной лимфоме

*Ю. А. Веряскина, С. Е. Титов, И. Б. Ковынев, С. С. Фёдорова,
Я. Ю. Шебуняева, О. В. Антоненко, Т. И. Поспелова, И. Ф. Жимулёв* 492

МОЛЕКУЛЯРНАЯ БИОЛОГИЯ КЛЕТКИ

Влияние рецептора-γ, активируемого пролифератором пероксисом,
на модуляцию синтеза ангиопозитин-подобного белка 4
в клетках Caco-2, обработанных *Clostridium butyricum*

X. Zhao, H. S. Huang, S. R. Shi 501

EBF1 повышает чувствительность клеток рака шейки матки к цисплатину
посредством активации транскрипции FBN1

N. N. Shen, J. H. Lin, P. P. Liu 503

СТРУКТУРНО-ФУНКЦИОНАЛЬНЫЙ АНАЛИЗ БИОПОЛИМЕРОВ И ИХ КОМПЛЕКСОВ

Лиганды, специфичные к определенным последовательностям пар оснований ДНК.
XIX. Синтез, спектральные характеристики, вирусологические и биохимические
исследования флуоресцентных симметричных димерных трисбензимидазолов DB₃(*n*)

*А. Ф. Арутюнян, А. А. Костюков, С. П. Королёв, М. Б. Готтих,
О. Ю. Сусова, Д. Н. Калюжный, А. Л. Жузе*

517

Окисление порфирином ZnP1 G4-структур в последовательностях промоторов
генов *MYC* и *TERT*

Г. В. Чащина, Д. Н. Калюжный

528

БИОИНФОРМАТИКА

Поиск прогностических маркеров гепатоцеллюлярной карциномы с помощью
интеллектуального анализа текстов и данных

Z. H. Yang, S. X. Wang

537

МЕТОДЫ

Рекомбиназная полимеразная амплификация для быстрого выявления возбудителей
бактериальной пневмонии человека

С. А. Лапа, С. А. Суржиков, С. А. Благодатских, В. Е. Шершов, А. В. Чудинов

539

Нуклеазы CRISPR-Cas II типа: алгоритм поиска и *in vitro* характеристика

*А. А. Васильева, С. А. Алюкас, П. А. Селькова, А. Н. Арсениев, В. Е. Чернова,
О. С. Мушарова, Е. И. Климук, М. А. Ходорковский, К. В. Северинов*

546

УДК 577.21:631.52

УЛУЧШЕНИЕ КУЛЬТУРНЫХ РАСТЕНИЙ ПРИ ПОМОЩИ СИСТЕМЫ CRISPR/Cas: НОВЫЕ ГЕНЫ-МИШЕНИ

© 2023 г. Ю. В. Ухатова^а, *, М. В. Ерастенкова^а, Е. С. Коршикова^а, Е. А. Крылова^а, А. С. Михайлова^а, Т. В. Семилет^а, Н. Г. Тихонова^а, Н. А. Швачко^а, Е. К. Хлесткина^а

^аФедеральный исследовательский центр Всероссийский институт генетических ресурсов растений им. Н.И. Вавилова, Санкт-Петербург, 190000 Россия

*e-mail: sci_secretary@vir.nw.ru

Поступила в редакцию 04.05.2022 г.

После доработки 23.09.2022 г.

Принята к публикации 07.10.2022 г.

Успехи геномного редактирования сельскохозяйственных культур с использованием системы CRISPR/Cas в большой степени зависят от правильного выбора генов-мишеней, направленные изменения в которых позволят повысить урожайность, улучшить качество растительного сырья и устойчивость к биотическим и абиотическим стрессирующим факторам. В настоящей работе систематизированы и каталогизированы сведения о генах-мишенях, использованных для улучшения культурных растений. В последнем систематическом обзоре рассмотрены статьи, индексируемые в базе данных Scopus, опубликованные на 17.08.2019 г. В нашей работе охвачен период с 18.08.2019 по 15.03.2022 гг. Поиск по заданному алгоритму позволил выявить 2090 статей, среди которых только 685 содержат результаты редактирования генов 28 видов культурных растений (поиск проведен по 56 культурам). В значительной части этих публикаций рассмотрено либо редактирование генов-мишеней, проведенное ранее в аналогичных работах, либо исследования относились к сфере обратной генетики, и только 136 статей содержат данные о редактировании новых генов-мишеней, модификация которых направлена на улучшение селекционно значимых признаков растений. Всего за весь период применения системы CRISPR/Cas с целью улучшения селекционно значимых свойств редактированию были подвергнуты 287 генов-мишеней культурных растений. В настоящем обзоре представлен подробный анализ редактирования новых генов-мишеней. Чаще всего целью этих работ было повышение урожайности и устойчивости растений к болезням, а также улучшение свойств растительного сырья. Отмечено, удалось ли на момент публикации получить стабильные трансформанты, применялось ли редактирование к немодельным сортам. Существенно расширен спектр модифицированных сортов ряда культур, в частности, пшеницы, риса, сои, томата, картофеля, рапса, винограда, кукурузы. В подавляющем большинстве случаев редактирующие конструкции доставляли с использованием агробактериальной трансформации, реже — биобаллистики, трансфекции протопластов и гаплоиндукторов. Желаемого изменения признаков чаще всего удавалось достичь при помощи нокаута генов. В отдельных случаях осуществляли нокдаун и замены нуклеотидов в гене-мишени. Для получения нуклеотидных замен в генах культурных растений все чаще используют редактирование отдельных оснований (base-editing) и технологию поиска и замены (prime-editing). Появление удобной системы редактирования CRISPR/Cas способствовало развитию молекулярной частной генетики многих культурных видов растений.

Ключевые слова: CRISPR/Cas, биотехнология, гены-мишени, геномное редактирование, культурные растения, направленный мутагенез

DOI: 10.31857/S0026898423030151, **EDN:** CHWGLZ

ВВЕДЕНИЕ

Развитие методов генетического редактирования культурных растений способствует ускорению селекционного процесса и, как следствие, повышению его экономической эффективности. Однако спектр культур, на которых отработаны методики редактирования, остается достаточно ограниченным. Число генов-мишеней, контролирующих проявление хозяйственно ценных признаков, продолжает расширяться, хотя и не быст-

рыми темпами. Часто практически одновременно публикуются результаты редактирования одних и тех же генов-мишеней в растениях с одним и тем же генотипом, что увеличивает число публикаций, но фактически не приводит к получению новых научных результатов. В 2017 г. [1] была начата аналитическая работа по систематизации и каталогизации сведений о подтвержденных генах-мишенях, используемых для улучшения культурных растений, которая получила развитие в 2019 г. [2].

В настоящем обзоре систематизированы и каталогизированы сведения о подтвержденных генах-мишенях, опубликованные за период с 18.08.2019 по 15.03.2022 гг.

По данным поиска в базе данных Scopus (15 марта 2022 г.) количество статей с ключевыми словами “CRISPR” в сочетании с названиями 56 сельскохозяйственных культур составило 2090 (377 в 2019 и 206 в 2017). Однако большинство публикаций не связано напрямую с применением редактирования новых генотипов и новых генов-мишеней. Чаще всего статьи содержали информацию о новых данных для обратной генетики, а не об улучшении свойств растений. Количество статей, описывающих результаты изменения растений с использованием метода редактирования CRISPR/Cas9, составило 685 (131 в 2019 г. и 88 — в 2017 г.). Суммарно число генов-мишеней, редактирование которых проведено с целью улучшения сельскохозяйственных культур, к настоящему моменту составляет 287 (рис. 1). Анализ, проведенный в настоящей работе, добавил в этот каталог еще 206 генов (табл. 1). Число культур, свойства которых были улучшены при помощи редактирования, составило 28, т.е. за 2.5 года увеличилось на 11.

Новые отредактированные культуры — это земляника [71, 72], актинидия [67], помело [70], банан [65, 66], голубика [69], нут [93], брокколи [129], пекинская капуста [130], перец острый [133], люцерна посевная [72], рыжик [134]. Этот перечень показывает возросший интерес к плодовым и ягодным культурам, к развитию селекции на качество плодов.

Всего в программы по улучшению генотипов культурных растений при помощи геномного редактирования вовлечено 186 генотипов (рис. 2).

Разным видам культурных растений свойственна разная степень генотип-специфической эффективности культивирования *in vitro*, трансформации и регенерации. Именно от этого зависит потенциал вовлечения методов геномного редактирования в исследования, ориентированные на практическое применение. Однако следует отметить, что получить новые линии ячменя с помощью редактирования генов удастся, как и прежде, только с использованием модельного сорта Golden Promise [22–25]. Впервые показана возможность редактирования немодельного сорта этой культуры (отечественный сорт Алей [26]) на уровне клетки, но о получении стабильных трансформантов данного сорта пока не сообщалось. За последние годы, только у 23 культур увеличилось число отредактированных образцов.

Расширилось число отредактированных генотипов пшеницы (в 3.4 раза), томатов (в 2 раза), риса (в 1.4 раза), сои (в 6 раз), рапса (в 3 раза), винограда (в 3 раза), картофеля (в 2 раза).

ГЕНЫ-МИШЕНИ И ОСНОВНЫЕ ПРИЗНАКИ, ПОДВЕРГШИЕСЯ МОДИФИКАЦИИ

Геномное редактирование применимо к моно- и олигогенно контролируемым признакам. Часто в качестве мишеней используют гены, регулирующие переход от вегетативной стадии развития растений к генеративной, что позволяет ускорять сроки цветения/созревания. У риса отредактировано три новых гена-мишени, у кукурузы — два, у сои — пять, у рапса, актинидии, голубики и пекинской капусты — по одному (табл. 1). К признакам с олигогенным контролем относятся признаки устойчивости к болезням. В качестве новых мишеней у томата и винограда отредактировано по 10 таких генов, у риса — шесть, у пшеницы — три, у сои и картофеля — по три, у яблони — два, у ячменя, перца, банана и помело — по одному (табл. 1).

В отличие от факторов биотического стресса, устойчивость к абиотическому стрессу чаще всего имеет полигенный характер, при этом среди новых генов-мишеней встречаются гены, редактирование которых повышает устойчивость растений к засухе, холоду и другим неблагоприятным условиям окружающей среды (у томата — пять, у риса — два, у сои и рапса — по одному; табл. 1).

К признакам с полигенным контролем традиционно относится продуктивность, однако, поиск и нокаут генов негативных регуляторов роста и развития позволяет улучшать продуктивность культур. По восемь таких новых мишеней отредактировано у риса и пшеницы, у кукурузы — пять, у земляники, сои и винограда — по три, у ячменя и рапса — по два, у капусты и банана — по одному (табл. 1). Кроме того, на устойчивость к полеганию, а, следовательно, и на повышение урожайности влияют гены короткостебельности. Так, к укорочению стебля у пшеницы (мягкой и твердой) и у риса привело редактирование двух генов.

Среди новых мишеней часто встречаются гены, редактирование которых позволяет улучшить биохимический состав растительного сырья. В число этих генов входят гены, связанные с улучшением свойств крахмала, изменением состава жирных кислот, снижением аллергенности, улучшением ароматических свойств, увеличением сахаристости или содержания белка (табл. 1). Редактированию подвергнуты 13 новых генов-мишеней сои, восемь генов пшеницы, девять томата, пять картофеля, три гена кукурузы, по два у ячменя и хлопчатника, а у риса, ячменя, капусты, винограда, рыжика и брокколи — по одному.

Появление методов геномного редактирования, особенно системы CRISPR/Cas, способствовало развитию частной молекулярной генетики многих культурных видов растений. В первую

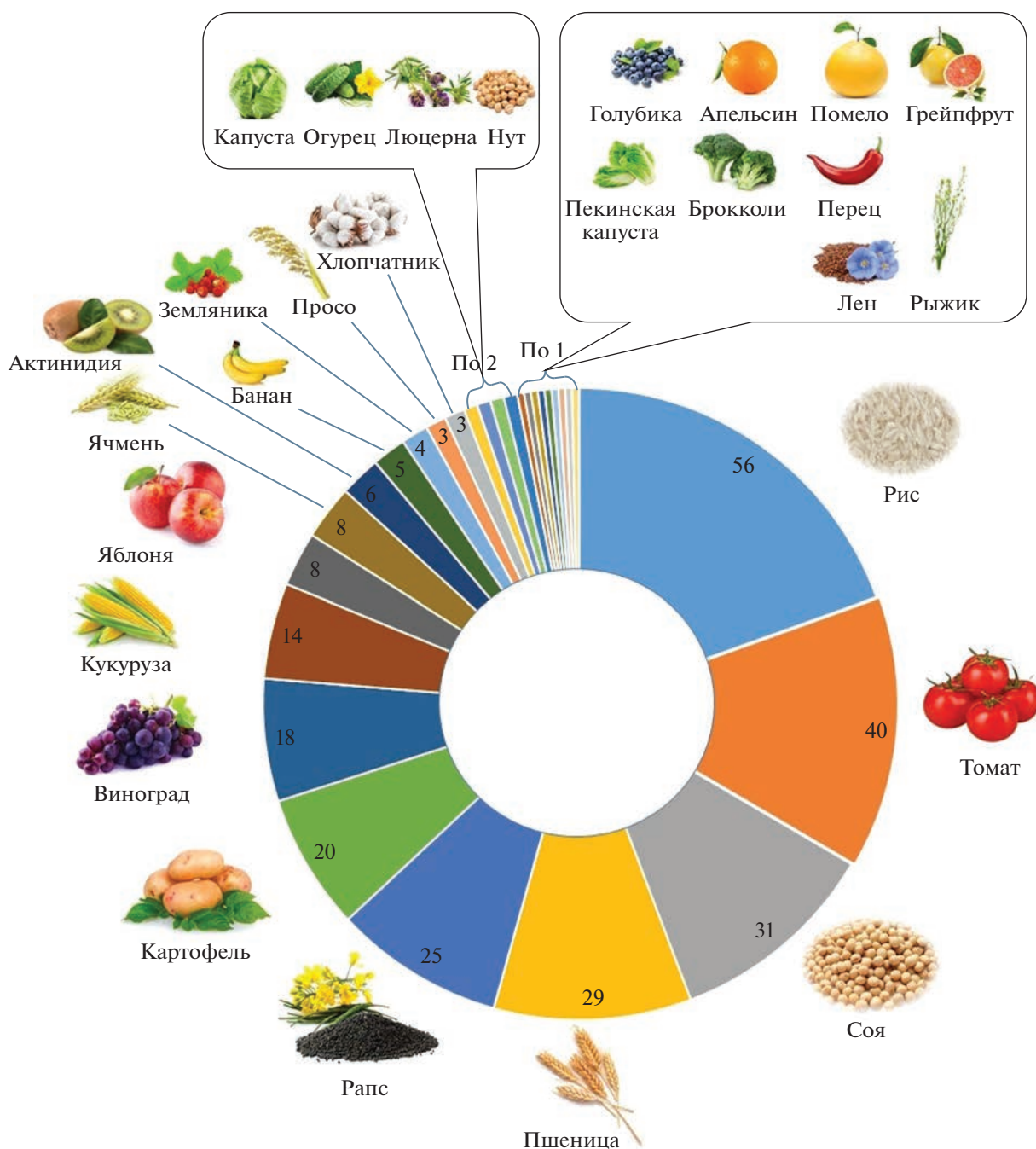


Рис. 1. Число генов-мишеней, модифицированных с использованием системы CRISPR/Cas, с целью изменения хозяйственно ценных признаков (суммарно из обзора [2] и табл. 1 за период с августа 2018 по март 2022 г.).

очередь, это связано с возможностью быстрого внедрения результатов фундаментальных исследований в селекционную практику. А во-вторых, возможности редактирования существенно облегчили проведение исследований в области обратной генетики, проверку генов-кандидатов.

Быстрое расширение списка генов-мишеней, основанное на использовании их ортологов, отражает закономерность, обнаруженную более 100 лет назад Николаем Ивановичем Вавиловым, кото-

рую он назвал Законом гомологических рядов в наследственной изменчивости [137].

Так, сходные нуклеотидные замены в генах *ALS* (кодирующих ацетоллактатсинтазу, ключевой фермент биосинтеза аминокислот с разветвленной цепью – мишеней гербицидов) использовали для повышения гербицидоустойчивости сои [138], риса [139–143], картофеля [144], кукурузы [48, 49], [145] и рапса [146].

Таблица 1. Гены сельскохозяйственных культур, модификация которых с помощью системы CRISPR/Cas привела к улучшению их хозяйственно ценных признаков (по материалам статей в журналах, индексируемых в Scopus, с сентября 2018 по март 2022 г.)

Ген	Функция продукта гена	Тип модификации	Целевое свойство	Сорт/Линия	Способ доставки	Частота мутантных линий	Получение негнотифицированных растений	Ссылка
РНС								
<i>OxFBK4</i>	Регулятор роста растений (нечувствительный к гиббереллину)	Нокаут	Укорочение стебля	RE9586 – мутант сорта Zhonghui8015 (ZH8015)	АТ	Не указано	Да	[3]
<i>OxPUB43</i>	Негативный регулятор массы и размера зерна	Нокаут	Увеличение массы и размера зерна	Nipponbare (Nip), Zhonghua 11 (ZH11)	АТ	Не указано	Да	[4]
<i>OxMORE1a*</i>	Негативный регулятор устойчивости к гемитрофным патогенам	Нокаут	Устойчивость к пирикулярнозу, бактериальному фитоторозу листьев, восприимчивость к корневой пятнистости риса	Dongjin	АТ	Не указано	Да	[5]
<i>PHUS*</i>	Фитохром С, один из ключевых компонентов геновой сети, обеспечивающей переход растений от вегетативной к генеративной фазе	Нокаут	Ускорение перехода к колошению	Nanjing46	АТ	10%	Да	[6]
<i>OxSPL4</i>	Транскрипционный фактор, один из регуляторов роста и развития растений	Делеция в первом экзоне, не приводящие к сдвигу рамки считывания	Увеличение размера и числа зерен в метелке (повышение продуктивности)	Nipponbare	АТ	Не указано	Да	[7]
<i>OxSWEET14</i>	Негативный регулятор устойчивости к бактериальному увяданию	Нокаут (делеция в промоторе 6 п.н.)	Устойчивость к бактериальному увяданию	TBR225	АТ	19–22%	Да	[8]
<i>OxPDCCD5</i>	Один из компонентов регуляторной сети, контролирующей процессы программируемой гибели клеток	Нокаут	Увеличение урожайности за счет плеiotропного эффекта (изменение архитектуры растений, повышение скорости фотосинтеза, увеличение массы и числа зерен в метелке)	Bing1B	АТ	Не указано	Да	[9]
<i>G4200x2 (SD1)</i>	20-оксидиза гиббереллина (один из генов “зеленой революции”)	Нокаут	Укорочение стебля, устойчивость к полеганию (как плеiotропный эффект)	IPB3S	АТ	19,3%	Да	[10]
<i>Pt21, Bsr-d1, Xa5</i>	Негативные регуляторы устойчивости к фитопатогенам	Тройной нокаут	Устойчивость к пирикулярнозу и бактериальному увяданию	Nipponbare	Не указано	7,7%	Да	[11]
<i>Pt21</i>	Негативный регулятор устойчивости к пирикулярнозу	Нокаут	Устойчивость к пирикулярнозу	VP 1636	АТ	66%	Да	[12]
<i>OxHXX1</i>	Один из компонентов сигнальной трансдукции и фосфорилирования глюкозы	Нокаут	Повышение урожайности и эффективности фотосинтеза	Huanghuazhan, Meixiangzhan, Wushansimiao	АТ	Не указано	Да	[13]
<i>OxPIN5b, GS3**, OxMYB30</i>	Негативные регуляторы длины метелки (<i>OxPIN5b</i>), размера зерна (<i>GS3</i>) и холодоустойчивости (<i>OxMYB30</i>)	Тройной нокаут	Повышение урожайности и холодоустойчивости	Nipponbare	АТ	21–66%	Да	[14]
<i>Ehd1*</i>	Один из компонентов геновой сети, обеспечивающей переход растений от вегетативной к генеративной фазе	Делеция со сдвигом и без сдвига рамки считывания	Более раннее цветение, увеличение урожайности (как плеiotропный эффект)	Nipponbare, Longdao16, Longdao24, Xushui134	АТ	46,6–55%	Да	[15]
<i>GCS1*</i>	Регулятор фертильности	Нокаут	Повышение содержания сахарозы в зерне, отсутствие оплодотворения	Tos17_NC0320, Nipponbare	АТ	Не указано	Да	[16]

Таблица 1. Продолжение

Ген	Функция продукта гена	Тип модификации	Целевое свойство	Сорт/Линия	Способ доставки	Частота мутантных линий	Получение нетрансгенных модифицированных растений	Ссылка
<i>O3g0603 100, O3g0568400, GL3.2</i>	Цитохром P-450	Множественный нокаут	Увеличение длины и ширины зерна, массы 1000 зерен (повышение продуктивности)	IR-96	АТ	44.4%	Да	[17]
<i>OxGhd7</i>	Один из компонентов генной сети, обеспечивающей переход растений от вегетативной к генеративной фазе	Нокаут	Более раннее цветение	Daohuaxiang (DHN), DF24	АТ	Не указано	Да	[18]
<i>OxRR2</i>	Регулятор передачи цитокининового сигнала	Нокаут	Повышение солеустойчивости	WPB106	АТ	15.2%	Да	[19]
<i>Ox8N3</i>	Негативный регулятор устойчивости к бактериальному увяданию	Нокаут	Устойчивость к бактериальному увяданию	Kitaake	АТ	25%	Да	[20]
<i>SD1 (GA20Ox2)</i>	20-оксидаза гиббереллина (один из генов "зеленой революции")	Нокаут	Укорочение стебля, устойчивое к полетанию (как плеiotропный эффект)	Kasalath, TeTePu	АТ	21–68%	Да	[21]
ЯЧМЕНЬ								
<i>SSIIa</i>	Крахмалсинтаза	Нокаут	Повышение содержания амилозы	Golden Promise	АТ	18–21%	Да	[22]
<i>MORC1**</i> , <i>MORC6a</i>	Негативные регуляторы устойчивости к патогенам	Нокаут/двойной нокаут	Устойчивость к мучнистой росе и фузариозу	Golden Promise	АТ	58%	Да	[23]
<i>HvCOMT1</i>	Регулятор биосинтеза лигнина	делеция, инсерция	Уменьшение содержания лигнина (повышение эффективности для получения биоэтанола)	Golden Promise	АТ	42%	Да	[24]
<i>NUD</i>	Фактор транскрипции AP2/ERF, регулирующий биосинтез липидов и образование дополнительного слоя клеток между перикарпом и цветковой чешуей	Нокаут	Голозерность	Golden Promise	Биобаллистика листьев/АТ	57%	Да	[25]
<i>NUD</i>	Фактор транскрипции AP2/ERF, регулирующий биосинтез липидов и образование дополнительного слоя клеток между перикарпом и цветковой чешуей	Нокаут	Голозерность	Алей	Транфекция протопластов	6–20%	Нет	[26]
<i>HvCKX1***</i>	Цитоксинидегидрогеназа	Нокаут	Повышение продуктивности	Golden Promise	АТ	81.8%	Да	[27]
<i>HvCKX1</i> , <i>HvCKX3***</i>	Цитоксинидегидрогеназа	Нокаут	Повышение продуктивности	Golden Promise	АТ	до 70%	Да	[28]
ПШЕНИЦА								
<i>TaIPK1</i>	Инозитол-пентакисфат 2-киназа I	Нокаут	Улучшение накопления железа и цинка	Borlaug-2016	АТ	12.7 и 10.8%	Да	[29]
<i>TaARE1</i>	Негативный регулятор усвоения азота	Нокаут	Повышение эффективности усвоения азота, замедление старения, повышение урожайности	Kenong199 (KN199)	Биобаллистика эмбрионидов	Не указано	Да	[30]
<i>TaPinb</i> , <i>TaWaxy (GBSS)</i> , <i>TaDA1</i>	Регуляция размера частиц муки (<i>Pinb</i>), гранул-связанная крахмалсинтаза (<i>GBSS</i>), Негативный регулятор массы зерна	Множественный нокаут	Мутации, которые могут способствовать повышению твердости, увеличению содержания амилпектина в зерне и увеличению показателя "масса тысячи зерен"	Fielder	АТ	54.2%	Да	[31]

Таблица 1. Продолжение

Ген	Функция продукта гена	Тип модификации	Целевое свойство	Сорт/Линия	Способ доставки	Частота мутантных линий	Получение нетрансгенных модифицированных растений	Ссылка
<i>TaASN2</i>	Аспарагинсинтаза	Нокаут	Уменьшение содержания свободного аспарагина в зерне и уменьшение содержания акриламида в выпечке	Cadenza	Биобаллистика эмбрионов	84%	Да	[32]
<i>Ta-eIF4E</i> , <i>Ta-eIF(iso)4E</i>	Негативный регулятор вирусостойчивости	Нокаут	Устойчивость к вирусу желтой мозаики и к вирусу полосатой мозаики	Cezanne, Goncourt, Prevent	Биобаллистика эмбрионов	3–20%	Да	[33]
<i>TaSBE1la</i>	Фермент, разветвляющий крахмальные цепи	Нокаут	Повышение содержания амилозы	Zhengmai 7698 (ZM), Bobwhite	Биобаллистика эмбрионов	38,4%	Да	[34]
<i>TaMs2</i>	Мужская стерильность	Нокаут	Восстановление фертильности	DMSW	АТ	9%	Да	[35]
<i>TaSD1</i> (<i>GA20ox2</i>), <i>TaSD1</i> (<i>GA20ox2</i>), <i>TaBR1</i> , <i>TaBR11</i>	20-оксидлаз гиббереллина (фермент биосинтеза гиббереллинов) и регулятор биосинтеза брассиностероидов	Нокаут всех гомеологичных вариантов	Укорочение стебля, устойчивость к полеганию (как плеiotропный эффект)	Пшеница мягкая яровая BW, W5 и K15, озимая S96 и пшеница твердая D6 и D7	Гаплоиндукция cas9/guide RNA (gRNA)-транскрипционной кукурузой	3,6–50%	Да	[36]
<i>TaGW7</i>	Регулятор клеточного деления	Нокаут гомеологичных копий	Увеличение ширины и массы зерна с одновременным укорачиванием его длины	Bobwhite Chinese Spring	Биобаллистика эмбрионов	1,1–8,3%	Да	[37]
<i>TaQsd2</i>	Аланинаминотрансфераза	Нокаут	Уменьшение рисков прорастания на корню	Fielder	АТ	2,3%	Да	[38]
<i>TaCKX2-1</i> , <i>TaGLW7</i> , <i>TaGW2*</i> , <i>TaGW8</i>	Негативные регуляторы массы зерна и озерненности колоса	Нокауты	Увеличение массы зерна и озерненности колоса (повышение продуктивности)	Fielder	АТ	10%	Да	[39]
<i>TaNFxL1</i>	Негативный регулятор устойчивости к фузариозу	Нокаут	Устойчивость к фузариозу	Roblin	Биобаллистика эмбрионов	Не указано	Да	[40]
<i>TaWTAI-CM3</i> , <i>TaWTAI-CM16</i>	Ингибиторы альфа-амилазы и трипсина	Множественный нокаут	Снижение содержания аллергенов белков	Svevo	Биобаллистика эмбрионов	14,4%	Да	[41]
КУКУРУЗА								
<i>ZmPA17</i>	Негативный регулятор ветвистости метелки	Нокаут	Повышение ветвистости метелки	B73 × CML247	АТ	1%	Да	[42]
<i>ZmPLD3</i>	Фосфолипаза	Нокаут	Повышение эффективности гаплоиндукции	LH244	АТ	до 5%	Да	[43]
<i>BADH2a</i> <i>BADH2b</i>	Бетаинаальдегиддегидрогеназа	Двойной нокаут	Улучшение ароматических качеств	N355, LN005M, XCW175	АТ	Не указано	Да	[44]
<i>ZmNL4</i>	Регулятор клеточного деления	Нокаут	Более узкие листовые пластинки (оптимизация архитектуры растений для более плотных посевов и повышения урожайности с единицы площади)	B73-329	АТ	Не указано	Нет	[45]
<i>ZmCLE1E5</i> , <i>ZmCLE7</i> , <i>ZmFCP1</i>	Регуляторы развития соцветия	Нокауты, множественные нокауты	Повышение озерненности початка (повышение продуктивности)	A188 × B73	АТ	Не указано	Да	[46]
<i>ZmPHYC1</i> , <i>ZmPHYC2</i>	Фитохром С, один из ключевых компонентов генной сети, обеспечивающей переход растений от вегетативной к генеративной фазе	Двойной нокаут	Более раннее цветение	ZC01	АТ	Не указано	Да	[47]

Таблица 1. Продолжение

Ген	Функция продукта гена	Тип модификации	Целевое свойство	Сорт/Линия	Способ доставки	Частота мутантных линий	Получение нетрансгенных модифицированных растений	Ссылка
<i>ZmALS1</i> , <i>ZmALS2</i> **	Ключевой фермент биосинтеза ветвящихся аминокислот – мишеней гербицидов	Нуклеотидные замены (методом prime-editing)	Устойчивость к гербициду хлорсульфурону	W542L, S621I	АТ	до 53.2%	Да	[48]
<i>ZmALS1</i> , <i>ZmALS2</i> **	Ключевой фермент биосинтеза ветвящихся аминокислот – мишеней гербицидов	Нуклеотидные замены (методом base-editing), двойные мутанты	Устойчивость к гербициду хлорсульфурону	ZC01	АТ	13.8%	Да	[49]
<i>Wx (GBSS)</i>	Гранулированная крахмалсинтаза	Нокаут	Увеличение содержания амилопектина в зерне	ZC01-DTMwx × 35Fu, 35M	Не указан	более 20%	Да	[50]
КАРТОФЕЛЬ								
<i>SBE1</i> , <i>SBE2</i>	Ферменты, разветвляющие крахмальные цепи	Нокауты, множественные нокауты	Повышение содержания амилозы	Desiree	АТ/ Трансфекция протопластов	52–72%	Да	[51]
<i>SBE1</i> , <i>SBE2</i>	Ферменты, разветвляющие крахмальные цепи	Нокауты, множественные нокауты	Повышение содержания амилозы	<i>S. phureja</i> , Desiree	АТ Трансфекция протопластов	1.5%	Да	[52]
<i>SBE3</i>	Фермент, разветвляющий крахмальные цепи	Нокаут четырех аллелей	Снижение содержания амилозы	Sayaka	АТ	8%	Да	[53]
<i>SISSR2</i>	Биосинтез стероидных гликокалоидов	Нокаут	Снижение концентрации стероидных гликокалоидов	Atlantic	АТ	46%	Да	[54]
<i>SIppo2</i>	Полифенолоксидаза	Нокаут	Снижение степени потемнения мякоти клубня	Desiree	Трансфекция протопластов	24–68%	Да	[55]
<i>SDND1</i> *, <i>SiCHL1</i> , <i>SiDMR6-1</i>	Негативные регуляторы устойчивости к фитофторозу	Нокауты, множественные нокауты	Устойчивость к фитофторозу	Desiree, King Edward	АТ	0–18%	Да	[56]
ВИНОГРАД								
<i>IdnDH</i>	<i>L</i> -идонатдегидрогеназа	Вставка	Изменение вкусовых характеристик, подавление биосинтеза винной кислоты	Chardonnay	АТ	Не указано	Нет, получены трансгенные растения	[57]
<i>VvDMR6-2</i>	Негативный регулятор устойчивости к мучнистой росе	Гомозиготная делеция 1 или 2 п.н. и гомозиготная вставку 1 п.н.	Повышение устойчивости к мучнистой росе	Crimson seedless	Трансфекция протопластов с последующей регенерацией растений	87.5%	Да	[58]
<i>VvAIR12</i> <i>VvSWE4</i> <i>VvLIN2 VvDEL1</i>	Негативные регуляторы устойчивости к грибным болезням	Делеция генов-мишеней	Повышение устойчивости к патогенам	Thompson Seedless	АТ	0–31.6%	Да	[59]
<i>VvMLO3 VvMLO4</i> **	Негативный регулятор устойчивости к мучнистой росе	Одно-нуклеотидные вставки и делеции различной длины	Повышение устойчивости к мучнистой росе	Thompson Seedless	АТ	0–38.5%	Нет, получены трансгенные растения	[60]
<i>VvEDR2</i>	Негативный регулятор устойчивости к мучнистой росе	Биаллельные мутации	Повышение устойчивости к мучнистой росе	Thompson Seedless	АТ	до 32%	Нет, получены трансгенные растения	[61]
<i>VvPR4b</i>	Негативный регулятор устойчивости к мучнистой росе	Бинуклеотидная делеция (CGCSA>CSA), приводящая к сдвигу рамки считывания	Снижение устойчивости к мучнистой росе	Thompson Seedless	АТ	20.2%	Нет, получены трансгенные растения	[62]

Таблица 1. Продолжение

Ген	Функция продукта гена	Тип модификации	Целевое свойство	Сорт/Линия	Способ доставки	Частота мутантных линий	Получение нетранслируемых модифицированных растений	Ссылка
<i>TAS4 MYB47</i>	Регулятор устойчивости к болезни Парса (Pearse disease), к вирусу красной пятнистости виноградной лозы (GRBV)	Вставка 1–2 п.н., делеция 5–6 п.н.	Повышение устойчивости к патогенам	Подвой 101-14	АТ	до 80%	Нет, получены трансгенные растения	[63]
<i>YsCCD8</i> <i>YsCCD7</i>	Регулятор ветвления побегов	Нокаут	Увеличение степени ветвления побегов, ведущее к повышению продуктивности	Подвой 41В	АТ	8,2%	Нет, получены трансгенные растения	[64]
БАНАН								
<i>MusaDMR6</i>	Регулятор устойчивости к бактериальному увяданию	Делеции, вставки, замены или делеции, сопровождающиеся вставками	Повышение устойчивости к патогенам	<i>Musa balbisiana</i> Сорт Sukali Ndiizi	АТ	100%	Да	[65]
<i>MaACO1</i>	Путь синтеза этилена, контроль созревания	Вставки	Повышение продуктивности	<i>Musa acuminata</i> сорт Brazilian	АТ	98,4–98,6%	Да	[66]
АКТИНИДИЯ								
<i>CEN</i> <i>CEN4</i> <i>SyGl</i>	Синтез цитокининов	Нокаут	Повышение продуктивности, получение более компактных растений, более раннее цветение	<i>Actinidia chinensis</i> Ranch. var. <i>chinensis</i> сорт Bruce	АТ	40,7%	Да	[67]
ЯБЛОНЯ								
<i>MaDIPM4**</i>	Устойчивость к бактериальному ожогу	Несколько типов мутаций: небольшая вставка (+1 н.), небольшая замена (R1 н.), маленькие и большие делеции	Повышение устойчивости к бактериальному ожогу	Gala, Golden Delicious	АТ	75%	Да	[68]
ГОЛУБИКА								
<i>CEN</i>	Регулятор роста и развития растения	Вставки/делеции длиной от 1 до 2 п.н.	Повышение продуктивности, получение более компактных растений, более раннее цветение	Blue Muffin, O'Neal.	АТ	17–37,5%	Да	[69]
ПОМЕЛО								
<i>LOB1</i>	Устойчивость к раку цитрусовых, вызванному <i>Xanthomonas citri</i> subsp. <i>citri</i>	Делеции и вставки	Повышение устойчивости к раку цитрусовых	<i>Citrus maxima</i>	АТ****	87,5%	Да	[70]
ЗЕМЛЯНИКА								
<i>TAA1</i> <i>ARF8</i>	Биосинтез ауксинов TAA1 и фактора ответа на ауксин 8 (ARF8)	Делеции и замены	Повышение продуктивности	<i>Fragaria vesca</i>	АТ	49–75%	Да	[71]
<i>FveYUC10</i>	Регулятор развития листьев, корней, цветков и плодов	Нокаут	Повышение продуктивности	Hawaii 4	АТ	Около 5%	Да	[72]
ЛЮЦЕРНА ПОСЕВНАЯ								
<i>ALS1</i> и <i>ALS2</i>	Ацетолактатсинтаза	Замена одного нуклеотида (С на Т) для изменения одной аминокислоты (Pro → Leu)	Устойчивость к гербицидам сульфониломочевинного и имидазопиридина ряда	Клон C23	АТ	6–11%	Да	[73]
СОЯ								
<i>Pod dehiscence 1</i> (<i>PDH1</i>) gene	Контроль растрескиваемости бобов	Нокаут	Уменьшение растрескиваемости бобов	Huachun6 (HSC6)	АТ	48%		[74]

Таблица 1. Продолжение

Ген	Функция продукта гена	Тип модификации	Целевое свойство	Сорт/Линия	Способ доставки	Частота мутантных линий	Получение нетрансгенных модифицированных растений	Ссылка
<i>GmAMS1</i> , <i>GmAMS2</i>	Развитие микроспор и клеток тапетума	Нокаут одного или двух генов	Получение стерильных линий на разнообразной генетической основе (контроль опыления в гибридной селекции)	Williams 82	АТ	14–25%	Да	[75]
<i>GmAITRs</i>	Транскрипционный фактор, принимающий участие в регуляции сигнала абсцизовой кислоты	Нокаут шести копий гена	Повышение засухо- и солеустойчивости	Williams 82	АТ	Не указано	Нет	[76]
<i>GmAGGE1</i> (<i>GmJAG</i>)	Регулятор числа семян на бобе, а также формы листьев	Замена нуклеотида С на G приводит к замене аминокислоты в консервативном домене EAR GmJAG1, что ведет к потере функции гена (две копии гена)	Увеличение числа семян на бобе, повышение урожайности	Huachun 6	АТ	Не указано	Да	[77]
<i>Night light-inducible and clock-regulate D 2 (LNK2)</i> , 4 гена	Контроль циркадного ритма растения	Нокаут четырех копий гена	Ранний переход к цветению в условиях длинного дня	Williams 82	АТ	92.3%	Да	[78]
<i>GmFATB1a</i> , <i>GmFATB1b</i>	Ферменты, участвующие в биосинтезе жирных кислот (тиоэстеразы)	Нокаут	Снижение содержания насыщенных жирных кислот (пальмитиновой и стеариновой) в соевом масле при стабильном уровне ненасыщенных кислот	Williams 82	АТ	Не указано	Нет	[79]
<i>GmKIX-1</i>	Регулятор размера органов растения, негативный регулятор пролиферации клеток	Нокаут	Увеличение размера листьев и веса семян (повышение продуктивности)	Maverick	АТ	Не указано	Да	[80]
<i>Gly m Bd 30 K</i>	Транспортёр жиров	Нокаут	Снижение вероятности возникновения аллергии	Jack	Биобаллистика эмбрионов	Не указано	Да	[81]
<i>GmGOLS1A</i> , <i>GmGOLS1B</i>	Ферменты биосинтеза олигосахаридов семейства рафинозы	Нокаут	Уменьшение содержания олигосахаридов семейства рафинозы (рафинозы, стахиозы, вербаскозы) (RFO) для улучшения питательных качеств	Maverick, DT26	АТ	До 100%	Нет	[82]
<i>Gly m Bd 28K</i> и <i>Gly m Bd 30K</i>	Основные аллерегены сои – вицилин-подобный гликопротеин Gly m Bd 28 K и белок, переносящий жиры Gly m Bd 30 K	Нокаут	Снижение вероятности возникновения аллергии на сою, получение гипоаллергенных сортов	Enrei, Kariyutaka	АТ	До 100%	Да	[83]
<i>GmAP1</i>	Транскрипционный фактор семейства MADS-box, определяющий специфичность меристемы цветка	Нокаут четырех копий гена	В условиях короткого дня поздний переход к цветению, большее число узлов и междоузлий, что непосредственно связывают с числом бобов на растении и урожайностью	Williams 82	АТ	20%	Да	[84]
<i>GmPRR37</i>	Супрессор перехода к цветению в условиях длинного дня	Нокаут	Переход к раннему цветению у мутантов очень позднецветущего сорта в условиях длинного дня	Zigongdongdou (ZGDD), Jack	АТ	Не указано	Нет	[85]
<i>GmLox</i> genes (<i>GmLox1</i> , <i>GmLox2</i> , <i>GmLox3</i>)	Липооксигеназы	Нокаут трех генов	Улучшение вкусовых качеств	Huachun 6	АТ	78.9%	Да	[85]
<i>GmF3H</i> , <i>GmF3H2</i> <i>GmFNSII-1</i>	Ферменты биосинтеза изофлавоноидов	Нокаут трех генов	Увеличение содержания изофлавоноидов в листьях и увеличение устойчивости к вирусу мозаики	Jack	АТ	44.4%	Да	[86]

Таблица 1. Продолжение

Ген	Функция продукта гена	Тип модификации	Целевое свойство	Сорт/Линия	Способ доставки	Частота мутантных линий	Получение нетрансгенных модифицированных растений	Ссылка
<i>FAD2-1A</i>	Десатураза	Нокаут	Увеличение содержания олеиновой кислоты	Nuaxia 3	АТ	Не указано	Нет	[87]
<i>GmFAD2-1B</i>	Десатуразы жирных кислот	Нокаут	Уменьшение содержания линолевой и линоленовой кислот	Maverick	АТ	12%	Да	[88]
<i>GmFAD2-2A</i>	Десатуразы жирных кислот	Нокаут	Увеличение содержания белка и снижение содержания линолевой кислоты	Jinong 38	АТ	55.5–95%	Да	[89]
<i>GmFT2a</i> и <i>GmFT3a</i>	Активаторы перехода к цветению	Нокаут одного или двух генов	Увеличение адаптационного потенциала (в зависимости от длины дня)	Jack	АТ	Не указано	Нет	[90]
<i>E1</i>	Регулятор перехода к цветению	Нокаут	Переход к раннему цветению в условиях длинного дня	Jack	АТ	75%	Да	[91]
<i>GmSPL9</i>	Транскрипционный фактор, регулятор перехода с ювенильной стадии	Нокаут	Изменение архитектуры растения, увеличение числа узлов главного стебля, ветвей	Williams 82	АТ	Не указано	Нет, получены трансгенные растения	[92]
НУТ								
<i>4CL</i> , <i>RIE7</i>	Регуляция накопления лигнина в условиях стресса; регуляция циркадных ритмов	Нокаут	Устойчивость к засухе	Kabuli	Трансфекция протопластов	Не указано	Нет	[93]
ТОМАТ								
<i>LeRBOH</i> <i>LeRBOHE</i>	Регуляторы накопления АФК в пыльниках	Нокаут	Создание форм с мужской стерильной зародышевой плазмой для использования в гибридной селекции	Alisa Craig	АТ	Не указано	Да	[94]
<i>SINVINH1</i>	Ингибитор инвертазы	Нокаут	Повышение содержания сахарозы и глюкозы в плодах	Suzukoma	АТ	Не указано	Да	[95]
<i>SINVINH1</i> , <i>SIVPE5</i>	Ингибитор накопления растворимых сахаров	Нокаут <i>SINVINH1</i> и <i>SIVPE5</i>	Повышение содержания растворимых сахаров в плодах	M82	АТ	42.9–61.5%	Да	[96]
<i>More Axillary Growth 1 (MAX1)</i>	Ингибитор уровня оранбохола	Нокаут	Устойчивость к корневому паразитическому сорняку египетской заразице (<i>Phelipanche aegyptiaca</i>)	МП-1	АТ	20%	Да	[97]
<i>SIWAT1</i>	Регуляция содержания ауксина	Нокаут	Повышение устойчивости к вертициллезу (<i>Verticillium dahliae</i> , <i>V. albo-atrum</i>) и фузариозу (<i>Fusarium oxysporum</i>)	Money maker	АТ	Не указано	Да	[98]
<i>SIBES1</i>	Участие в сигнальном каскаде стероидов	Нокаут	Получение более плотных плодов и увеличение сроков хранения	Alisa Craig	АТ	Не указано	Да	[99]
<i>DMR6</i>	Восприимчивость к бактериальным грибным заболеваниям	Нокаут	Устойчивость к широкому спектру болезней	Fia. 8000	АТ	Не указано	Нет	[100]
<i>SIHyPRP1 domain(s)</i>	Синтез гликопротеинов клеточной стенки	Нокаут делеция	Устойчивость к солевому стрессу	Hongkwang 15T01	АТ	не указано	Да	[101]
<i>Ms10³⁵</i> , <i>GST4A</i>	Мужская стерильность, кодирование глутатион S-трансферазы	Двойной нокаут <i>Ms10³⁵</i> и <i>GST4A</i>	Создание линий томатов с мужской стерильностью	Линия 6Q71и 5N21	АТ	46.7 и 47.4% для <i>Ms10³⁵</i> ; 28 и 26.3% для гена <i>GST4A</i>	Да	[102]
<i>eIF4E1</i>	Негативный регулятор устойчивости к вирусным патогенам	Нокаут	Повышение устойчивости к Y-вирусу картофеля и к вирусу мозаики огурца	Линия S8	АТ	15%	Да	[103]
<i>eIF4E1</i>	Негативный регулятор устойчивости к вирусным патогенам	Нокаут	Устойчивость томата к вирусу крапчатости перца	Micro-Tom	АТ	72%	Да	[104]

Таблица 1. Продолжение

Ген	Функция продукта гена	Тип модификации	Целевое свойство	Сорт/Линия	Способ доставки	Частота мутантных линий	Получение нетрансгенных модифицированных растений	Ссылка
<i>eIF4E2</i>	Негативный регулятор устойчивости к вирусным патогенам	Нокаут	Устойчивость томата к вирусу крапчатости перца	West Virginia 106 (WVa106)	АТ	Не указано	Да	[105]
<i>SIMS10</i>	Транскрипционный фактор bHLH, регулирующий продолжительность жизни клеток	Нокаут	Получение форм с мужской стерильностью	KS-13	АТ	46.6%	да	[106]
<i>GID1a</i>	Рецептор гиббереллина	Нокаут	Повышение устойчивости к засухе	M82	АТ	Не указано	Да	[107]
<i>PMR4</i>	Кодирование каллозасинтазы	Нокаут	Повышение устойчивости к мучнистой росе	MoneyMaker	АТ	Не указано	Нет	[108]
<i>rgCaM</i>	Кальмодулин-подобный белок	Нокаут	Повышение устойчивости к вирусу желтой курчавости листьев	MoneyMaker	АТ	Не указано	Да	[109]
<i>CCD8</i>	Каротиндиоксигеназа 8	Нокаут	Повышение устойчивости к паразитическому сорняку египетской зарази	Линия T5	АТ	71.4%	Да	[110]
<i>SLAZ2</i>	Негативный регулятор биосинтеза жасмоната	Нокаут	Устойчивость к бактериальной пятнистости	MoneyMaker	АТ	Не указано	Да	[111]
<i>SGR1, LCY-E, Bc, LCY-B1, LCY-B2</i>	Биосинтез ликопина (<i>SGR1</i> -encoding A STAY-GREEN protein, lycopene ϵ -cyclase, beta-lycopene cyclase, lycopene β -cyclase $\frac{1}{2}$)	Нокаут	Повышение содержания ликопина	Ailsa Craig	АТ	0–95.8%	Да	[112]
РАПС								
<i>BnaSVp</i>	Регулятор продолжительности вегетативной стадии	Нокаут	Переход к раннему цветению	Линия J9707	АТ	45%	Да	[113]
<i>BnF5H</i>	Ферулят-5-гидроксилаза	Нокаут	Снижение содержания лигнина S/G и повышение устойчивости к белой плесени	Westar	АТ	Не указано	Да	[114]
<i>BnaA03.BP</i>	Регулятор морфогенеза листьев и изгиба цветоножки	Нокаут	Получение компактной формы куста	Zhong shuang11	АТ	79%	Да	[115]
<i>BnQCR8</i>	Консервативная субъединица комплекса цитохрома b-c1	Снижение числа копий <i>BnQCR8</i>	Повышение устойчивости к белой плесени и серой гнили	Westar	АТ	60.9%	Да	[116]
<i>BnaC04EPPS</i>	Енолпирувилликимат-3-фосфатакиназа	Нокаут	Устойчивость к глифосату	627R	АТ	20%	Да	[117]
<i>BnSHPIA09, BnSHPIA04, BnSHPIA04-A, BnSHPIA04-B, BnSHPIA05</i>	Контроль раскрываемости стручков	Нокаут	Предотвращение растрескивания стручков	Zhongshuang 6 (ZS6)	АТ	27.5–35%	Да	[118]
<i>CRT1a</i>	Кальтрекуллин	Нокаут	Устойчивость к вертициллезу (возб. <i>Verticillium longisporum</i>)	Express 617	АТ	20%	Да	[119]
<i>BnITPK</i>	Инозитол-тетрафосфатакиназа	Нокаут трех паралогов <i>BnITPK</i>	Уменьшение содержания фитиновой кислоты в семенах – улучшение качества для корма	Haydn	АТ	не указано	Да	[120]
<i>BnSFAR1, BnSFAR4, BnSFAR5</i>	Эстеразы	Нокаут	Повышение содержания масел в семенах	Линия RS306	АТ	0.2–1.1%	Да	[121]
<i>BnALS1</i>	Ацетолактатакиназа	Замена P197S	Повышение устойчивости к гербицидам	Линия J9712	АТ	3.2%	Да	[122]
<i>BnTT2</i>	Биосинтез проантоцианидинов	Нокаут	Увеличение содержания линолевой и линоленовой кислот	Линия J9712	АТ	4.9%	Да	[123]

Таблица 1. Окончание

Ген	Функция продукта гена	Тип модификации	Целевое свойство	Сорт/Линия	Способ доставки	Частота мутантных линий	Получение нестранных модифицированных растений	Ссылка
<i>bnaab-rga-D</i>	DELLA белок	Нокаут	Засухоустойчивость, желтые семена	Westar	АТ	Не указано	Да	[124]
<i>BnTT8</i>	Биосинтез проантоцианидинов	Двойной нокаут	Повышение содержания масел и белков в семенах	Линия J9707	АТ	10.4%	Да	[125]
<i>BnaMAX1</i>	Моноксигеназа цитохрома P450	Нокаут двух гомологов <i>BnaMAX1</i>	Улучшение архитектуры, повышение продуктивности	862	АТ	56.3–67.4%	Да	[126]
<i>BnaEOD3</i>	Цитохром P450/CYP78A6	Нокаут четырех гомологичных локусов <i>BnaEOD3</i>	Увеличение числа и массы семян (повышение продуктивности)	Линия J9707	АТ	45.6%	Да	[127]
<i>BnWRKY70</i>	Устойчивость к патогенам	Инсерции, делеции	Повышение устойчивости к белой гнили	Линия J9712	АТ	50%	Да	[128]
БРОККОЛИ								
<i>BaMYB28</i>	Регулятор транскрипции глюкозинолатного пути	Делеция 9 п.н. в 3 экзоне	Повышенное содержание глюкорафанина	Beneforté, Kingdome (KD)	Транфекция протопластов	До 100%	Да	[129]
ПЕКИНСКАЯ КАПУСТА								
<i>BnaFLC2, BnaFLC3</i>	Регуляция биосинтеза глюкозинолата (A-GSL)	Двойной нокаут <i>BnaFLC2</i> и <i>BnaFLC3</i>	Переход к раннему цветению	Линия Kenshin (<i>Brassica rapa</i> L. ssp. <i>pekinensis</i>)	АТ	97.7–100%	Да	[130]
КАПУСТА								
<i>MYB28</i>	Регулятор транскрипции глюкозинолатного пути	Нокаут	Улучшение продуктивности	Дигаллоид AG DH1012 (<i>Brassica oleracea</i> ssp. <i>alboglabra</i> (A12DHd) и <i>B. oleracea</i> ssp. <i>italica</i> (Green Duke GDDH33))	АТ	12%	Да	[131]
<i>BoCER1</i>	Биосинтез кутикулярного воска	Нокаут	Развитие бриллиантовой зеленой капусты (безвосковой)	Линия 'CW1'	АТ	3.4%	Да	[132]
ПЕРЕЦ								
<i>CaERT28</i>	Первичный негативный регулятор устойчивости к антракнозу	Делеция, инсерции	Повышение устойчивости к антракнозу	Arka Lohit (AL) (<i>Capsicum annuum</i>)	АТ	72.5–85.7%	Да	[133]
РЫЖИК								
<i>CsFAD2</i>	Десатуразы жирных кислот	Нокаут	Увеличение содержания мононенасыщенных жирных кислот	Same, Suneson	АТ	0.4–84.1%	Да	[134]
ХЛОПЧАТНИК								
<i>GoPGF</i>	Регулятор образования госсиполов	Нокаут	Удаление госсипола из семян хлопчатника	Stoneville 7A	АТ	Не указано	Да	[135]
<i>GnFAD2</i>	Десатураза жирных кислот	Нокаут	Увеличение содержания олеиновых кислот	Jin668	АТ	71.5%	Да	[136]

* Гены с неоднозначным эффектом, нокаут которых приводит к улучшению одних признаков, но ухудшает показатели по другим.
** Гены, которые и ранее использовали в качестве мишеней, но в новых работах они включены вместе с новыми генами для мультиплексного нокаута.
*** Ожидаемое целевое свойство не достигнуто, так как не нокаутированные гомологи компенсируют отсутствие продукта нокаутированного гена; сделано заключение о необходимости нокаута большего числа гомологов.
**** Применен другой белок (не Cas9. АТ – агробактериальная трансформация штаммами *A. tumefaciens* (если не указано иное).

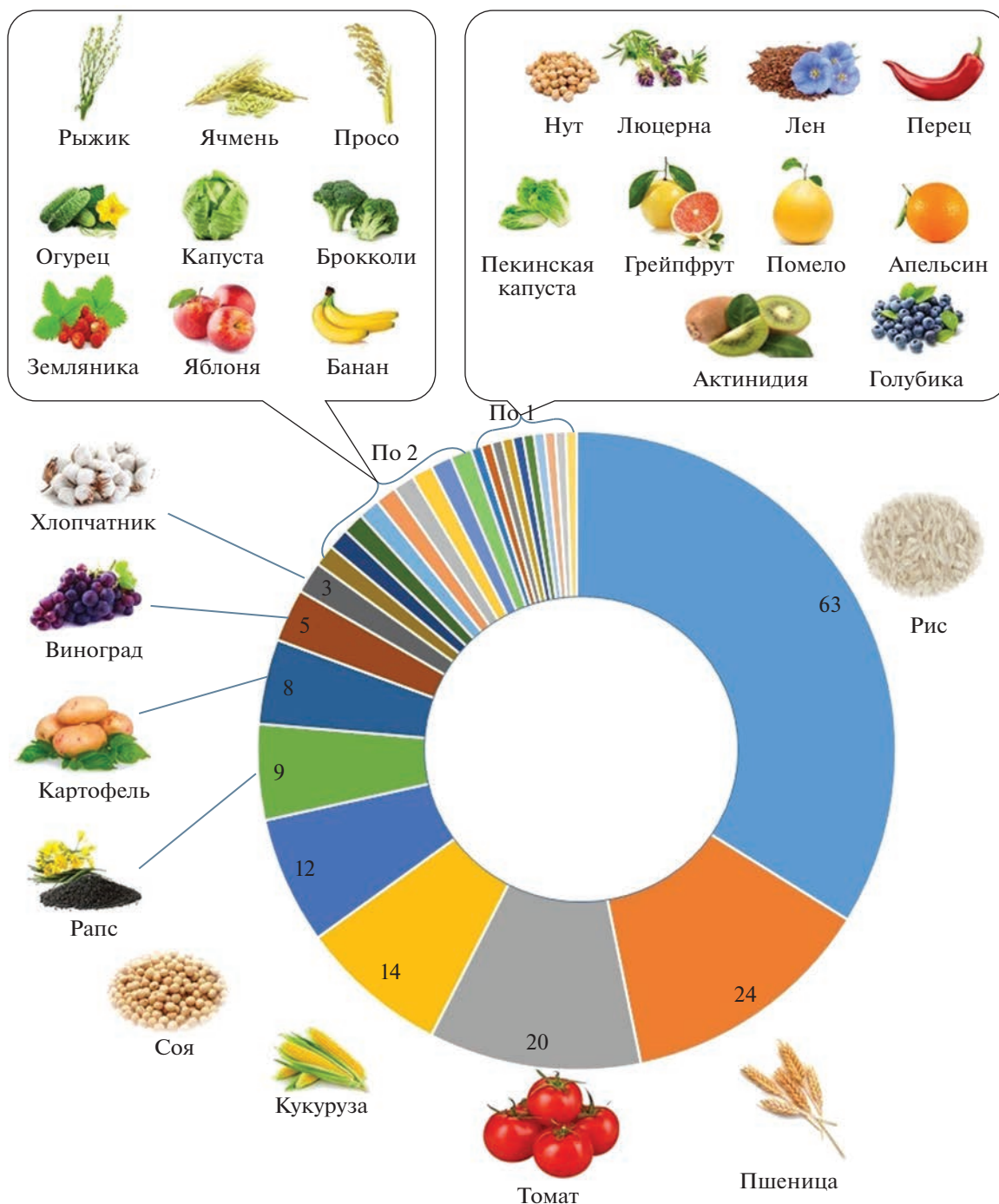


Рис. 2. Число генотипов, модифицированных с использованием системы CRISPR/Cas с целью изменения хозяйственно ценных признаков (суммарно из обзора [2] и табл. 1 за период с августа 2018 по март 2022 г.).

С целью повышения содержания амилопектина осуществлен нокаут генов крахмалсинтазы, связанной с крахмальными гранулами (*GBSS/Waxy/Wx*), у риса [147–149], картофеля [150–152], кукурузы [153] и пшеницы [39]. Проведен нокаут генов десатуразы жирных кислот *FAD2* (для улучшения состава жирных кислот) у риса [154], рапса [155], хлопчатника [136], сои [87–89] и рыжика [134].

Нокаут генов бетаинальдегид-дегидрогеназы позволил улучшить ароматические свойства зерна кукурузы [45].

В растениях огурца [156], пшеницы [33] и томата [103–105] осуществлен нокаут гена *eIF* (негативного регулятора вирусоустойчивости); генов *LOB1* (негативный регулятор устойчивости к раку

цитрусовых) в растениях грейпфрута [157], апельсина [159] и помело [70], а также двух генов негативных регуляторов устойчивости к мучнистой росе: *MLO* — у винограда [60, 159], томата [160,], пшеницы [161]; *EDR* — у пшеницы [162] и винограда [61].

При помощи нокаута генов *CEN* изменена архитектура растений киви и голубики [67, 69].

Гены “зеленой революции” *SD1*, нокаут которых приводит к короткостебельности и устойчивости к полеганию, описаны в работах по редактированию риса [163] и пшеницы [36]. Сорты риса с рецессивным аллелем гена *SD1* (которые участвовали в зеленой революции) были известны и ранее, тогда как у пшеницы в зеленой революции в свое время участвовал совсем другой ген, несущий мутацию приобретения функции вместо мутации потери функции [164]. У гексаплоидной пшеницы три копии гена *SD1*, поэтому вероятность потери функции всеми тремя копиями ничтожно мала. Такая возможность появилась при разработке методов геномного редактирования.

При проведении нокаута у полиплоидных видов растений (пшеница, картофель) редактирующую конструкцию часто приходится направлять сразу на несколько аллельных/гомеоаллельных генов (см. табл. 1). Еще и поэтому система CRISPR/Cas с ее возможностью мультиплексного редактирования стала настоящей находкой для растений, у которых часто встречается полиплоидия. Мультиплексное редактирование разных генов все активнее применяется и для улучшения сортов культурных растений как по одному [36], так и по разным [14, 165] признакам.

НОКДАУН vs НОКАУТА И ПЛЕЙОТРОПНЫЕ ЭФФЕКТЫ ГЕНОВ-МИШЕНЕЙ

Важное значение имеет комплексная оценка мутантов и изучение возможного плейотропного действия генов-мишеней. Тщательное изучение отредактированных растений показывает, что далеко не всегда внесенные модификации вызывают однозначный эффект — улучшая одни признаки, они могут негативно влиять на другие. Так, в работе Li и соавт. [6] подчеркнуто, что при использовании мутантов риса по гену *PHYC*, отличающихся более ранним переходом к колошению, следует учитывать условия окружающей среды и стараться подбирать микроклиматические условия с умеренными температурами во время налива зерна во избежание ухудшения его качества. Сообщается также, что нокаут гена *OsMORE1a* (устойчивость к поражению гемифототрофными фитопатогенами), несмотря на его положительный эффект, не лишен таких побочных эффектов, как повышенная восприимчивость к некротрофному патогену [166].

При выборе в качестве мишеней генов, являющихся негативными регуляторами, важное значение может иметь более тонкое изменение настройки генов экспрессии, чем просто нокаут. Так, делеции в первом экзоне, не приводящие к сдвигу рамки считывания, позволили модифицировать функцию гена *SP4* таким образом, что увеличился размер и число зерен в метелке, а следовательно, повысилась продуктивность растений риса [7]. При этом были изучены эффекты делеций разной длины и показан больший эффект делеции 3 п.н., чем 15 п.н. Нокаут гена негативного регулятора *OsIRO3* [167] не привел к желаемому эффекту, зато нокдаун этого гена, выявленный среди сортов, полученных методами традиционной селекции), способствовал повышению толерантности к защелачиванию почв. Следовательно, в дальнейшем редактирование можно использовать для внесения необходимых изменений в промотор гена *OsIRO3*. Так, делеция 6 п.н. в промоторе *OsSWEET14* риса, полученная с использованием системы CRISPR/Cas, привела к нокдауну этого гена, что способствовало снижению восприимчивости к возбудителю бактериального увядания. Проведенная при этом комплексная проверка мутантных растений не выявила какого-либо побочного негативного воздействия на основные хозяйственно ценные признаки [8]. Нокаут одного из структурных генов фенилпропаноидного пути (*C3'H*) приводил к получению стерильных карликовых растений риса, тогда как нокдаун этого гена при помощи РНК-интерференции имел положительные эффекты. Потенциально редактирование можно использовать для того, чтобы подобрать необходимые модификации в промоторе *C3'H* для достижения эффекта, аналогичного полученному при помощи РНК-интерференции [168].

Есть, однако, и примеры, когда преютропное действие нокаута оказывает положительный кумулятивный эффект. Так, изучение генетической системы, регулирующей процессы программируемой гибели клеток, позволило выбрать удачную мишень — ген *OsPDCD5* риса, нокаут которого привел к повышению урожайности риса за счет плейотропного действия (изменения архитектуры растений, повышения скорости фотосинтеза, увеличения массы и числа зерен в метелке). Аналогичным образом нокаут гена гексокиназы (*OsHKK1*), одного из компонентов сигнальной трансдукции и фосфорилирования глюкозы, позволил повысить урожайность за счет одновременного увеличения озерненности метелки, удлинения эффективных побегов метелки, повышения эффективности фотосинтеза; при этом качество зерна у мутантных линий не уступало качеству исходных линий.

МЕТОДЫ ДОСТАВКИ И СПОСОБЫ РЕДАКТИРОВАНИЯ

В подавляющем большинстве случаев редактирующую конструкцию доставляют в растения с помощью агробактериальной трансформации (87.6% работ; табл. 1), реже — биобаллистики (6.6%), трансфекции протопластов (5.1%) и с использованием гаплоиндукторов (0.7%).

Агробактериальную трансформацию проводят как на зародышах [45, 47, 117], так и на каллусах [168, 169]. Наиболее часто применяют штаммы *Agrobacterium tumefaciens* [18, 54, 127, 132], реже — *A. rhizogenes* [82, 88].

Метод биобаллистики чаще всего используют для доставки редактирующей конструкции в клетки пшеницы [32–34], сои [81], брокколи [129]. Проводят также трансфекцию протопластов ячменя [26], пшеницы [37], картофеля [51, 52, 55] и нута [93], при помощи гаплоиндукции доставляют редактирующие конструкции в растения пшеницы [36].

Не во всех работах представлены данные о растениях-регенерантах и стабильных трансформантах (табл. 1). Вероятно, это связано с желанием зафиксировать приоритетность выбора нового гена-мишени. Это несколько снижает общий уровень опубликованных исследований.

Частота получения мутантных линий сильно зависит от генотипа. Так, например, частота мутаций среди семи трансгенных растений помело составила 100% (#Pum_т 1–6 и 8), в то время как у одного растения мутации не были обнаружены [72].

На эффективность трансформации может влиять и сама редактирующая конструкция. У растений T0 люцерны посевной частота мутаций редактируемых генов-мишеней (*Gmams1*, *Gmams2* или *Gmams1,2*) в зависимости от выбранной конструкции составила 52% (32 из 61), 40% (21 из 52) и 45% (30 из 66) соответственно [75]. Эффективность редактирования одного из целевых генов — *GmFAD2-1A* или *GmFAD2-2A* — в растениях сои составила 95 и 55.56% соответственно, при этом в 66.67% случаев отмечено появление двойных мутантов [89].

Желаемого изменения признаков чаще всего удавалось достичь при помощи нокаута генов (табл. 1). В отдельных случаях свойство белкового продукта изменяли путем небольших делеций в кодирующих последовательностях, не влияющих на сдвиг рамки считывания [7], или изменяли экспрессию генов путем внесения делеций в промоторы [8]. В единичных случаях осуществляли нуклеотидные замены. Причем в сравнении с работами, опубликованными до 2019 года, когда одонуклеотидные замены вносили при помощи гомологической рекомбинации с использованием донорного фрагмента, в более свежих работах

наблюдается технологический прорыв, связанный с применением редактирования отдельных оснований (base-editing) и редактирования с помощью поиска и замены (prime-editing) [48, 49].

ЗАКЛЮЧЕНИЕ

К настоящему моменту достигнуты существенные успехи в улучшении сортов почти 30 сельскохозяйственных культур при помощи редактирования генов. Чаще всего эти работы направлены на повышение урожайности и устойчивости к болезням, а также на улучшение свойств растительного сырья. Появление удобной системы редактирования CRISPR/Cas способствовало активному развитию частной молекулярной генетики многих культурных видов растений. Перспектива дальнейшего развития практико-ориентированных работ по редактированию генов связана с расширением применения таких способов модификации, как base-editing и prime-editing, позволяющих вносить нуклеотидные замены. Следует ожидать расширения спектра редактируемых мишеней за счет новых генов, связанных с улучшением признаков, путем замены в них нуклеотидов, а не нокаута. Возможности широкого практического применения будут зависеть от степени преодоления генотипзависимой эффективности трансформации и регенерации конкретной культуры. Определенные перспективы связаны также с диверсификацией способов доставки редактирующих конструкций.

Исследование выполнено при финансовой поддержке Российского научного фонда (№ 21-66-00012).

Настоящая статья не содержит описания исследований, выполненных с участием людей или использованием животных в качестве объектов.

Авторы заявляют об отсутствии конфликта интересов.

СПИСОК ЛИТЕРАТУРЫ

1. Korotkova A.M., Gerasimova S.V., Shumny V.K., Khlestkina E.K. (2017) Crop genes modified using the CRISPR/Cas system. *Russ. J. Genet.* 7(8), 822–832. <https://doi.org/10.1134/S2079059717050124>
2. Короткова А.М., Герасимова С.В., Хлесткина Е.К. (2019) Текущие достижения в области модификации генов культурных растений с использованием системы CRISPR/Cas. *Вавил. журн. генет. селекции.* 23(1), 29–37.
3. Zegeye W.A., Chen D., Islam M., Wang H., Riaz A., Rani M.H., Hussain K., Liu Q., Zhan X., Cheng S., Cao L., Zhang Y. (2022) *OsFBK4*, a novel GA insensitive gene positively regulates plant height in rice (*Oryza sativa* L.). *Ecol. Genet. Genom.* 23. <https://doi.org/10.1016/j.egg.2022.100115>

4. Wu Q., Liu Y., Huang J. (2022) CRISPR-Cas9 mediated mutation in *OsPUB43* improves grain length and weight in rice by promoting cell proliferation in spikelet hull. *Int. J. Mol. Sci.* **23**(4), 2347. <https://doi.org/10.3390/ijms23042347>
5. Kim C.Y., Park J.Y., Choi G., Kim S., Vo K.T.X., Jeon J.S., Kang S., Lee Y.H. (2022) A rice gene encoding glycosyl hydrolase plays contrasting roles in immunity depending on the type of pathogens. *Mol. Plant Pathol.* **23**(3), 400–416. <https://doi.org/10.1111/mpp.13167>
6. Li B., Du X., Fei Y., Wang F., Xu Y., Li X., Li W., Chen Z., Fan F., Wang J., Tao Y., Jiang Y., Zhu Q.-H., Yang J. (2021) Efficient breeding of early-maturing rice cultivar by editing *PHYC* via CRISPR/Cas9. *Rice*. **14**, 86. <https://doi.org/10.1186/s12284-021-00527-3>
7. Hu J., Huang L., Chen G., Liu H., Zhang Y., Zhang R., Zhang S., Liu J., Hu Q., Hu F., Wang W., Ding Y. (2021) The elite alleles of *OsSPL4* regulate grain size and increase grain yield in rice. *Rice*. **14**, 90. <https://doi.org/10.1186/s12284-021-00531-7>
8. Duy P.N., Lan D.T., Thu H.P., Thanh H.N., Pham N.P., Auguy F., Thi B.T.H., Manh T.B., Cunnac S., Pham X.H. (2021) Improved bacterial leaf blight disease resistance in the major elite Vietnamese rice cultivar TBR225 via editing of the *OsSWEET14* promoter. *PLoS One*. **16**(9), e0255470. <https://doi.org/10.1371/journal.pone.0255470>
9. Dong S., Dong X., Han X., Zhang F., Zhu Y., Xin X., Wang Y., Hu Y., Yuan D., Wang J., Huang Z., Niu F., Hu Z., Yan P., Cao L., He H., Fu J., Xin Y., Tan Y., Mao B., Zhao B., Yang J., Yuan L., Luo X. (2021) *OsPDCD5* negatively regulates plant architecture and grain yield in rice. *Proc. Natl. Acad. Sci. USA*. **118**(29), e2018799118. <https://doi.org/10.1073/pnas.2018799118>
10. Nurhayati, Ardie S.W., Santoso T.J., Sudarsono. (2021) CRISPR/Cas9-mediated genome editing in rice cv. IPB3S results in a semi-dwarf phenotypic mutant. *Biodiversitas J. Biol. Diversity*. **22**(9), 3792–3800. <https://doi.org/10.13057/biodiv/d220924>
11. Tao H., Shi X., He F., Wang D., Xiao N., Fang H., Wang R., Zhang F., Wang M., Li A., Liu X., Wang G.L., Ning Y. (2021) Engineering broad-spectrum disease-resistant rice by editing multiple susceptibility genes. *J. Integr. Plant Biol.* **63**(9), 1639–1648. <https://doi.org/10.1111/jipb.13145>
12. Nawaz G., Usman B., Peng H., Zhao N., Yuan R., Liu Y., Li R. (2020) Knockout of *Pi21* by CRISPR/Cas9 and iTRAQ-based proteomic analysis of mutants revealed new insights into *M. oryzae* resistance in elite rice line. *Genes*. **11**, 735. <https://doi.org/10.3390/genes11070735>
13. Zheng S., Ye C., Lu J., Liufu J., Lin L., Dong Z., Li J., Zhuang C. (2021) Improving the rice photosynthetic efficiency and yield by editing *OsHXX1* via CRISPR/Cas9 system. *Int. J. Mol. Sci.* **22**, 9554. <https://doi.org/10.3390/ijms22179554>
14. Zeng Y., Wen J., Zhao W., Wang Q., Huang W. (2020) Rational improvement of rice yield and cold tolerance by editing the three genes *OsPIN5b*, *GS3*, and *OsMYB30* with the CRISPR-Cas9 system. *Front. Plant Sci.* **10**, 1663. <https://doi.org/10.3389/fpls.2019.01663>
15. Wu M., Liu H., Lin Y., Chen J., Fu Y., Luo J., Zhang Z., Liang K., Chen S., Wang F. (2020) In-frame and frame-shift editing of the *Ehd1* gene to develop japonica rice with prolonged basic vegetative growth periods. *Front. Plant Sci.* **11**, 307. <https://doi.org/10.3389/fpls.2020.00307>
16. Honma Y., Adhikari P.B., Kuwata K., Kagenishi T., Yokawa K., Notaguchi M., Kurotani K., Toda E., Bessho-Uehara K., Liu X., Zhu S., Wu X., Kasahara R.D. (2020) High-quality sugar production by osgcs1 rice. *Commun. Biol.* **3**, 617. <https://doi.org/10.1038/s42003-020-01329-x>
17. Usman B., Nawaz G., Zhao N., Liu Y., Li R. (2020) Generation of high yielding and fragrant rice (*Oryza sativa* L.) lines by CRISPR/Cas9 targeted mutagenesis of three homoeologs of cytochrome P450 gene family and *OsBADH2* and transcriptome and proteome profiling of revealed changes triggered by mutations. *Plants*. **9**, 788. <https://doi.org/10.3390/plants9060788>
18. Wang G., Wang C., Lu G., Wang W., Mao G., Habben J.E., Song C., Wang J., Chen J., Gao Y., Liu J., Greene T.W. (2020) Knockouts of a late flowering gene via CRISPR-Cas9 confer early maturity in rice at multiple field locations. *Plant. Mol. Biol.* **104**, 137–150. <https://doi.org/10.1007/s11103-020-01031-w>
19. Zhang A., Liu Y., Wang F., Li T., Chen Z., Kong D., Bi J., Zhang F., Luo X., Wang J., Tang J., Yu X., Liu G., Luo L. (2019) Enhanced rice salinity tolerance via CRISPR/Cas9-targeted mutagenesis of the *OsRR22* gene. *Mol. Breeding*. **39**(3), 47. <https://doi.org/10.1007/s11032-019-0954-y>
20. Kim Y.A., Moon H., Park C.J. (2019) CRISPR/Cas9-targeted mutagenesis of *Os8N3* in rice to confer resistance to *Xanthomonas oryzae* pv. *oryzae*. *Rice*. **12**, 67. <https://doi.org/10.1186/s12284-019-0325-7>
21. Hu X., Cui Y., Dong G., Feng A., Wang D., Zhao C., Zhang Y., Hu J., Zeng D., Guo L., Qian Q. (2019) Using CRISPR-Cas9 to generate semi-dwarf rice lines in elite landraces. *Sci. Rep.* **9**, 19096. <https://doi.org/10.1038/s41598-019-55757-9>
22. Yang Q., Ding J., Feng X., Zhong X., Lan J., Tang H., Harwood W., Li Z., Guzmán C., Xu Q., Zhang Y., Jiang Y., Qi P., Deng M., Ma J., Wang J., Chen G., Lan X., Wei Y., Zheng Y., Jiang Q. (2022) Editing of the starch synthase IIa gene led to transcriptomic and metabolomic changes and high amylose starch in barley. *Carbohydrate Polymers*. **285**, 119238. <https://doi.org/10.1016/j.carbpol.2022.119238>
23. Galli M., Martiny E., Imani J., Kumar N., Koch A., Steinbrenner J., Kogel K.H. (2022) CRISPR/Sp-Cas9-mediated double knockout of barley microRNA *MORC1* and *MORC6a* reveals their strong involvement in plant immunity, transcriptional gene silencing and plant growth. *Plant Biotechnol. J.* **20**(1), 89–102. <https://doi.org/10.1111/pbi.13697>
24. Lee J.H., Won H.J., Tran P.H.N., Lee S.-Mi, Je H.-Y.K., Jung H. (2021) Improving lignocellulosic biofuel production by CRISPR/Cas9-mediated lignin modifica-

- tion in barley. *GCB Bioenergy*. **13**, 742–752.
<https://doi.org/10.1111/gcbb.12808>
25. Gerasimova S.V., Hertig C., Korotkova A.M., Kolosovskaya E.V., Otto I., Hiekel S., Kochetov A.V., Khlestkina E.K., Kumlehn J. (2020) Conversion of hulled into naked barley by Cas endonuclease-mediated knockout of the *NUD* gene. *BMC Plant Biol.* **20**, 255.
<https://doi.org/10.1186/s12870-020-02454-9>
 26. Герасимова С.В., Короткова А.М., Хертинг К., Хикель С., Хоффи Р., Будхагатапалли Н., Отто И., Хензель Г., Шумный В.К., Кочетов А.В., Кумлен Й., Хлесткина Е.К. (2018) Применение РНК-направленной нуклеазы Cas9 для сайт-специфической модификации генома в протопластах сибирского сорта ячменя с высокой способностью к регенерации. *Вавил. журн. генет. селекции*. **22**(8), 1033–1039.
<https://doi.org/10.18699/VJ18.447>
 27. Holubova K., Hensel G., Vojta P., Tarkowski P., Bergounoux V., Galuszka P. (2018) Modification of barley plant productivity through regulation of cytokinin content by reverse-genetics approaches. *Front. Plant. Sci.* **9**, 1676.
<https://doi.org/10.3389/fpls.2018.01676>
 28. Gasparis S., Przyborowski M., Kała M., Nadolska-Orczyk A. (2019) Knockout of the *HvCKX1* or *HvCKX3* gene in barley (*Hordeum vulgare* L.) by RNA-guided Cas9 nuclease affects the regulation of cytokinin metabolism and root morphology. *Cells*. **8**(8), 782.
<https://doi.org/10.3390/cells8080782>
 29. Ibrahim S., Saleem B., Rehman N., Zafar A.S., Naem M.K., Khan M.R. (2021) CRISPR/Cas9 mediated disruption of inositol pentakisphosphate 2-kinase 1 (TaIPK1) reduces phytic acid and improves iron and zinc accumulation in wheat grains. *J. Adv. Res.* 1–9.
<https://doi.org/10.1016/j.jare.2021.07.006>
 30. Guo M., Wang Q., Zong Y., Nian J., Li H., Li J., Wang T., Gao C., Zuo J. (2021) Genetic manipulations of *TaARE1* boost nitrogen utilization and grain yield in wheat. *J. Genet. Genomics*. **48**(10), 950–953.
<https://doi.org/10.1016/j.jgg.2021.07.003>
 31. Zhang S., Zhang R., Gao J., Gu T., Song G., Li W., Li D., Li Y., Li G. (2019) Highly efficient and heritable targeted mutagenesis in wheat via the *Agrobacterium tumefaciens*-mediated CRISPR/Cas9 system. *Int. J. Mol. Sci.* **20**(17), 4257.
<https://doi.org/10.3390/ijms20174257>
 32. Raffan S., Sparks C., Huttly A., Hyde L., Martignago D., Mead A., Hanley S.J., Wilkinson P.A., Barker G., Edwards K.J., Curtis T.Y., Usher S., Kosik O., Halford N.G. (2021) Wheat with greatly reduced accumulation of free asparagine in the grain, produced by CRISPR/Cas9 editing of asparagine synthetase gene *TaASN2*. *Plant Biotechnol. J.* **19**(8), 1602–1613.
<https://doi.org/10.1111/pbi.13573>
 33. Hahn F., Loures S.L., Sparks C.A., Kanyuka K., Nekrasov V. (2021) Efficient CRISPR/Cas-mediated targeted mutagenesis in spring and winter wheat varieties. *Plants*. **10**, 1481.
<https://doi.org/10.3390/plants10071481>
 34. Li J., Jiao G., Sun Y., Chen J., Zhong Y., Yan L., Jiang D., Ma., Xia, L. (2021) Modification of starch composition, structure and properties through editing of *TaSBEIIa* in both winter and spring wheat varieties by CRISPR/Cas9. *Plant Biotechnol. J.* **19**, 937–951.
<https://doi.org/10.1111/pbi.13519>
 35. Tang H., Liu H., Zhou Y., Liu H., Du L., Wang K., Ye X. (2020) Fertility recovery of wheat male sterility controlled by Ms2 using CRISPR/Cas9. *Plant Biotechnol. J.* **19**(2), 224–226.
<https://doi.org/10.1111/pbi.13482>
 36. Budhagatapalli N., Halbach T., Hiekel S., Büchner H., Müller A.E., Kumlehn J. (2020) Site-directed mutagenesis in bread and durum wheat via pollination by Cas9/guide RNA-transgenic maize used as haploidy inducer. *Plant Biotechnol. J.* **18**(12), 2376–2378.
 37. Wang W., Pan Q., Tian B., He F., Chen Y., Bai G., Akhunov E. (2019) Gene editing of the wheat homologs of *TONNEAU1*-recruiting motif encoding gene affects grain shape and weight in wheat. *Plant J.* **100**, 251–264.
<https://doi.org/10.1111/tpj.14440>
 38. Abe F., Haque E., Hisano H., Tanaka T., Kamiya Y., Mikami M., Sato K. (2019) Genome-edited triple-recessive mutation alters seed dormancy in wheat. *Cell Rep.* **28**(5), 1362–1369.e4.
<https://doi.org/10.1016/j.celrep.2019.06.090>
 39. Zhang Z., Hua L., Gupta A., Tricoli D., Edwards K.J., Yang B., Li W. (2019) Development of an *Agrobacterium*-delivered CRISPR/Cas9 system for wheat genome editing. *Plant Biotechnol. J.* **17**, 1623–1635.
<https://doi.org/10.1111/pbi.13088>
 40. Brauer E.K., Balcerzak M., Rocheleau H., Leung W., Scherthaner J., Subramaniam R., Ouellet T. (2020) Genome editing of a deoxynivalenol-induced transcription factor confers resistance to fusarium graminearum in wheat. *Mol. Plant Microbe Interact.* **33**(3), 553–560.
<https://doi.org/10.1094/MPMI-11-19-0332-R>
 41. Camerlengo F., Frittelli A., Sparks C., Doherty A., Martignago D., Larre C., Lupi R., Sestili F., Masci S. (2020) CRISPR-Cas9 multiplex editing of the α -amylase/trypsin inhibitor genes to reduce allergen proteins in durum wheat. *Front. Sustain. Food Syst.* **4**, 104.
<https://doi.org/10.3389/fsufs.2020.00104>
 42. Guan H., Chen X., Wang K., Liu X., Zhang D., Li Y., Song Y., Shi Y., Wang T., Li C., Li Y. (2022) Genetic variation in *ZmPAT7* contributes to tassel branch number in maize. *Int. J. Mol. Sci.* **23**(5), 2586.
<https://doi.org/10.3390/ijms23052586>
 43. Li Y., Lin Z., Yue Y., Zhao H., Fei X., Lizhu E., Liu C., Chen S., Lai J., Song W. (2021) Loss-of-function alleles of *ZmPLD3* cause haploid induction in maize. *Nat. Plants*. **7**, 1579–1588.
<https://doi.org/10.1038/s41477-021-01037-2>
 44. Wang Y., Liu X., Zheng X., Wang W., Yin X., Liu H., Ma C., Niu X., Zhu J.K., Wang F. (2021) Creation of aromatic maize by CRISPR/Cas. *J. Integr. Plant. Biol.* **63**(9), 1664–1670.
<https://doi.org/10.1111/jipb.13105>
 45. Gao L., Yang G., Li Y., Sun Y., Xu R., Chen Y., Wang Z., Xing J., Zhang Y. (2021) A Kelch-repeat superfamily

- gene, *ZmNL4*, controls leaf width in maize (*Zea mays* L.). *Plant J.* **107**(3), 817–830.
<https://doi.org/10.1111/tpj.15348>
46. Liu L., Gallagher J., Arevalo E.D., Chen R., Skopelitis T., Wu Q., Bartlett M., Jackson D. (2021) Enhancing grain-yield-related traits by CRISPR-Cas9 promoter editing of maize *CLE* genes. *Nat. Plants.* **7**(3), 287–294.
<https://doi.org/10.1038/s41477-021-00858-5>
 47. Li Q., Wu G., Zhao Y., Wang B., Zhao B., Kong D., Wei H., Chen C., Wang H. (2020) CRISPR/Cas9-mediated knockout and overexpression studies reveal a role of maize phytochrome C in regulating flowering time and plant height. *Plant Biotechnol. J.* **18**(12), 2520–2532.
<https://doi.org/10.1111/pbi.13429>
 48. Jiang Y.Y., Chai Y.P., Lu M.H., Han X.L., Lin Q., Zhang Y., Zhang Q., Zhou Y., Wang X.C., Gao C., Chen Q.J. (2020) Prime editing efficiently generates W542L and S621I double mutations in two *ALS* genes in maize. *Genome Biol.* **21**(1), 257.
<https://doi.org/10.1186/s13059-020-02170-5>
 49. Li Y., Zhu J., Wu H., Liu C., Huang C., Lan J., Zhao Y., Xie C. (2020) Precise base editing of non-allelic acetolactate synthase genes confers sulfonylurea herbicide resistance in maize. *Crop J.* **8**, 449–456.
<https://doi.org/10.1016/j.cj.2019.10.001>
 50. Qi X., Wu H., Jiang H., Zhu J., Huang C., Zhang X., Liu C., Cheng B. (2020) Conversion of a normal maize hybrid into a *waxy* version using *in vivo* CRISPR/Cas9 targeted mutation activity. *Crop J.* **8**, 440–448.
<https://doi.org/10.1016/j.cj.2020.01.006>
 51. Zhao X., Jayarathna S., Tureson H., Fält A., Nestor G., González M.N., Olsson N., Beganovic M., Hofvander P., Andersson R., Andersson M. (2021) Amylose starch with no detectable branching developed through DNA-free CRISPR-Cas9 mediated mutagenesis of two starch branching enzymes in potato. *Sci. Rep.* **11**, 4311.
<https://doi.org/10.1038/s41598-021-83462-z>
 52. Tuncel A., Corbin K.R., Ahn-Jarvis J., Harris S., Hawkins E., Smedley M.A., Harwood W., Warren F.J., Patron N.J., Smith A.M. (2019) Cas9-mediated mutagenesis of potato starch-branching enzymes generates a range of tuber starch phenotypes. *Plant Biotechnol. J.* **17**, 2259–2271.
<https://doi.org/10.1111/pbi.13137>
 53. Takeuchi A., Ohnuma M., Teramura H., Asano K., Noda T., Kusano H., Tamura K., Shimada H. (2021) Creation of a potato mutant lacking the starch branching enzyme gene *StSBE3* that was generated by genome editing using the CRISPR/dMac3-Cas9 system. *Plant Biotechnol.* (Tokyo). **38**(3), 345–353.
<https://doi.org/10.5511/plantbiotechnology.21.0727a>
 54. Zheng Z., Ye G., Zhou Y., Pu X., Su W., Wang J. (2021) Editing sterol side chain reductase 2 gene (*StSSR2*) via CRISPR/Cas9 reduces the total steroidal glycoalkaloids in potato. *All Life.* **14**(1), 401–413.
<https://doi.org/10.1080/26895293.2021.1925358>
 55. Gonzalez M.N., Massa G.A., Andersson M., Tureson H., Olsson N., Fält A.-S., Storani L., Oneto D.C.A., Hofvander P., Feingold S.E. (2020) Reduced enzymatic browning in potato tubers by specific editing of a polyphenol oxidase gene via ribonucleoprotein complexes delivery of the CRISPR/Cas9 system. *Front. Plant Sci.* **10**, 1649.
<https://doi.org/10.3389/fpls.2019.01649>
 56. Kieu N.P., Lenman M., Wang E.S., Petersen B.L., Andreasson E. (2021) Mutations introduced in susceptibility genes through CRISPR/Cas9 genome editing confer increased late blight resistance in potatoes. *Sci. Rep.* **11**(1), 4487.
<https://doi.org/10.1038/s41598-021-83972-w>
 57. Osakabe Y., Liang Z., Ren C., Nishitani C., Osakabe K., Wada M., Komori S., Malnoy M., Velasco R., Poli M., Jung M.-H., Koo O.-J., Viola R., Kanchiswamy C.N. (2018) CRISPR–Cas9-mediated genome editing in apple and grapevine. *Nat. Protoc.* **13**(12), 2844–2863.
<https://doi.org/10.1038/s41596-018-0067-9>
 58. Scintilla S., Salvagnin U., Giacomelli L., Zeilmaker T., Malnoy M.A., van der Voort J.R., Moser C. (2021) Regeneration of plants from DNA-free edited grapevine protoplasts.
<https://doi.org/10.1101/2021.07.16.452503>
 59. Olivares F., Loyola R., Olmedo B., Miccono M.D.L.Á., Aguirre C., Vergara R., Riquelme D., Madrid G., Plantat P., Mora R., Espinoza D., Prieto H. (2021) CRISPR/Cas9 targeted editing of genes associated with fungal susceptibility in *Vitis vinifera* L. cv. Thompson Seedless using geminivirus-derived replicons. *Front. Plant Sci.* **12**, 791030.
<https://doi.org/10.3389/fpls.2021.791030>
 60. Wan D.Y., Guo Y., Cheng Y., Hu Y., Xiao S., Wang Y., Wen Y.Q. (2020) CRISPR/Cas9-mediated mutagenesis of *VvMLO3* results in enhanced resistance to powdery mildew in grapevine (*Vitis vinifera*). *Horticulture Res.* **7**, 116.
<https://doi.org/10.1038/s41438-020-0339-8>
 61. Yang L., Guo Y., Hu Y., Wen Y. (2020) CRISPR/Cas9-mediated mutagenesis of *VviEDR2* results in enhanced resistance to powdery mildew in grapevine (*Vitis vinifera*). *Acta Horticulturae Sinica.* **47**(4), 623–634.
<https://doi.org/10.16420/j.issn.0513-353x.2019-0660>
 62. Li M., Jiao Y., Wang Y., Zhang N., Wang B., Liu R., Yin X., Xu Y., Liu G. (2020) CRISPR/Cas9-mediated *VvPR4b* editing decreases downy mildew resistance in grapevine (*Vitis vinifera* L.). *Hortic. Res.* **7**(1), 149.
<https://doi.org/10.1038/s41438-020-00371-4>
 63. Sunitha S., Rock C.D. (2020) CRISPR/Cas9-mediated targeted mutagenesis of TAS4 and MYBA7 loci in grapevine rootstock 101-14. *Transgenic Res.* **29**(3), 355–367.
<https://doi.org/10.1007/s11248-020-00196-w>
 64. Ren C., Guo Y., Kong J., Lecourieux F., Dai Z., Li S., Liang Z. (2020) Knockout of *VvCCD8* gene in grapevine affects shoot branching. *BMC Plant Biol.* **20**(1), 47.
<https://doi.org/10.1186/s12870-020-2263-3>
 65. Tripathi J.N., Ntui, V.O., Shah T., Tripathi L. (2021) CRISPR/Cas9-mediated editing of *DMR6* orthologue in banana (*Musa* spp.) confers enhanced resistance to bacterial disease. *Plant Biotechnol. J.* **1**(7), 1291–1293.
<https://doi.org/10.1111/pbi.13614>
 66. Hu C., Sheng O., Deng G., He W., Dong T., Yang Q., Dou T., Li C., Gao H., Liu S., Yi G., Bi F. (2021)

- CRISPR/Cas9-mediated genome editing of *MaACO1* (aminocyclopropane-1-carboxylate oxidase 1) promotes the shelf life of banana fruit. *Plant Biotechnol. J.* **19**(4), 654–656.
<https://doi.org/10.1111/pbi.13534>
67. Varkonyi-Gasic E., Wang T., Cooney J., Jeon S., Voogd C., Douglas M.J., Pilkington S.M., Akagi T., Allan A.C. (2021) *Shy Girl*, a kiwifruit suppressor of feminization, restricts gynoceium development via regulation of cytokinin metabolism and signalling. *New Phytol.* **230**, 1461–1475.
<https://doi.org/10.1111/nph.17234>
 68. Pompili V., Costa L. D., Piazza S., Pindo M., Malnoy M. (2019) Reduced fire blight susceptibility in apple cultivars using a high-efficiency CRISPR/Cas9-FLP/FRT-based gene editing system. *Plant Biotechnol. J.* **18**(3), 845–858.
<https://doi.org/10.1111/pbi.13253>
 69. Omori M., Yamane H., Li K., Matsuzaki R., Ebihara S., Li T., Tao R. (2020) Expressional analysis of *FT* and *CEN* genes in a continuously flowering highbush blueberry “*Blue Muffin*”. *Acta Hort.* **1280**, 197–201.
<https://doi.org/10.17660/ActaHortic.2020.1280.27>
 70. Jia H., Wang Y., Su H., Huan X., Wang N. (2022) *LbCas12a-D156R* efficiently edits *LOB1* effector binding elements to generate canker-resistant citrus plants. *Cells.* **11**(3), 315.
<https://doi.org/10.3390/cells11030315>
 71. Zhou J., Wang G., Liu Z. (2018) Efficient genome editing of wild strawberry genes, vector development and validation. *Plant Biotechnol. J.* **16**(11), 1868–1877.
<https://doi.org/10.1111/pbi.12922>
 72. Feng J., Dai C., Luo H., Han Y., Liu Z., Kang C. (2019) Reporter gene expression reveals precise auxin synthesis sites during fruit and root development in wild strawberry. *J. Exp. Botany.* **70**(2), 563–574.
<https://doi.org/10.1093/jxb/ery384>
 73. Bottero E., Gomez C., Stritzler M., Tajima H., Frare R., Pascuan C., Blumwald E., Ayub N., Soto G. (2022) Generation of a multi-herbicide-tolerant alfalfa by using base editing. *Plant Cell Rep.* **41**, 493–495.
<https://doi.org/10.1007/s00299-021-02827-w>
 74. Zhang Z., Wang J., Kuang H., Hou Z., Gong P., Bai M., Zhou S., Yao X., Song S., Yan L., Guan Y. (2022) Elimination of an unfavorable allele conferring pod shattering in an elite soybean cultivar by CRISPR/Cas9. *ABIOTECH.*
<https://doi.org/10.1007/s42994-022-00071-8>
 75. Chen X., Yang S., Zhang Y., Zhu X., Yang X., Zhang C., Li H., Feng X. (2021) Generation of male-sterile soybean lines with the CRISPR/Cas9 system. *Crop J.* **9**(6), 1270–1277.
<https://doi.org/10.1016/j.cj.2021.05.003>
 76. Wang T., Xun H., Wang W., Ding X., Tian H., Hussain S., Dong Q., Li Y., Cheng Y., Wang C., Lin R., Li G., Qian X., Pang J., Feng X., Dong Y., Liu B., Wang S. (2021) Mutation of *GmAITR* genes by CRISPR/Cas9 genome editing results in enhanced salinity stress tolerance in soybean. *Front. Plant Sci.* **12**, 779598.
<https://doi.org/10.3389/fpls.2021.779598>
 77. Cai Z., Xian P., Cheng Y., Ma Q., Lian T., Nian H., Ge L. (2021) CRISPR/Cas9-mediated gene editing of *GmJAGGED1* increased yield in the low-latitude soybean variety Huachun 6. *Plant Biotechnol. J.* **19**(10), 1898–1900.
<https://doi.org/10.1111/pbi.13673>
 78. Li Z., Cheng Q., Gan Z., Hou Z., Zhang Y., Li Y., Li H., Nan H., Yang C., Chen L., Lu S., Shi W., Chen L., Wang Y., Fang C., Kong L., Su T., Li S., Kou K., Wang L., Kong F., Liu B., Dong L. (2021) Multiplex CRISPR/Cas9-mediated knockout of soybean *LNK2* advances flowering time. *Crop J.* **9**(4), 767–776.
<https://doi.org/10.1016/j.cj.2020.09.005>
 79. Ma J., Sun S., Whelan J., Shou H. (2021) CRISPR/Cas9-mediated knockout of *GmFATB1* significantly reduced the amount of saturated fatty acids in soybean seeds. *Int. J. Mol. Sci.* **22**, 3877.
<https://doi.org/10.3390/ijms22083877>
 80. Nguyen C.X., Paddock K.J., Zhang Z., Stacey M.G. (2021) *GmKIX8-1* regulates organ size in soybean and is the causative gene for the major seed weight QTL qSw17-1. *New Phytol.* **229**, 920–934.
<https://doi.org/10.1111/nph.16928>
 81. Adachi K., Hirose A., Kanazashi Y., Hibara M., Hirata T., Mikami M., Endo M., Hirose S., Maruyama N., Ishimoto M., Abe J., Yamada T. (2021) Site-directed mutagenesis by biolistic transformation efficiently generates inheritable mutations in a targeted locus in soybean somatic embryos and transgene-free descendants in the T1 generation. *Transgenic Res.* **30**, 77–89.
<https://doi.org/10.1007/s11248-020-00229-4>
 82. Le H., Nguyen N.H., Ta D.T., Le T.N.T., Bui T.P., Le N.T., Nguyen C.X., Rolletschek H., Stacey G., Stacey M.G., Pham N.B., Do P.T., Chu H.H. (2020) CRISPR/Cas9-mediated knockout of galactinol synthase-encoding genes reduces raffinose family oligosaccharide levels in soybean seeds. *Front. Plant Sci.* **11**, 612942.
<https://doi.org/10.3389/fpls.2020.612942>
 83. Sugano S., Hirose A., Kanazashi Y., Adachi K., Hibara M., Itoh T., Mikami M., Endo M., Hirose S., Maruyama N., Abe J., Yamada T. (2020) Simultaneous induction of mutant alleles of two allergenic genes in soybean by using site-directed mutagenesis. *BMC Plant Biol.* **20**, 513.
<https://doi.org/10.1186/s12870-020-02708-6>
 84. Chen L., Nan H., Kong L., Yue L., Yang H., Zhao Q., Fang C., Li H., Cheng Q., Lu S., Kong F., Liu B., Dong L. (2020) Soybean AP1 homologs control flowering time and plant height. *J. Integr. Plant Biol.* **62**, 1868–1879.
<https://doi.org/10.1111/jipb.12988>
 85. Wang L., Sun S., Wu T., Liu L., Sun X., Cai Y., Li J., Jia H., Yuan S., Chen L., Jiang B., Wu C., Hou W., Han T. (2020) Natural variation and CRISPR/Cas9-mediated mutation in *GmPRR37* affect photoperiodic flowering and contribute to regional adaptation of soybean. *Plant Biotechnol. J.* **18**, 1869–1881.
<https://doi.org/10.1111/pbi.13346>
 86. Zhang P., Du H., Wang J., Pu Y., Yang C., Yan R., Yang H., Cheng H., Yu D. (2020) Multiplex CRISPR/Cas9-mediated metabolic engineering in-

- creases soya bean isoflavone content and resistance to soya bean mosaic virus. *Plant Biotechnol. J.* **18**(6), 1384–1395.
<https://doi.org/10.1111/pbi.13302>
87. Hou Z.H., Wu Y., Cheng Q., Gan Z.R., Liu B.H. (2019) Creation of high oleic acid soybean mutation plants by CRISPR/Cas9. *Acta Agronomica Sinica* (China). **45**(6), 839–847.
 88. Do P.T., Nguyen C.X., Bui H.T., Tran L.T.N., Stacey G., Gillman J.D., Zhang Z.J., Stacey M.J. (2019) Demonstration of highly efficient dual gRNA CRISPR/Cas9 editing of the homeologous *GmFAD2-1A* and *GmFAD2-1B* genes to yield a high oleic, low linoleic and α -linolenic acid phenotype in soybean. *BMC Plant Biol.* **19**, 311.
<https://doi.org/10.1186/s12870-019-1906-8>
 89. Wu N., Lu Q., Wang P., Zhang Q., Zhang J., Qu J., Wang N. (2020) Construction and analysis of *GmFAD2-1A* and *GmFAD2-2A* soybean fatty acid desaturase mutants based on CRISPR/Cas9 technology. *Int. J. Mol. Sci.* **21**, 1104.
<https://doi.org/10.3390/ijms21031104>
 90. Cai Y., Wang L., Chen L., WT., Liu L., Sun S., Wu C., Yao W., Jiang B., Yuan S., Han T., Hou W. (2020) Mutagenesis of *GmFT2a* and *GmFT5a* mediated by CRISPR/Cas9 contributes for expanding the regional adaptability of soybean. *Plant Biotechnol. J.* **18**, 298–309.
<https://doi.org/10.1111/pbi.13199>
 91. Han J., Guo B., Guo Y., Zhang B., Wang X., Qiu L.-J. (2019) Creation of early flowering germplasm of soybean by CRISPR/Cas9 technology. *Front. Plant Sci.* **10**, 1446.
<https://doi.org/10.3389/fpls.2019.01446>
 92. Bao A., Chen H., Chen L., Chen S., Hao Q., Guo W., Qiu D., Shan Z., Yang Z., Yuan S., Zhang C., Zhang X., Liu B., Kong F., Li X., Zhou X., Trna L.-S.P., Cao D. (2019) CRISPR/Cas9-mediated targeted mutagenesis of *GmSPL9* genes alters plant architecture in soybean. *BMC Plant Biol.* **19**, 131.
<https://doi.org/10.1186/s12870-019-1746-6>
 93. Badhan S., Ball A.S., Mantri N. (2021) First report of CRISPR/Cas9 mediated DNA-free editing of *4CL* and *RVE7* genes in chickpea protoplasts. *Int. J. Mol. Sci.* **22**, 396.
<https://doi.org/10.3390/ijms22010396>
 94. Dai X., Han H., Huang W., Zhao L., Song M., Cao X., Liu C., Niu X., Lang Z., Ma C., Xie H. (2022) Generating novel male sterile tomatoes by editing respiratory burst oxidase homolog genes. *Front. Plant Sci.* **12**, 817101.
<https://doi.org/10.3389/fpls.2021.817101>
 95. Kawaguchi K., Takei-Hoshi R., Yoshikawa I., Nishida K., Kobayashi M., Kusano M., Lu Y., Ariizumi T., Ezura H., Otagaki S., Matsumoto S., Shiratake K. (2021) Functional disruption of cell wall invertase inhibitor by genome editing increases sugar content of tomato fruit without decrease fruit weight. *Sci. Rep.* **11**(1), 21534.
 96. Wang B., Li N., Huang S., Hu J., Wang Q., Tang Y., Yang T., Asmutola P., Wang J., Yu Q. (2021) Enhanced soluble sugar content in tomato fruit using CRISPR/Cas9-mediated *SlINVINH1* and *SlVPE5* gene editing. *Peer. J.* **9**, e12478.
<https://doi.org/10.7717/peerj.12478>
 97. Bari V.K., Nassar J.A., Aly R. (2021) CRISPR/Cas9 mediated mutagenesis of *MORE AXILLARY GROWTH 1* in tomato confers resistance to root parasitic weed *Phelipanche aegyptiaca*. *Sci. Rep.* **11**, 3905.
<https://doi.org/10.1038/s41598-021-82897-8>
 98. Hanika K., Schipper D., Chinnappa S., Oortwijn M., Schouten H.J., Thomma B.P.H.J., Bai Y. (2021) Impairment of tomato *WAT1* enhances resistance to vascular wilt fungi despite severe growth defects. *Front. Plant Sci.* **12**, 721674.
<https://doi.org/10.3389/fpls.2021.721674>
 99. Liu H., Lihong Liu, Liang D., Zhang M., Jia C., Qi M., Liu Y., Shao Z., Meng F., Hu S., Yin Y., Li C., Wang Q. (2021) SIBES1 promotes tomato fruit softening through transcriptional inhibition of *PMEU1*. *Science*. **24**(8), 102926.
<https://doi.org/10.1016/j.isci.2021.102926>
 100. Thomazella D.P.D.T., Seong K., Mackelprang R., Dahlbeck D., Geng Y., Gill U.S., Qi T., Pham J., Giuseppe P., Lee C.Y., Ortega A., Cho M., Hutton S.F., Staskawicz B. (2021) Loss of function of a DMR6 ortholog in tomato confers broad-spectrum disease resistance. *Proc. Natl Acad. Sci. USA*. **118**(27), e2026152118.
<https://doi.org/10.1073/pnas.2026152118>
 101. Tran M.T., Doan D.T.H., Kim J., Song Y.J., Sung Y.W., Das S., Kim E.J., Son G.H., Kim S.H., Van Vu T., Kim J.Y. (2021) CRISPR/Cas9-based precise excision of *SlHyPRP1* domain(s) to obtain salt stress-tolerant tomato. *Plant Cell Rep.* **40**(6), 999–1011.
<https://doi.org/10.1007/s00299-020-02622-z>
 102. Liu J., Wang S., Wang H., Luo B., Cai Y., Li X., Zhang Y., Wang X. (2021) Rapid generation of tomato male-sterile lines with a marker use for hybrid seed production by CRISPR/Cas9 system. *Mol. Breed.* **41**(3), 25.
<https://doi.org/10.1007/s11032-021-01215-2>
 103. Atarashi H., Jayasinghe W.H., Kwon J., Kim H., Taninaka Y., Igarashi M., Ito K., Yamada T., Masuta C., Nakahara K.S. (2020) Artificially edited alleles of the eukaryotic translation initiation factor *4E1* gene differentially reduce susceptibility to cucumber mosaic virus and potato virus Y in tomato. *Front. Microbiol.* **11**, 564310.
<https://doi.org/10.3389/fmicb.2020.564310>
 104. Yoon Y.J., Venkatesh J., Lee J.H., Kim J., Lee H.E., Kim D.S., Kang B.C. (2020) Genome editing of *EIF4E1* in tomato confers resistance to pepper mottle virus. *Front. Plant Sci.* **11**, 1–11.
<https://doi.org/10.3389/fpls.2020.01098/full>
 105. Kuroiwa K., Thenault C., Nogue F., Perrot L., Maziera M., Galloisa J.L. (2022) CRISPR-based knockout of *EIF4E2* in a cherry tomato background successfully recapitulates resistance to pepper veinal mottle virus. *Plant Sci.* **316**, 111160.
<https://doi.org/10.1016/j.plantsci.2021.111160>
 106. Jung Y.J., Kim D.H., Lee H.J., Nam K.H., Bae S., Nou I.S., Cho Y.-G., Kim M.K., Kang K.K. (2020) Knockout of *SIMS10* gene (Soly02g079810) encoding bHLH transcription factor using CRISPR/Cas9 system confers male sterility phenotype in tomato.

- Plants*. **9**(9), 1189.
<https://doi.org/10.3390/plants9091189>
107. Illouz-Eliaz N., Nissan I., Nir I., Ramon U., Shohat H., Weiss D. (2020) Mutations in the tomato gibberellin receptors suppress xylem proliferation and reduce water loss under water-deficit conditions. *J. Exp. Botany*. **71**(12), 3603–3612.
<https://doi.org/10.1093/jxb/eraa137>
 108. Santillán Martínez M.I., Bracuto V., Koseoglou E., Appiano M., Jacobsen E., Visser R.G.F., Wolters A.M.A., Bai Y. (2020) CRISPR/Cas9-targeted mutagenesis of the tomato susceptibility gene *PMR4* for resistance against powdery mildew. *BMC Plant Biol.* **20**, 284.
<https://doi.org/10.1186/s12870-020-02497-y>
 109. Faal G.P., Farsi M., Seifi A., Kakhki A.M. (2020) Virus-induced CRISPR-Cas9 system improved resistance against tomato yellow leaf curl virus. *Mol. Biol. Rep.* **47**(5), 3369–3376.
<https://doi.org/10.1007/s11033-020-05409-3>
 110. Bari V.K., Nassar J.A., Kheredin S.M., Gal-On A., Ron M., Britt A., Steele D., Yoder J., Aly R. (2019) CRISPR/Cas9-mediated mutagenesis of carotenoid cleavage dioxygenase 8 in tomato provides resistance against the parasitic weed *Phelipanche aegyptiaca*. *Sci. Rep.* **9**, 11438.
<https://doi.org/10.1038/s41598-019-47893-z>
 111. Ortigosa A., Gimenez-Ibanez S., Leonhardt N., Solano R. (2019) Design of a bacterial speck resistant tomato by CRISPR/Cas9-mediated editing of *SIJAZ2*. *Plant Biotechnol. J.* **17**(3), 665–673.
<https://doi.org/10.1111/pbi.13006>
 112. Li X., Wang Y., Chen S., Tian H., Fu D., Zhu B., Luo Y., Zhu H. (2018) Lycopene is enriched in tomato fruit by CRISPR/Cas9-mediated multiplex genome editing. *Front. Plant Sci.* **9**, 559.
<https://doi.org/10.3389/fpls.2018.00559>
 113. Ahmar S., Zhai Y., Huang H., Yu K., Hafeez M., Khan U., Shahid M., Samad R.A., Khan S.U., Amoo O., Fan C., Zhou Y. (2020) Development of mutants with varying flowering times by targeted editing of multiple *SVP* gene copies in *Brassica napus* L. *Crop J.* **10**(1), 67–74.
<https://doi.org/10.1016/j.cj.2021.03.023>
 114. Cao Y., Yan X., Ran S., Ralph J., Smith R.A., Chen X., Qu C., Li J., Liu L. (2022) Knockout of the lignin pathway gene *BnF5H* decreases the S/G lignin compositional ratio and improves *Sclerotinia sclerotiorum* resistance in *Brassica napus*. *Plant Cell Environ.* **45**(1), 248–261.
<https://doi.org/10.1111/pce.14208>
 115. Fan S., Zhang L., Tang M., Cai Y., Liu J., Liu H., Liu J., Terzaghi W., Wang H., Hua W., Zheng M. (2021) CRISPR/Cas9-targeted mutagenesis of the *BnaA03.BP* gene confers semi-dwarf and compact architecture to rapeseed (*Brassica napus* L.). *Plant Biotechnol. J.* **19**(12), 2383–2385.
<https://doi.org/10.1111/pbi.13703>
 116. Zhang X., Cheng J., Lin Y., Fu Y., Xie J., Li B., Bian X., Feng Y., Liang W., Tang Q., Zhang H., Liu X., Zhang Y., Liu C., Jiang D. (2021) Editing homologous copies of an essential gene affords crop resistance against two cosmopolitan necrotrophic pathogens. *Plant Biotechnol. J.* **19**(11), 2349–2361.
<https://doi.org/10.1111/pbi.13667>
 117. Wang Z., Wan L., Xin Q., Zhang X., Song Y., Wang P., Hong D., Fan Z., Yang G. (2021) Optimizing glyphosate tolerance in rapeseed by CRISPR/Cas9-based geminiviral donor DNA replicon system with Csy4-based single-guide RNA processing. *J. Exp. Bot.* **72**(13), 4796–4808.
<https://doi.org/10.1093/jxb/erab167>
 118. Zaman Q.U., Wen C., Yuqin S., Mengyu H., Desheng M., Jacqueline B., Baohong Z., Chao L., Qiong H. (2021) Characterization of SHATTERPROOF homologs and CRISPR-Cas9-mediated genome editing enhances pod-shattering resistance in *Brassica napus* L. *CRISPR J.* **4**(3), 360–370.
<https://doi.org/10.1089/crispr.2020.0129>
 119. Probsting M., Schenke D., Hossain R., Thureau C., Wighardt L., Schuster A., Zhou Z., Ye W., Rietz S., Leckband G., Cai D. (2020) Loss-of-function of CRT1a (calreticulin) reduces susceptibility to *Verticillium longisporum* in both *Arabidopsis thaliana* and oilseed rape (*Brassica napus*). *Plant Biotechnol. J.* **18**, 2328–2344.
<https://doi.org/10.1111/pbi.13394>
 120. Sashidhar N., Harloff H.J., Potgieter L., Jung C. (2020) Gene editing of three *BnITPK* genes in tetraploid oilseed rape leads to significant reduction of phytic acid in seeds. *Plant Biotechnol. J.* **18**(11), 2241–2250.
<https://doi.org/10.1111/pbi.13380>
 121. Karunarathna N.L., Wang H., Harloff H.-J., Jiang L., Jung C. (2020) Elevating seed oil content in a polyploid crop by induced mutations in *SEED FATTYACID REDUCER* genes. *Plant Biotechnol. J.* **18**(11), 2251–2266.
<https://doi.org/10.1111/pbi.13381>
 122. Wu J., Chen C., Xian G., Liu D., Lin L., Yin S., Sun Q., Fang Y., Zhang H., Wang Y. (2020) Engineering herbicide-resistant oilseed rape by CRISPR/Cas9-mediated cytosine base-editing. *Plant Biotechnol. J.* **18**(9), 1857–1859.
<https://doi.org/10.1111/pbi.13368>
 123. Xie T., Chen X., Guo T., Rong H., Chen Z., Sun Q., Batley J., Jiang J., Wang Y. (2020) Targeted knockout of *BnTT2* homologues for yellow-seeded *Brassica napus* with reduced flavonoids and improved fatty acid composition. *J. Agric. Food Chem.* **68**(20), 5676–5690.
<https://doi.org/10.1021/acs.jafc.0c01126>
 124. Wu J., Yan G., Duan Z., Wang Z., Kang C., Guo L., Liu K., Tu J., Shen J., Yi B., Fu T., Li X., Ma C., Dai C. (2020) Roles of the *Brassica napus* DELLA protein BnaA6.RGA, in modulating drought tolerance by interacting with the ABA signaling component BnaA10.ABF2. *Front. Plant Sci.* **11**, 577.
<https://doi.org/10.3389/fpls.2020.00577>
 125. Zhai Y., Yu K., Cai S., Hu L., Amoo O., Xu L., Yang Y., Ma B., Jiao Y., Zhang C., Khan M.H.U., Khan S.U. (2020) Targeted mutagenesis of *BnTT8* homologs controls yellow seed coat development for effective oil production in *Brassica napus* L. *Plant Biotechnol. J.* **18**(5), 1153–1168.
<https://doi.org/10.1111/pbi.13281>

126. Zheng M., Zhang L., Tang M., Liu J., Liu H., Yang H., Fan S., Terzaghi, W., Wang H., Hua W. (2020) Knock-out of two *Bna MAX 1* homologs by CRISPR/Cas9-targeted mutagenesis improves plant architecture and increases yield in rapeseed (*Brassica napus* L.). *Plant Biotechnol. J.* **18**(3), 644–654.
<https://doi.org/10.1111/pbi.13228>
127. Khan M.H.U., Hu L., Zhu M., Zhai Y., Khan S.U., Ahmar S., Amoo O., Zhang K., Fan C., Zhou Y. (2021) Targeted mutagenesis of *EOD3* gene in *Brassica napus* L. regulates seed production. *J. Cell. Physiol.* **236**(3), 1996–2007.
<https://doi.org/10.1002/jcp.29986>
128. Sun Q., Lin L., Liu D., Wu D., Fang Y., Wu J., Wang Y. (2018) CRISPR/Cas9-mediated multiplex genome editing of the *BnWRKY11* and *BnWRKY70* genes in *Brassica napus* L. *Int. J. Mol. Sci.* **19**(9), 2716.
<https://doi.org/10.3390/ijms19092716>
129. Kim Y.C., Ahn W.S., Cha A., Jie E. Y., Kim S. W., Hwang B.-H., Lee S. (2022) Development of glucoraphanin-rich broccoli (*Brassica oleracea* var. *italica*) by CRISPR/Cas9-mediated DNA-free *BolMYB28* editing. *Plant Biotechnol. Rep.* **16**, 123–132.
<https://doi.org/10.1007/s11816-021-00732-y>
130. Jeong S.Y., Ahn H., Ryu J., Oh Y., Sivanandhan G., Won K.-H., Park Y.D., Kim J.-S., Kim H., Lim Y.P., Kim S.-G. (2019) Generation of early-flowering chinese cabbage (*Brassica rapa* spp. *pekinensis*) through CRISPR/Cas9-mediated genome editing. *Plant Biotechnol. Rep.* **13**(5), 491–499.
<https://doi.org/10.1007/s11816-019-00566-9>
131. Neequaye M., Stavnstrup S., Harwood W., Lawrenson T., Hundleby P., Irwin J., Troncoso-Rey P., Saha S., Traka M.H., Mithen R., Østergaard L. (2021) CRISPR-Cas9-mediated gene editing of *MYB28* genes impair glucoraphanin accumulation of *Brassica oleracea* in the field. *CRISPR J.* **4**(3), 416–426.
<https://doi.org/10.1089/crispr.2021.0007>
132. Cao W., Dong X., Ji J., Yang L., Fang Z., Zhuang M., Zhang Y., Lv H., Wang Y., Sun P., Liu Y., Li Z., Han F. (2021) BoCER1 is essential for the synthesis of cuticular wax in cabbage (*Brassica oleracea* L. var. *capitata*). *Scientia Horticulturae.* **277**, 109801.
<https://doi.org/10.1016/j.scienta.2020.109801>
133. Mishra R., Mohanty J.N., Mahanty B., Joshi R.K. (2021) A single transcript CRISPR/Cas9 mediated mutagenesis of *CaERF28* confers anthracnose resistance in chilli pepper (*Capsicum annuum* L.). *Planta.* **254**(1), 5.
<https://doi.org/10.1007/s00425-021-03660-x>
134. Lee K.-R., Jeon I., Yu H., Kim S.-G., Kim H.-S., Ahn S.-J., Lee J., Lee S.-K., Kim H.U. (2021) Increasing monounsaturated fatty acid contents in hexaploid *Camelina sativa* seed oil by *FAD2* gene knockout using CRISPR-Cas9. *Front. Plant Sci.* **12**, 702930.
<https://doi.org/10.3389/fpls.2021.702930>
135. Janga M.R., Pandeya D., Campbell L.M., Konganti K., Villafuerte S.T., Puckhaber L., Pepper A., Stipanovic R.D., Scheffler J.A., Rathore K.S. (2019). Genes regulating gland development in the cotton plant. *Plant Biotechnol. J.* **17**(6), 1142–1153.
<https://doi.org/10.1111/pbi.13044>
136. Chen Y., Fu M., Li H., Wang L., Liu R., Liu Z., Zhang X., Jin S. (2021) High-oleic acid content, nontransgenic allotetraploid cotton (*Gossypium hirsutum* L.) generated by knockout of *GhFAD2* genes with CRISPR/Cas9 system. *Plant Biotechnol. J.* **19**(3), 424–426.
<https://doi.org/10.1111/pbi.13507>
137. Вавилов Н.И. (1920) Закон гомологических рядов в наследственной изменчивости. Доклад на 3-й Всероссийской встрече селекционеров. Саратов, с. 16.
138. Li Z., Liu Z. B., Xing A., Moon B.P., Koellhoffer J.P., Huang L., Cigan A.M. (2015). Cas9-guide RNA directed genome editing in soybean. *Plant Physiol.* **169**(2), 960–970.
139. Butt H., Eid A., Ali Z., Atia M.A., Mokhtar M.M., Hassan N., Mahfouz M.M. (2017) Efficient CRISPR/Cas9-mediated genome editing using a chimeric single-guide RNA molecule. *Front. Plant Sci.* **8**, 1441.
140. Shimatani Z., Kashojiya S., Takayama M., Terada R., Arazoe T., Ishii H., Teramura H., Yamamoto T., Komatsu H., Miura K., Ezura H., Nishida K., Ariizumi T., Kondo A. (2017) Targeted base editing in rice and tomato using a CRISPR-Cas9 cytidine deaminase fusion. *Nat. Biotechnol.* **35**, 441–443.
<https://doi.org/10.1038/nbt.3833>
141. Shimatani Z., Fujikura U., Ishii H., Matsui Y., Suzuki M., Ueki Y., Taoka K., Terada R., Nishida K., Kondo A. (2018) Inheritance of co-edited genes by CRISPR-based targeted nucleotide substitutions in rice. *Plant Physiol. Biochem.* **131**, 78–83.
<https://doi.org/10.1016/J.PLAPHY.2018.04.028>
142. Shimatani Z., Fujikura U., Ishii H., Terada R., Nishida K., Kondo A. (2018) Herbicide tolerance-assisted multiplex targeted nucleotide substitution in rice. *Data Brief.* **20**, 1325–1331.
<https://doi.org/10.1016/J.DIB.2018.08.124>
143. Sun Y., Jiao G., Liu Z., Zhang X., Li J., Guo X., Du W., Du J., Francis F., Zhao Y., Xia L. (2017) Generation of high-amylose rice through CRISPR/Cas9-mediated targeted mutagenesis of starch branching enzymes. *Front. Plant Sci.* **8**, 298.
<https://doi.org/10.3389/fpls.2017.00298>
144. Butler N.M., Baltes N.J., Voytas D.F., Douches D.S. (2016) Geminivirus-mediated genome editing in potato (*Solanum tuberosum* L.) using sequence-specific nucleases. *Front. Plant Sci.* **7**, 1045.
<https://doi.org/10.3389/fpls.2016.01045>
145. Svitashv S., Young J.K., Schwartz C., Gao H., Falco S.C., Cigan A.M. (2015) Targeted mutagenesis, precise gene editing, and site-specific gene insertion in maize using Cas9 and guide RNA. *Plant Physiol.* **169**(2), 931–945.
<https://doi.org/10.1104/pp.15.00793>
146. Braatz J., Harloff H.J., Mascher M., Stein N., Himmelbach A., Jung C. (2017) CRISPR-Cas9 targeted mutagenesis leads to simultaneous modification of different homoeologous gene copies in polyploid oil-seed rape (*Brassica napus*). *Plant Physiol.* **174**(2), 935–942.
<https://doi.org/10.1104/pp.17.00426>

147. Liu X., Ding Q., Wang W., Pan Y., Tan C., Qiu Y., Chen Y., Li H., Li Y., Ye N., Xu N., Wu X., Ye R., Liu J., Ma C. (2022) Targeted deletion of the first intron of the *Wxb* allele via CRISPR/Cas9 significantly increases grain amylose content in rice. *Rice* (N.Y.). **15**(1), 1. <https://doi.org/10.1186/s12284-021-00548-y>
148. Yunyan F., Jie Y., Fangquan W., Fangjun F., Wenqi L., Jun W., Yang X., Jinyan Z., Weigong Z. (2019) Production of two elite glutinous rice varieties by editing *Wx* gene. *Rice Sci.* **26**, 118–124. <https://doi.org/10.1016/j.rsci.2018.04.007>
149. Zhang J., Zhang H., Botella J.R., Zhu J.K. (2018) Generation of new glutinous rice by CRISPR/Cas9-targeted mutagenesis of the *Waxy* gene in elite rice varieties. *J. Integr. Plant Biol.* **60**(5), 369–375. <https://doi.org/10.1111/jipb.12620>
150. Veillet F., Chauvin L., Kermarrec M.P., Sevestre F., Chauvin J.E. (2019) The *Solanum tuberosum* *GBSSI* gene: a target for assessing gene and base editing in tetraploid potato. *Plant Cell Rep.* **38**, 1065–1080. <https://doi.org/10.1007/s00299-019-02426-w>
151. Kusano H., Ohnuma M., Mutsuro-Aoki H., Asahi T., Ichinosawa D., Onodera H., Asano K., Noda T., Horie T., Fukumoto K., Kihira M., Teramura H., Yazaki K., Umemoto N., Muranaka T., Shimada H. (2018) Establishment of a modified CRISPR/Cas9 system with increased mutagenesis frequency using the translational enhancer dMac3 and multiple guide RNAs in potato. *Sci. Rep.* **8**, 13753. <https://doi.org/10.1038/s41598-018-32049-2>
152. Andersson M., Turesson H., Nicolai A., Fält A.S., Samuelsson M., Hofvander P. (2017) Efficient targeted multiallelic mutagenesis in tetraploid potato (*Solanum tuberosum*) by transient CRISPR-Cas9 expression in protoplasts. *Plant Cell Rep.* **36**(1), 117–128. <https://doi.org/10.1007/s00299-016-2062-3>
153. Qi X., Wu H., Jiang H., Zhu J., Huang C., Zhang X., Liu C., Cheng B. (2020) Conversion of a normal maize hybrid into a *waxy* version using *in vivo* CRISPR/Cas9 targeted mutation activity. *Crop J.* **8**, 440–448. <https://doi.org/10.1016/j.cj.2020.01.006>
154. Abe K., Araki E., Suzuki Y., Toki S., Saika H. (2018) Production of high oleic/low linoleic rice by genome editing. *Plant Physiol. Biochem.* **131**, 58–62. <https://doi.org/10.1016/J.PLAPHY.2018.04.033>
155. Okuzaki A., Ogawa T., Koizuka C., Kaneko K., Inaba M., Imamura J., Koizuka N. (2018) CRISPR/Cas9-mediated genome editing of the fatty acid desaturase 2 gene in *Brassica napus*. *Plant Physiol. Biochem.* **131**, 63–69. <https://doi.org/10.1016/J.PLAPHY.2018.04.025>
156. Chandrasekaran J., Brumin M., Wolf D., Leibman D., Klap C., Pearlsman M., Gal-On A. (2016) Development of broad virus resistance in non-transgenic cucumber using CRISPR/Cas9 technology. *Mol. Plant Pathol.* **17**(7), 1140–1153. <https://doi.org/10.1111/mpp.12375>
157. Jia H., Zhang Y., Orbović V., Xu J., White F.F., Jones J.B., Wang N. (2017) Genome editing of the disease susceptibility gene *CsLOB1* in citrus confers resistance to citrus canker. *Plant Biotechnol. J.* **15**(7), 817–823. <https://doi.org/10.1111/pbi.12677>
158. Peng A., Chen S., Lei T., Xu L., He Y., Wu L., Yao L., Zou X. (2017) Engineering canker-resistant plants through CRISPR/Cas9-targeted editing of the susceptibility gene *CsLOB1* promoter in citrus. *Plant Biotechnol. J.* **15**(12), 1509–1519. <https://doi.org/10.1111/pbi.12733>
159. Malnoy M., Viola R., Jung M.H., Koo O.J., Kim S., Kim J.S., Nagamangala Kanchiswamy C. (2016) DNA-free genetically edited grapevine and apple protoplast using CRISPR/Cas9 ribonucleoproteins. *Front. Plant Sci.* **7**, 1904.
160. Nekrasov V., Wang C., Win J., Lanz C., Weigel D., Kamoun S. (2017) Rapid generation of a transgene-free powdery mildew resistant tomato by genome deletion. *Sci. Rep.* **7**(1), 1–6. <https://doi.org/10.1038/s41598017-00578-x>
161. Wang Y., Cheng X., Shan Q., Zhang Y., Liu J., Gao C., Qiu J.L. (2014) Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. *Nat. Biotechnol.* **32**(9), 947–951. <https://doi.org/10.1038/nbt.2969>
162. Zhang Y., Bai Y., Wu G., Zou S., Chen Y., Gao C., Tang D. (2017) Simultaneous modification of three homoeologs of *TaEDR1* by genome editing enhances powdery mildew resistance in wheat. *Plant J.* **91**(4), 714–724. <https://doi.org/10.1111/tpj.13599>
163. Hu X., Cui Y., Dong G., Feng A., Wang D., Zhao C., Zhang Y., Hu J., Zeng D., Guo L., Qian Q. (2019) Using CRISPR-Cas9 to generate semi-dwarf rice lines in elite landraces. *Sci. Rep.* **9**, 19096. <https://doi.org/10.1038/s41598-019-55757-9>
164. Khlestkina E.K., Shvachko N.A., Zavarzin A.A., Börner A. (2020) Vavilov's series of the “green revolution” genes. *Russ. J. Genetics.* **56**(11), 1371–1380. <https://doi.org/10.1134/S1022795420110046>
165. Zhang S., Zhang R., Song G., Gao J., Li W., Han X., Chen M., Li Y., Li G. (2018) Targeted mutagenesis using the *Agrobacterium tumefaciens*-mediated CRISPR-Cas9 system in common wheat. *BMC Plant Biol.* **18**, 302. <https://doi.org/10.1186/s12870-018-1496-x>
166. Kim C.Y., Park J.Y., Choi G., Kim S., Vo K.T.X., Jeon J.S., Lee Y.H. (2022) A rice gene encoding glycosyl hydrolase plays contrasting roles in immunity depending on the type of pathogens. *Mol. Plant Pathol.* **23**(3), 400–416. <https://doi.org/10.1111/mpp.13167>
167. Carey-Fung O., O'Brien M., Beasley J.T., Johnson A.A.T. (2022) A model to incorporate the bHLH transcription factor *OsIRO3* within the rice iron homeostasis regulatory network. *Int. J. Mol. Sci.* **23**, 1635. <https://doi.org/10.3390/ijms23031635>
168. Takeda Y., Tobimatsu Y., Karlen S.D., Koshihara T., Suzuki S., Yamamura M., Murakami S., Mukai M., Hattori T., Osakabe K., Ralph J., Sakamoto M., Umezawa T. (2018) Downregulation of p-COUMAROYL ESTER 3-HYDROXYLASE in rice leads to altered cell wall structures and improves biomass saccharification. *Plant J.* **95**(5), 796–811. <https://doi.org/10.1111/tpj.13988>

Improvement of Crops Using the CRISPR/Cas System: New Target Genes

Y. V. Ukhatova¹, *, M. V. Erastenkova¹, E. S. Korshikova¹, E. A. Krylova¹, A. S. Mikhailova¹,
T. V. Semilet¹, N. G. Tikhonova¹, N. A. Shvachko¹, and E. K. Khlestkina¹

¹*Federal Research Center N.I. Vavilov All-Russian Institute of Plant Genetic Resources (VIR),
Ministry of Science and Higher Education, St. Petersburg, 190000 Russia*

**e-mail: sci_secretary@vir.nw.ru*

Successful application of the CRISPR/Cas genome editing system to various crops largely depends on the correct choice of target genes that may be purposefully changed to improve yield, quality, and resistance to biotic and abiotic stressors. The objective of this work was systematizing and cataloguing the information on the confirmed target genes for crop improvement. The latest systematic review was presented on peer-reviewed scientific papers (indexed in the Scopus database) published before August 17, 2019. The present study covers the period from August 18, 2019 to March 15, 2022. The search according to the given algorithm revealed 2090 publications, and their analysis showed that only 685 original papers contained the results of gene editing for 28 crops (the search included 56 crops). A significant part of these publications described the application of genome editing to target genes previously identified in similar works or the studies were associated with reverse genetics, while only 136 publications contained data on editing new target genes whose modification was aimed at improving plant traits important for breeding. The total number of target genes in cultivated plants that were edited to improve properties of breeding value over the entire period of the CRISPR/Cas system application was 287. A detailed analysis of the editing of new target genes is presented in this review. The studies were most often aimed at increasing plant productivity and disease resistance as well as improving the properties of plant materials. Observations are made whether it was possible to obtain stable transformants at the time of publication and whether the editing technique was applied to non-model cultivars. For a number of crops, however, the range of modified cultivars was significantly expanded, specifically for wheat, rice, soybean, tomato, potato, rapeseed, grapevine, and maize. In a vast majority of cases, agrobacterium-mediated transformation was used to deliver the editing construct; less often it was bioballistics, protoplast transfection or haploinducers. The desired change in traits was most often achieved by gene knockout. In some cases, knockdown and nucleotide substitutions were applied. The base-editing and prime-editing approaches have increasingly been used to make nucleotide substitutions in crop genes. The emergence of a convenient CRISPR/Cas editing system helped to significantly intensify the development of molecular genetics specific to many crop species.

Keywords: base-editing, CRISPR/Cas, prime-editing, biotechnology, target genes, genome editing, cultivated plants, targeted mutagenesis

УДК 577.2.01

СТРОЕНИЕ, ЭКСПРЕССИЯ И НЕКАНОНИЧЕСКИЕ ФУНКЦИИ рДНК ЧЕЛОВЕКА: РОЛЬ НЕКОДИРУЮЩИХ РЕГИОНОВ

© 2023 г. А. А. Садова^{a, b, *}, Д. Ю. Пантелеев^a, Г. В. Павлова^{a, c, d}

^aИнститут высшей нервной деятельности и нейрофизиологии Российской академии наук,
Москва, 117485 Россия

^bРоссийский национальный исследовательский медицинский университет им. Н.И. Пирогова,
Москва, 117997 Россия

^cНациональный медицинский исследовательский центр нейрохирургии им. академика Н.Н. Бурденко,
Москва, 125047 Россия

^dПервый Московский государственный медицинский университет им. И.М. Сеченова,
Москва, 119435 Россия

*e-mail: 89652410866@mail.ru

Поступила в редакцию 09.09.2022 г.

После доработки 08.11.2022 г.

Принята к публикации 16.11.2022 г.

Гены рРНК — одни из самых, на первый взгляд, эволюционно-консервативных последовательностей, при ближайшем рассмотрении удивляют разнообразием строения и набором выполняемых функций. Некодирующие области рибосомных генов содержат регуляторные последовательности, сайты узнавания различных белков, псевдогены, повторяющиеся элементы и гены микроРНК. Рибосомные межгенные спейсеры отвечают не только за морфологические особенности ядрышка и его функционирование — экспрессию рРНК и биогенез рибосом, но и контролируют образование гетерохроматина в ядре, опосредуя дифференцировку клеток. Изменение экспрессии отдельных частей огромных (по сравнению с кодирующими) не кодирующих регионов рДНК в ответ на поступающие сигналы помогает клетке реагировать на различные виды стресса. Нарушение этого процесса может приводить к возникновению различных заболеваний — от онкологических до нейродегенеративных и психических расстройств. В представленном обзоре рассмотрены строение и транскрипция рибосомного межгенного спейсера, а также роль этого спейсера в экспрессии рРНК, развитии наследственных заболеваний и опухолей у человека.

Ключевые слова: рДНК, рРНК, рибосомный межгенный спейсер, IGS-РНК, промоторная РНК, РНК PAPAS

DOI: 10.31857/S0026898423030084, **EDN:** CHNVAJ

ВВЕДЕНИЕ

Рибосомные гены, кодирующие рРНК — главный компонент рибосом, играют важную роль в жизнедеятельности клеток, при этом гены рРНК человека остаются недостаточно изученными.

Сокращения: днРНК — длинные не кодирующие РНК; нкРНК — не кодирующие РНК; рМГС — рибосомный межгенный спейсер; ETS — внешний транскрибируемый спейсер (External Transcribed Spacer); IGS-РНК — РНК, синтезируемые в области рМГС; ITS — внутренний транскрибируемый спейсер (Internal Transcribed Spacer); NOR — ядрышковый организатор (Nucleolus Organizer Region); NoRC — ядрышковый комплекс ремоделирования хроматина (Nucleolar Remodeling Complex); NuRD — комплекс ремоделирования и деацетилирования нуклеосом (Nucleosome Remodeling and Deacetylation); PAPAS — антисмысловые промоторные и пре-рРНК (Promoter and Pre-rRNA Antisense).

Нуклеотидная последовательность рДНК была расшифрована более двух десятилетий назад, но структурные и функциональные особенности повторяющихся элементов, составляющих рибосомный межгенный спейсер (рМГС), не установлены. Эти последовательности не относятся к “мусорной” ДНК, как считалось ранее, они играют важную роль в регуляции роста и пролиферации клеток, а также в их дифференцировке [1]. При нарушении работы рибосомных генов под угрозой находится весь организм, поскольку эти области ДНК не только участвуют в формировании рибосом, но и обеспечивают нормальное функционирование генома, контролируя процессы образования гетерохроматина и подавления экспрессии различных групп генов [2]. Так, в клетках, подвергающихся малигнизации, ядрыш-

ки, главным компонентом которых являются гены рРНК, морфологически отличаются от ядрышек в здоровых клетках, что свидетельствует о перераспределении гетерохроматина в ядре и возможном подавлении транскрипции генов, ответственных за нормальное функционирование клеток [3]. Однако точный механизм этого переключения не до конца понятен. В настоящее время ведется поиск онкомаркеров, которые могли бы оказаться полезными при диагностике опухоли, прогнозировании опухолевой прогрессии и эффективности проводимой терапии [4]. Например, большое внимание привлекает изучение профиля экспрессии некодирующих РНК (нкРНК) в опухолях человека различного происхождения и изменений этого профиля в зависимости от степени злокачественности опухоли [5]. При этом создание панели нкРНК, изменение экспрессии которых может быть характерным для той или иной опухоли и стадии ее развития, затрудняется постоянным обнаружением все новых некодирующих транскриптов [6].

Некодирующие транскрипты, образующиеся в области рМГС, представляют огромный научный и практический интерес, поскольку они могут вносить вклад в регуляцию экспрессии рибосомных генов. Это предположение основано на различиях в профиле экспрессии этих транскриптов, а также в профиле экспрессии рРНК в нормальных и опухолевых клетках. Подобные различия могут использоваться для ранней диагностики и прогноза развития опухоли, определения ее злокачественности [7]. С другой стороны, показано, что в предпромоторной области рДНК млекопитающих синтезируются регуляторные транскрипты, которые стимулируют дифференцировку или специфический ответ клеток на изменение условий среды [8, 9]. Существование аналогичных молекул у человека предсказано на основе биоинформатического анализа данных секвенирования ДНК и транскриптомного анализа, их поиск ведется в настоящее время.

В нашем обзоре проведен последовательный анализ данных о строении рДНК человека, механизмах регуляции экспрессии рРНК, а также о роли рМГС в транскрипции рибосомных генов. Представлено также краткое описание патологий, возникающих при нарушениях функционирования ядрышка.

СТРУКТУРНО-ФУНКЦИОНАЛЬНАЯ ОРГАНИЗАЦИЯ РИБОСОМНЫХ ГЕНОВ ЧЕЛОВЕКА

Рибосомные гены — последовательности ДНК, кодирующие рРНК — составляют значительную часть генома человека и формируют в ядре клетки особую зону синтеза рибосом, называемую ядрышком [10]. Четыре вида рРНК эукариот коди-

руются двумя типами рДНК: 5S рДНК располагается на хромосоме 1, транскрибируется РНК-полимеразой III (PolIII) и кодирует лишь 5S рРНК; тогда как гены 18S, 28S и 5.8S рРНК транскрибируются РНК-полимеразой I (PolI) в виде единого предшественника (45S рРНК), включающего внешние (5'- и 3'-ETS) и внутренние (ITS) транскрибируемые спейсеры, который затем подвергается процессингу [11]. Последовательности, кодирующие 45S рРНК, имеют длину около 13 т.п.н., они локализуются на коротких плечах пяти акроцентрических хромосом (13, 14, 15, 21 и 22), образуя кластеры, в которых гены расположены один за другим в ориентации “голова-к-хвосту” и разделены участками длиной около 30 т.п.н., называемыми рМГС [12]. Кластеры рибосомных генов формируют ядрышковые организаторы (NOR, nucleolus organizer regions), которые вместе с другими участками генома и многочисленными белками определяют структуру и морфологию ядрышка (рис. 1) [1].

5S рДНК и 45S рДНК являются повторяющимися областями генома: в геноме человека обнаружено более 200 копий последовательностей, кодирующих 5S рРНК, и более 400 копий 45S рРНК. Сообщение о существовании корреляции между числом копий 5S и 45S рДНК [13] было недавно опровергнуто [14]. Число копий рДНК варьирует в геномах и популяциях человека, а также может изменяться при малигнизации клеток [15]. Интересно отметить, что варьирует не только число копий, но и последовательность каждого рДНК-повтора. Все это приводит к нестабильности локуса, которая, с одной стороны, считается причиной патологии, а с другой, частью нормальной физиологии клетки [16].

К 3'-области рДНК, кодирующей 28S рРНК, прилегают так называемые R-повторы, содержащие Sal-боксы, необходимые для терминирования транскрипции (рис. 1) [17]. Показано, что при участии РНК-полимеразы I и фактора терминирования транскрипции TTF-1 (transcription termination factor 1) Sal-боксы функционируют также как барьеры вилки репликации (RFB, replication fork barrier), причем в клетках человека, в отличие от других млекопитающих, эти RFB останавливают репликативные вилки независимо от направления их движения [18].

рМГС содержит множество регуляторных последовательностей, способных к связыванию с различными белками и к образованию некодирующих РНК [19]. Они включают микросателлитные и Alu-повторы, консервативные у человека и приматов, псевдоген *cdc27* и многочисленные сайты узнавания фактора транскрипции с-Мус [20, 21].

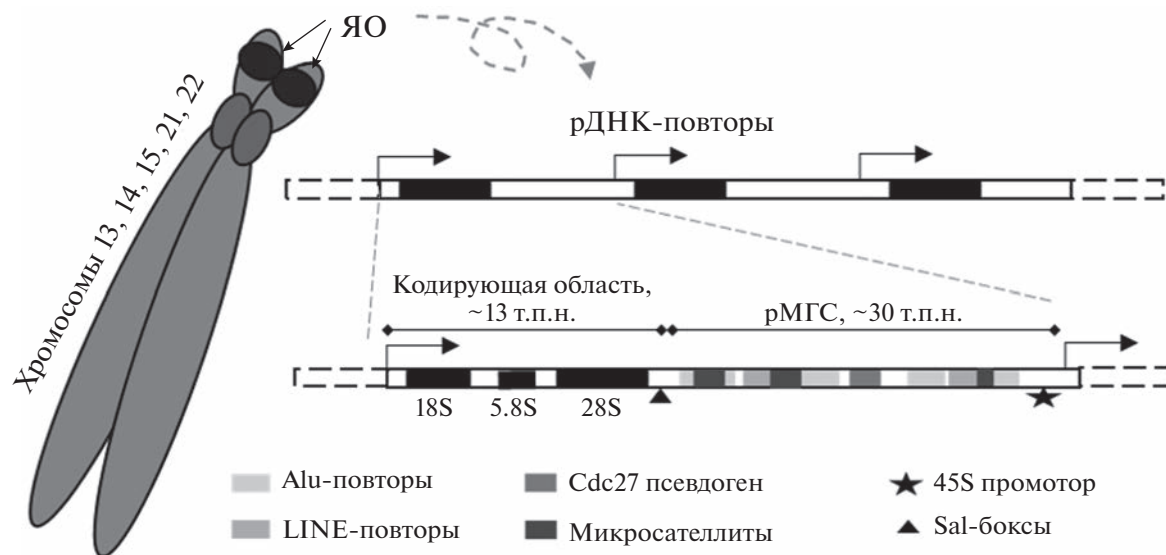


Рис. 1. Схема организации рибосомных генов человека на пяти акроцентрических хромосомах. ЯО — ядрышковые организаторы; рМГС — рибосомный межгенный спейсер; LINE — длинные диспергированные повторы. (Рисунок автора на основе последовательности U13369.1 в базе данных GenBank.)

УЧАСТИЕ РИБОСОМНЫХ ГЕНОВ В ОБРАЗОВАНИИ ГЕТЕРОХРОМАТИНА

Регионы, кодирующие 45S рРНК, относятся к одним из самых транскрипционно активных участков генома, при этом немалая их часть упакована в конститутивный гетерохроматин. Формирование гетерохроматина ядрышка влияет на упаковку в неактивный хроматин и других областей генома: удаление части генов рРНК приводит к снижению образования гетерохроматина во всем ядре клеток дрозофилы и уменьшению репрессии тех или иных генов [2].

Предполагается, что, соседствуя с прицентромерными (у мышей) или прителомерными (у человека) областями, рибосомные гены могут распространять гетерохроматизацию на эти участки генома, формируя перинуклеолярный гетерохроматин [20]. При этом ассоциированными с ядрышком оказываются не только теломерные повторы и центромерная сателлитная ДНК, но и “молчащие” участки большинства хромосом человека, названные доменами, ассоциированными с ядрышком (NAD, nucleolar-associated domains) [21]. В NAD идентифицировано более тысячи генов, которые относятся к разным семействам (например, гены обонятельных рецепторов, Т-клеточных рецепторов и иммуноглобулинов), но обладают при этом важными общими чертами: эти гены входят в состав больших генных кластеров, а их экспрессия тканеспецифична [22]. Методом Hi-C показана ассоциация 45S рДНК с генами, продукты которых локализуются в митохондриях, а также с генами, кодирующими рибосомные белки [23].

Образование гетерохроматина в ядрышке регулируется хроматинре моделирующим комплексом NoRC (см. раздел Регуляция экспрессии рибосомных генов), ключевым компонентом которого является белок TIP5 [24]. Подавление экспрессии белка TIP5 приводит к нарушению формирования не только внутриядрышкового, но и перинуклеолярного гетерохроматина, включающего центромерные и перичентромерные участки хромосом [25]. Нокдаун всего комплекса вызывает нарушение сегрегации хромосом в митозе, что ведет к хромосомным aberrациям и дестабилизации генома [26].

С TIP5 взаимодействует множество белков и других регуляторных молекул, локализующихся как в ядрышке, так и в нуклеоплазме, и участвующих в ремоделировании хроматина. Показано, что в образовании ядрышкового хроматина участвует поли(ADP-рибоза)полимераза 1 (PARP1), которая контактирует с TIP5 посредством промоторной РНК: после связывания этого комплекса с неактивными промоторами рибосомных генов PARP1 модифицирует гистоны и другие белки хроматина с помощью поли(ADP-рибозы), вызывая эпигенетически наследуемое подавление экспрессии этих участков рДНК [27]. Однако если PARP1 локализуется в нуклеоплазме, то это, напротив, способствует формированию открытого хроматина и активирует транскрипцию [28]. PARP1 контролирует активное состояние промоторов генов плюрипотентности в стволовых клетках, поэтому при его нокдауне наблюдаются признаки дифференцировки клеток [29]. Как и все семейство поли(ADP-рибоза)полимераз, PARP1

выполняет множество функций, начиная от регуляции экспрессии генов и управления механизмом репарации ДНК, до участия в процессах воспаления и апоптоза [30]. Тот факт, что PARP1 обнаруживается в активно транскрибируемых участках генома и присутствует при этом в неактивном гетерохроматине ядрышка, говорит о многогранной роли этого фермента и подтверждает сложный характер взаимодействий между доменами хроматина ядерных структур [31].

Еще один партнер TTP5 — Тау. В ядре клетки он ассоциирован с перичентромерным хроматином и отвечает за его стабилизацию [32]. Нокдаун этого гена приводит к снижению количества репрессивных гистоновых меток в ядрышковом гетерохроматине и способствует повышению транскрипции рибосомных генов. По-видимому, Тау может препятствовать внесению модификаций хроматинремоделирующим комплексом NoRC [33].

РЕГУЛЯЦИЯ ЭКСПРЕССИИ РИБОСОМНЫХ ГЕНОВ

Поскольку молекулы рРНК являются основным компонентом рибосом, регуляция транскрипции их генов и процессинга тесно связаны с функционированием, ростом и делением клеток. Инициация транскрипции генов 45S рРНК происходит в результате связывания белка UBF (upstream binding factor) — главного транскрипционного фактора РНК-полимеразы I — с элементом UCE (upstream control element) в промоторной области рДНК. UBF стимулирует инициацию транскрипции, вытесняя гистон H1, расплетая ДНК и привлекая РНК-полимеразу I на промотор [34]. Именно уровень этого белка в клетке определяет активность доступных для транскрипции промоторов рДНК; в отсутствие UBF экспрессия рибосомных генов подавляется [35]. Для формирования полного преинициаторного комплекса необходимо, чтобы с коровым промотором (CP, core promoter element) связался комплекс белков SL1, в состав которого входят TBP (TATA-binding protein) и ассоциированные с ним многочисленные факторы TAF_I (TBP-associated factors) [36]. Последние отвечают за рекрутирование фермента и его специфическое взаимодействие с последовательностью промотора, а также за привлечение дополнительных факторов транскрипции [37]. Таким образом, представление об РНК-полимеразе I как о холоферменте, состоящем из многих белков, необходимых как для инициации транскрипции, так и для элонгации, определяет возможность сложноорганизованной регуляции экспрессии рибосомных генов (рис. 2а) [38].

Как уже сказано, геном человека содержит большое количество копий рибосомных генов, при этом значительная часть рДНК в дифференцированных клетках не транскрибируется, по-

скольку упакована в гетерохроматин. Сочетания модификаций гистонов контролируют экспрессию рибосомных генов. Несмотря на присутствие консервативных маркеров в определенных частях рДНК, клеточные популяции могут иметь собственный профиль (набор и распределение) гистоновых модификаций. Например, в эмбриональных стволовых клетках, кератиноцитах, эндотелиальных и опухолевых клетках H3K4me2 присутствует на участке 28–29 т.п.н. рМГС и в предпромоторной области, а распределение H3K27me3 во всех четырех типах клеток кардинально отличается [43]. Неактивное состояние рибосомных генов поддерживается благодаря сложной согласованной работе нескольких белковых комплексов, отвечающих за ремоделирование хроматина [44].

Ядрышковый ремоделирующий комплекс (NoRC, nucleolar remodeling complex) состоит из субъединицы SNF2h (sucrose-nonfermenting protein 2 homolog), обладающей АТФазной активностью, и белка TTP5 (TTF-I-interacting protein 5), содержащего ДНК-связывающий домен и взаимодействующего с фактором терминации транскрипции TTF-I. Формирование комплекса начинается с TTF-I, который, связываясь с проксимальным элементом промотора T₀, рекрутирует остальные компоненты NoRC. Комплекс подавляет экспрессию нижележащего рибосомного гена за счет сдвига нуклеосомы относительно промотора, что препятствует инициации транскрипции на данном участке [24, 45]. При этом к промотору привлекаются ферменты, ответственные за метилирование ДНК (DNMT1 и DNMT3b) и деацетилирование гистонов (HDAC1), что приводит к образованию гетерохроматина в этой области и долговременному сайленсингу рибосомных генов (рис. 2б) [8].

Таким образом, NoRC отвечает за формирование конститутивного гетерохроматина, который эпигенетически наследуется клетками в процессе дифференцировки [46]. Здесь необходимо отметить, что метилирование ДНК в области промоторов и энхансеров рДНК у млекопитающих служит своеобразным маркером “молчащих” рибосомных генов, а также старения клеток. Эти данные позволяют судить о возрасте организма и даже предсказывать продолжительность жизни [47]. Белок TTF-I не только участвует в терминации транскрипции рибосомных генов, но может также определять их топологию. Показано, что активно транскрибирующиеся гены рРНК образуют ДНК-петли, в основании которых находятся промотор и терминатор, а их взаимодействие опосредуется TTF-I [48].

Более сложно устроен комплекс NuRD (nucleosome remodeling and deacetylation), состоящий из нескольких субъединиц и осуществляющий обра-

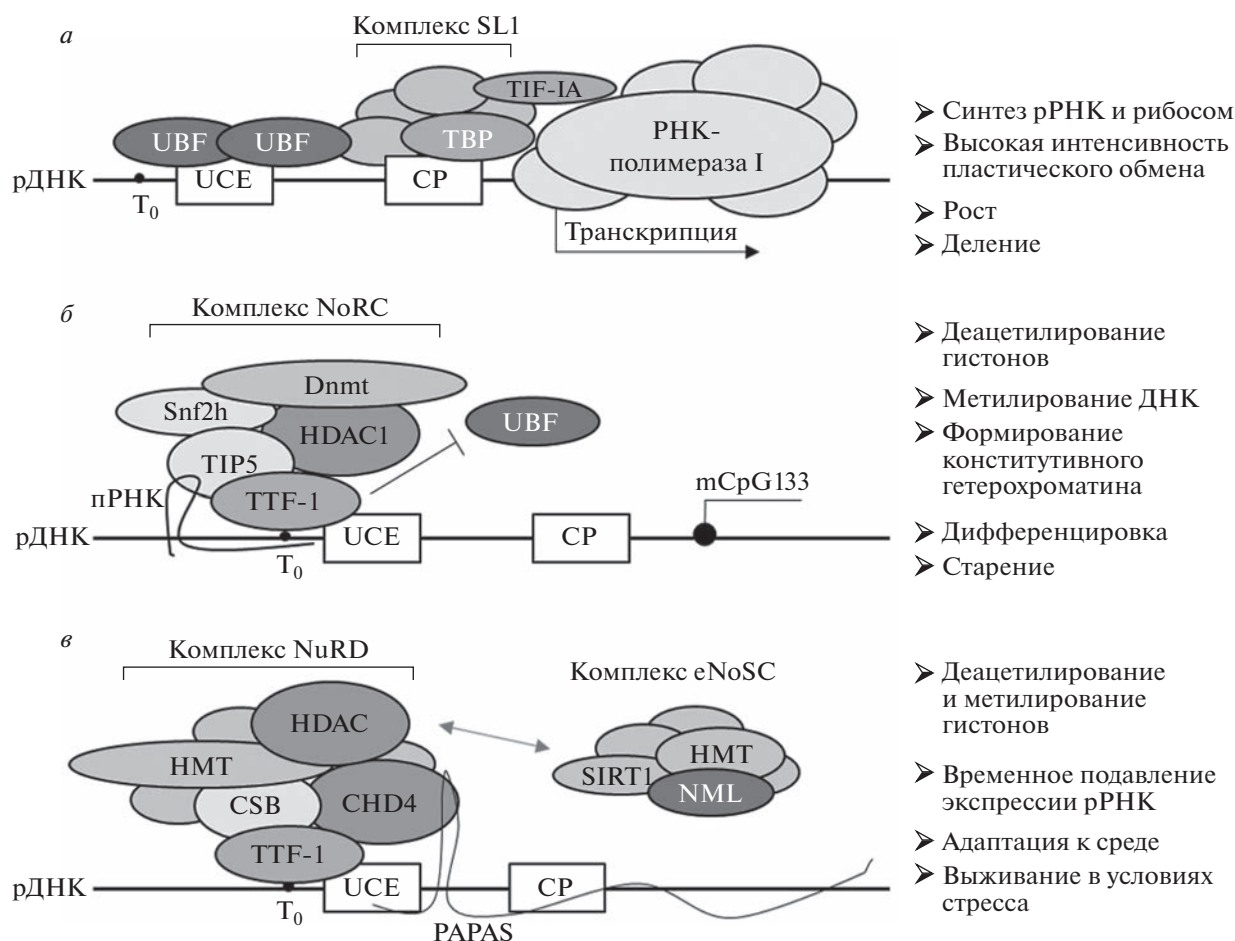


Рис. 2. Схема регуляции экспрессии рДНК. *а* — Преенициаторный комплекс (PIC) РНК-полимеразы I. Сборка PIC осуществляется за счет взаимодействия UBF, который связывается с элементом UCE, комплекса SL1, ассоциированного с коровым промотором через TBP, а также транскрипционного фактора TIF-1A и РНК-полимеразы I. Кроме TBP, в состав SL1 входят факторы специфичности TAF₁ (согласно [39]). *б* — Сборка комплекса NoRC на промоторе рДНК: TTF-1 узнает T₀, привлекает в комплекс TIP5, ассоциированный с ним белок Snf2, затем включаются деацетилаза гистона H4 (HDAC1) и ДНК-метилтрансфераза (Dnmt). Дополнительно белок TIP5 может рекрутироваться к промоторной РНК, комплементарно взаимодействующей с промоторной областью рДНК. Таким образом, NoRC обеспечивает репрессию транскрипции рДНК, поскольку UBF не может связаться с промотором, а транскрипция блокирована (согласно [40]). *в* — Сборка комплекса NuRD, опосредованная TTF-1, происходит после взаимодействия дефосфорилированного белка CHD4 и некодирующей РНК PAPAS, активация которых может быть вызвана повышением температуры (согласно [41]). В условиях недостатка глюкозы активируется деацетилаза SIRT1, которая после взаимодействия с TTF-1 рекрутирует другие белки комплекса eNoSC, осуществляющие образование гетерохроматина в этой области (согласно [42]).

тимое подавление экспрессии рибосомных генов в дифференцированных клетках (рис. 2*в*). В состав NuRD входит множество белков: деацетилазы гистонов, АТФ-зависимые хеликазы, а также белки, не имеющие каталитической активности, но взаимодействующие с ДНК. Как и NoRC, NuRD привлекается на промотор рДНК белком TTF-1, связанным с T₀. При этом в отличие от длительного подавления экспрессии, наблюдаемого при воздействии NoRC, NuRD временно останавливает транскрипцию рибосомных генов, а рДНК остается доступной для транскрипционных факторов [49]. Так, при взаимодействии это-

го комплекса с белком CSB (Cockayne syndrome group B) на рДНК привлекаются гистон-метилтрансферазы, вносящие модификации, характерные для активно транскрибируемых генов [50].

Вообще, экспрессия рДНК характеризуется высокой пластичностью в ответ на такие стимулы окружающей среды, как доступность питательных веществ и факторов роста [51].

Специализированный трехкомпонентный комплекс eNoSC (energy-dependent nucleolar silencing complex), обратимо подавляющий экспрессию рибосомных генов в условиях голодания (недостатка глюкозы), состоит из белка SIRT1, гистон-

метилтрансферазы SUV39H1 и нуклеометилина NML (nucleomethilin). Сборка комплекса на рДНК, как и в случае с NuRD, опосредуется TTF-I. Повышение уровня NAD^+ в клетке (маркер низкоэнергетического состояния) напрямую активирует SIRT1, который деацетилюет гистон H3 в области промотора рДНК, переводя рибосомный ген в неактивное состояние (рис. 2б) [52]. Интересно, что этот путь регулирует не только биосинтез рРНК и рибосом, но и определяет дальнейшую судьбу клетки в условиях недостатка питания. Снижение содержания РНК в ядрышке в присутствии eNoSC вызывает транслокацию белков из ядрышка в нуклеоплазму и активацию белка p53, ответственного за запуск репарации ДНК, задержку клеточного цикла и инициацию апоптоза [53].

В регуляции экспрессии рибосомных генов участвуют не только белки, но и множество нкРНК, функция которых чаще всего связана с определением специфичности того или иного белкового комплекса. нкРНК обеспечивают очень точную регуляцию экспрессии в ответ на стимулы различной природы и стремительно меняющиеся условия среды [54].

Длинные некодирующие РНК (днРНК), происходящие из самих рибосомных генов, могут как прямо, так и опосредованно регулировать экспрессию рДНК. Например, в условиях теплового шока нкРНК IGS₁₆RNA и IGS₂₂RNA, транскрибируемые с соответствующих областей рМГС, участвуют в структурном и функциональном ремоделировании ядрышка: они контролируют изменение локализации белков и подавление транскрипции рДНК [55].

Однако основная роль некодирующих транскриптов, синтезирующихся с рДНК, заключается в привлечении компонентов ремоделирующих комплексов и облегчении их сборки. В клетках млекопитающих в предпромоторной области рДНК экспрессируется промоторная РНК, которая участвует в подавлении экспрессии рибосомных генов за счет двух функционально отличающихся доменов [56]. Эта молекула длиной ~200 нуклеотидов имеет 20-нуклеотидную последовательность на 5'-конце, комплементарную промотору рДНК и T₀ – сайту связывания TTF-I, и может образовывать в этом месте триплекс ДНК-РНК, который специфически узнается ДНК-метилтрансферазой DNMT3b, вносящей репрессивные метки в эту область. Именно участок узнавания промотора определяет функциональность и эффективность всего процесса сайленсинга рДНК [57]. Вторая ключевая последовательность промоторной РНК состоит из ~90 нуклеотидов и способна образовывать устойчивую структуру типа стебель с петлей, которая необходима для взаимодействия с TIP5 и сборки NoRC (рис. 2б) [58].

Область, кодирующая промоторную РНК, может транскрибироваться РНК-полимеразой II в обратной ориентации, при этом образуются гетерогенные антисмысловые днРНК, названные PAPAS (promoter and pre-rRNA antisense), которые могут комплементарно взаимодействовать с промотором и началом рРНК-кодирующей последовательности [56]. При недостатке питания или остановке роста в клеточных культурах повышается содержание РНК PAPAS, которые привлекают гистон-метилтрансферазу Suv4-20h2, что стимулирует образование гетерохроматина и снижение синтеза рРНК [59]. Повышение продукции PAPAS происходит также при других видах стресса (например, при тепловом шоке), однако в этом случае они по-другому регулируют образование гетерохроматина. Показано, что РНК PAPAS, имеющая участки, комплементарные энхансерной и промоторной области рДНК, связывается с ней, образуя РНК-ДНК-триплекс за счет хугстиновских взаимодействий, и привлекает дефосфорилированный в условиях стресса белок CHD4 – компонент комплекса NuRD, отвечающего за сайленсинг рибосомных генов (рис. 2б) [41].

Показана особая регуляция процессов обучения и синаптической пластичности в клетках гиппокампа, где днРНК (LoNA) подавляет экспрессию рРНК на уровне транскрипции, привлекая белок нуклеолин (NLC) и стимулируя образование гетерохроматина в области промотора рДНК за счет модификации гистонов. В этом случае ни UBF, ни РНК-полимераза I не могут связаться с промоторной областью, что приводит к нарушению формирования преинициаторного комплекса [60]. Сходным образом, по-видимому, действуют продукты транскрипции Alu-повторов, входящих в состав интронов белкокодирующих генов, транскрибируемых РНК-полимеразой II (AluРНК): показано их взаимодействие с нуклеолином. Очевидно, что эти транскрипты исключительно важны для организации гетерохроматина ядрышка, поскольку при специфическом ингибировании AluРНК антисмысловыми олигонуклеотидами, как и при ингибировании РНК-полимеразы II α -аманитином, ядрышко распадается на мелкие домены, что не наблюдается при ингибировании рибосом или других РНК-полимераз [61]. В эмбриональных стволовых клетках человека экспрессируется необычная РНК SLERT (snoRNA-ended lncRNA enhances pre-ribosomal RNA transcription), модулирующая активность РНК-хеликазы DDX21, которая ингибирует РНК-полимеразу I. Подавление или потеря SLERT приводит к замедлению роста и пролиферации клеток в колониях, а в организме голых мышей (Nude) с ксенографтами клеток с нокаутом SLERT – к сниженной массе опухоли по сравнению с контролем [62].

ТРАНСКРИПЦИЯ РИБОСОМНОГО МЕЖГЕННОГО СПЕЙСЕРА

Результаты ряда исследований подтверждают факт транскрипции рМГС у млекопитающих и человека [63, 64]. Зоны рМГС, обогащенные консервативными последовательностями, ассоциированы с точками начала транскрипции и транскрибируются во многих тканях и клеточных линиях [65]. Показано, что рДНК транскрибируется как в прямом (смысловом), так и в обратном (антисмысловом) направлении, в результате образуются нкРНК, выполняющие определенные функции [54, 64, 65].

Изучение малых нкРНК, синтезирующихся в придатках яичек человека, позволило предсказать существование микроРНК hsa-mir-6724, гомологичной последовательности предпромоторной области рМГС [66]. И хотя роль hsa-mir-6724 пока не установлена, предполагается, что она может быть связана со старением стволовых клеток [67].

В клеточных линиях человека обнаружены днРНК (IGS₂₁RNA и IGS₂₈RNA), транскрибирующиеся с соответствующих областей рМГС в условиях ацидоза, а также IGS₁₆RNA и IGS₂₂RNA, которые транскрибируются и выполняют определенные функции в условиях теплового шока, привлекая в ядрышко различные белки [53, 68]. Предполагается, что эти молекулы синтезируются независимо от рРНК и других транскриптов в этой области, то есть они могут иметь собственные промоторы, однако пока это не установлено [68].

Показано, что в модуляции ядрышковой организации и регуляции транскрипции рибосомных генов в ходе дифференцировки клеток мышей может участвовать промоторная РНК [69], которая транскрибируется РНК-полимеразой I и контролируется собственным спейсерным промотором. Предшественник этой молекулы (длинной ~2 т.н.) подвергается процессингу с образованием транскриптов длиной 150–300 нуклеотидов, которые разрушаются, если не оказываются связанными и стабилизированными фактором TIP5 [70]. Обнаружено повышение уровня первичного транскрипта, дающего начало промоторной РНК, в эмбриональных стволовых клетках (ESC), в то время как в нейральных стволовых клетках (NPC) обнаружена процессированная молекула промоторной РНК. Предполагается, что созревание промоторной РНК происходит в ходе дифференцировки клеток, когда эта РНК способствует формированию гетерохроматина (рис. 3) [9].

В рДНК человека выявлены области, гомологичные спейсерному промотору у мышей и других грызунов, в которых идентифицированы сайты связывания белков UBF и TBP, а также РНК-полимеразы I, причем к TBP прилегает консенсусная последовательность, узнаваемая TTF-I (T_{sp}).

По-видимому, спейсерный промотор человека является активным и регулируется подобно промотору у мышей (рис. 4) [71]. Экспрессия этого участка выявлена в линиях клеток рака легкого человека, а образующаяся в результате нкРНК ассоциирована с регуляцией транскрипции рДНК [72]. Соответствующий промоторной РНК транскрипт обнаружен в клеточной линии K562 хронического миелолейкоза человека [43].

В области рМГС синтезируется еще одна хорошо изученная РНК — PAPAS. Эта длинная (10 т.н.) антисмысловая РНК транскрибируется РНК-полимеразой II независимо от промоторов (инициация происходит в случайных сайтах) в условиях стресса и теплового шока. Как и промоторная РНК, PAPAS способствует структурированию хроматина в области промотора рибосомных генов [59].

Изучение экспрессии РНК в различных частях рМГС в клетках человека показало, что время жизни этих транскриптов неодинаково [73]. Более стабильными и имеющими более длинный период полураспада по сравнению с рРНК оказались молекулы, синтезирующиеся в предпромоторной области и соответствующие промоторной РНК. При этом более устойчивыми к распаду оказались РНК, происходящие из областей IGS₂₄ и IGS₂₈ [73], также исследованные в этой работе.

Обнаружено, что гиперпродукция белка РНФ6, наблюдаемая при различных заболеваниях, снижает синтез пре-рРНК, мешая РНК-полимеразе I во время элонгации, но повышает экспрессию некодирующих регионов IGS₃₆ и IGS₃₉ рибосомных генов [74]. Предполагается, что экспрессия этих участков индуцируется при транскрипционном стрессе, вызванном ингибированием транскрипции генов рРНК, а образующиеся в результате транскрипты могут способствовать сайленсингу рРНК-кодирующих областей. Интересно, что указанные молекулы синтезируются РНК-полимеразой I, а их количество обратно пропорционально количеству рРНК. Это значит, что в нормальных условиях фермент генерирует предшественник рРНК, а в условиях стресса “переключается” на некодирующие участки и транскрибирует регуляторные молекулы, необходимые для защиты ядра и клетки.

Недавно мы обнаружили экспрессию предпромоторной области рДНК в клеточных культурах глиом и глиобластом [75], а также различия в экспрессии отдельных участков исследуемой области в опухолевых и в условно нормальных клетках. Участие рМГС и транскриптов, происходящих из его отдельных областей, в борьбе клетки с различными видами стресса подтверждается резким повышением экспрессии области IGS₄₀ с образованием антисмысловых транскриптов при обработке

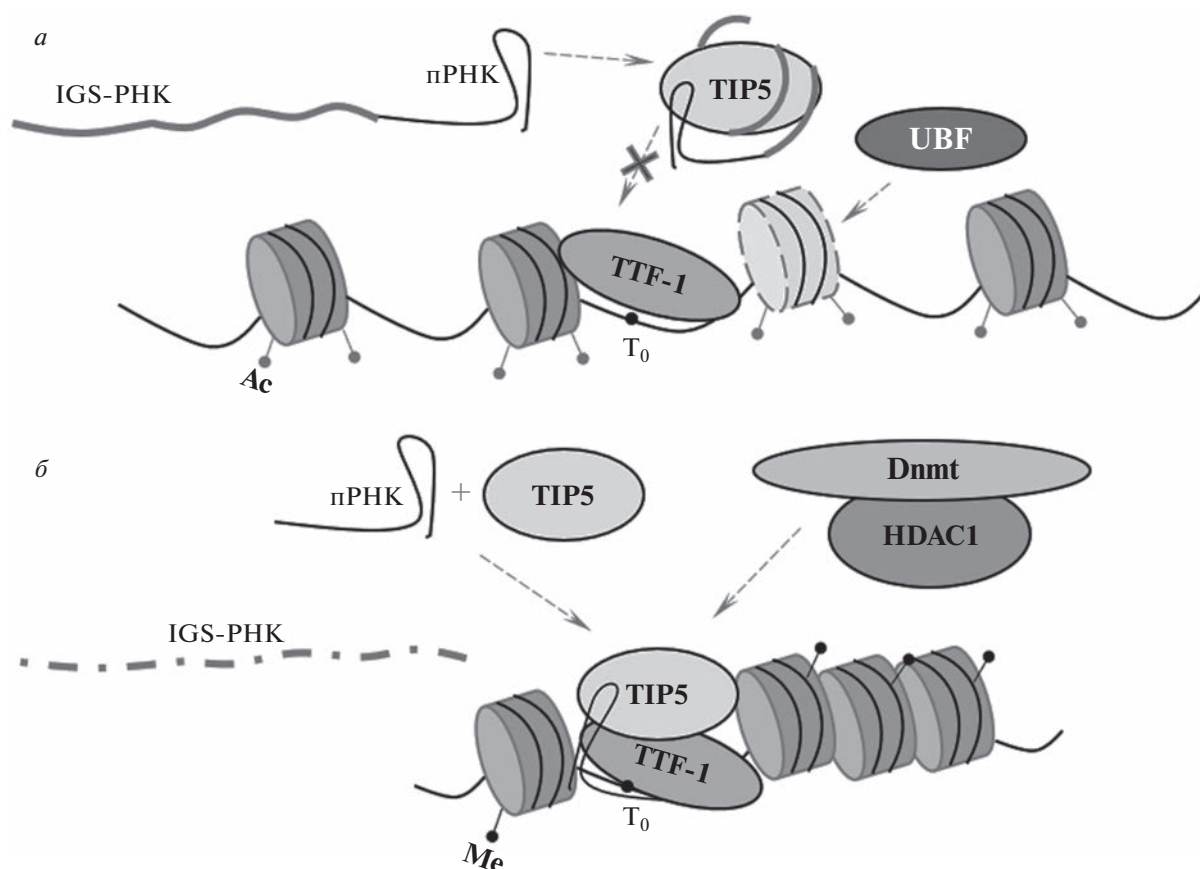


Рис. 3. Модель организации хроматина в ядре и ядрышке эмбриональных стволовых клеток (ESC) (а) и дифференцированных клеток (б). В ESC непротранскрибированный транскрипт IGS-PHK препятствует взаимодействию TIP5 с TTF-1. В этом случае не происходит привлечения ремоделирующего комплекса NoRC и образования гетерохроматина в этой области. По мере дифференцировки клеток IGS-PHK подвергается процессингу с образованием промоторной РНК (пРНК), которая способствует взаимодействию TIP5 с TTF-1, в результате чего рекрутируется ремоделирующий комплекс, транскрипция подавляется, формируется гетерохроматин. (Рисунок автора по описанию, данному в [9].)

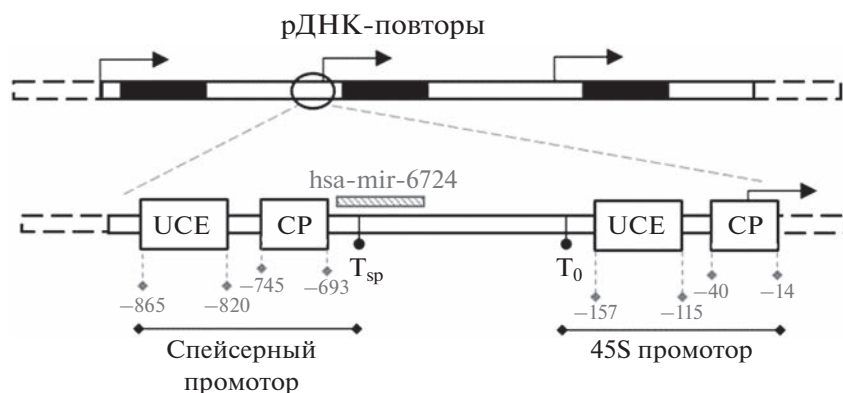


Рис. 4. Схема организации предпромоторного участка рибосомных генов человека. Стрелка указывает точку начала транскрипции рРНК; UCE – вышележащий контрольный элемент; CP – коровый промотор; T_{sp} и T₀ – сайты связывания транскрипционного фактора TTF-I.

клеток альфа-аманитином. Этот эффект нивелировался актиномицином Д, что может указывать на участие РНК-полимеразы I в транскрипции рДНК в обратном направлении [76].

Наконец, недавно было показано прямое участие РНК-полимеразы II в биогенезе рибосом [77]. И хотя о присутствии этого фермента в ядрышке ранее не сообщалось, оказалось, что он ассоциирован со спейсерными участками рибосомных генов и вовлечен в транскрипцию рМГС в антисмысловом направлении. Образующиеся при этом длинные РНК можно сравнить со своеобразным “щитом”, который закрывает некодирующие области для РНК-полимеразы I. Эта гипотеза хорошо согласуется с приведенными выше данными, поскольку при нарушении работы РНК-полимеразы II — транскрипционном стрессе — РНК-полимераза I начинает синтезировать упомянутые ранее нкРНК (IGS₁₆RNA, IGS₂₁RNA, IGS₂₂RNA, IGS₂₈RNA), что вызывает реорганизацию ядрышка и прекращение синтеза рРНК и сборки рибосом.

НАРУШЕНИЕ СИНТЕЗА РИБОСОМ ПРИ НАСЛЕДСТВЕННЫХ ЗАБОЛЕВАНИЯХ

Нарушение образования гетерохроматина в рДНК может привести к дестабилизации генома, дезинтеграции ядрышка, старению клетки и развитию множества заболеваний [78]. Поскольку активность рибосомных генов регулируется различными механизмами, изменение биогенеза рибосом может быть обусловлено как нарушениями в самих рибосомных генах, так и в генах, кодирующих регуляторные белки и РНК.

Так, в рДНК из фибробластов, полученных от пациентов с синдромом Вернера, увеличено количество метилированных CpG-островков, а продолжительность жизни этих клеток снижена [79]. Изменение профиля метилирования рДНК, наблюдаемое при болезни Альцгеймера, приводит к снижению синтеза рРНК и уменьшению количества рибосом [80, 81]. Снижение метилирования ДНК в области промотора и 5'-ETS рибосомных генов выявлено у пациентов с пограничными расстройствами [82]. Сообщается также о существовании связи между гиперметилированными промоторами генов рРНК и сниженной экспрессией рРНК, соответственно, и склонностью к суициду, ассоциированному с жестоким обращением в детстве [83]. Нестабильность рДНК характерна и для других нейродегенеративных заболеваний, при которых профиль метилирования промоторов рибосомных генов не отличается от профиля в здоровых клетках [84]. Возрастные изменения профиля метилирования CpG-островков рибосомных генов предлагается использовать в качестве маркера старения [47].

Особое значение имеет тонкая настройка экспрессии рДНК в головном мозге млекопитающих. Транскрипция рДНК, как первый этап синтеза рибосом, может изменяться под воздействием стимулов окружающей среды, что лежит в основе синаптической пластичности [85]. Во время обучения и запоминания информации в ядрах гиппокампа происходит увеличение синтеза рРНК *de novo* и формирование пула рибосом, необходимого, по-видимому, для синтеза большого количества новых белков, участвующих в формировании нейронных связей. Если при этом специфически ингибировать активность РНК-полимеразы I, то произойдет нарушение формирования долговременной памяти, но не обучения и запоминания [86]. Регуляция этого процесса, по-видимому, связана с увеличением синтеза поли(ADP-рибозы) в ядрышках клеток гиппокампа ферментом PARP1, который активируется киназами PKA и ERK, отвечающими за долговременную синаптическую пластичность [87].

Некоторые симптомы тяжелых наследственных заболеваний, вызванных хромосомными перестройками и геномными мутациями, связывают с изменением числа копий рибосомных генов. Так, при синдроме Дауна отмечено повышенное содержание активных рибосомных генов, причем у новорожденных оно варьирует в более широких пределах, чем у взрослых индивидов [88]. Тем не менее, увеличение транскрипционной активности этого локуса может быть обусловлено общим повышением копийности рДНК из-за наличия дополнительной хромосомы 21, поскольку на ее коротких плечах расположены повторы рибосомных генов. Недавно обнаружили механизм эпигенетической компенсации трисомии: у взрослых индивидов с синдромом Дауна синтез рРНК снижен за счет гиперметилирования промоторов рДНК [89].

У индивидов без хромосомных аномалий значения копийности рибосомных генов, отличные от средних, ассоциированы с другими заболеваниями. Так, число копий рДНК в клетках крови больных шизофренией выше, чем в контрольной группе [90, 91]. Здесь необходимо отметить, что копийность рибосомных генов, в отличие от других повторяющихся элементов генома (например, теломерных повторов), не изменяется под воздействием психоэмоционального стресса, в том числе хронического, как при шизофрении [92]. Увеличенное число копий рДНК выявлено также у больных муковисцидозом [93], а сниженное — при ревматоидном артрите [94]. В настоящее время обсуждается взаимосвязь между числом рибосомных генов и патогенезом таких заболеваний, как ревматоидный артрит и шизофрения, а также генетические предпосылки обеих патологий, способные объяснить феномен их обратной коморбидности (дистропии) [95, 96].

Интересно, что число копий рибосомных генов индивидуально для каждого человека, однако в отдельно взятых популяциях среднее значение остается постоянным, а пределы вариации сужены в группе лиц старческого возраста по сравнению с группой более молодых индивидов. Предполагается, что малое число копий рДНК, как и слишком большое количество рибосомных генов, может быть ассоциировано с более короткой продолжительностью жизни [97]. Сниженную копиюность рибосомных генов предлагается использовать в качестве маркера прогноза когнитивных нарушений в пожилом возрасте [98].

Профиль экспрессии молекул, отвечающих за регуляцию биосинтеза рРНК, может изменяться как в течение жизни клетки, так и при развитии заболеваний. Например, отмечена повышенная экспрессия антисмысловой РНК PAPAS, участвующей в ремоделировании хроматина рДНК, в стареющих клетках и сниженная — в опухолевых клетках [56]. Обнаружена кольцевая РНК, способная регулировать биогенез рибосом на стадии контроля процессинга рРНК, профиль экспрессии которой изменяется при атеросклерозе [99].

Заболевания, связанные с нарушением синтеза рибосом в результате повреждения генов, кодирующих белки, отвечающие за контроль синтеза рРНК и сборку рибонуклеопротеиновых комплексов в ходе созревания рибосом, носят название рибосомопатий [100, 101]. Так, мутация в гене *TCOF1*, кодирующем кофактор РНК-полимеразы I, вызывает синдром Тричера—Коллинза (TCS), при котором наблюдается нарушение формирования тканей лицевой части черепа [102]. Снижение уровня транскрипционных факторов РНК-полимеразы I и, как следствие, уменьшение синтеза 45S рРНК показано при детском церебральном параличе [103]. В клетках, моделирующих болезнь Хантингтона, повышена экспрессия гистон-метилтрансферазы ESET, которая осуществляет триметилирование транскрипционного фактора UBF. При этом наблюдается сайленсинг рибосомных генов [104]. Аномально высокий уровень экспрессии самого UBF и повышенный синтез рРНК обнаружен в клетках головного мозга эмбрионов с недостатком фолиевой кислоты, что приводит к неполному закрытию нервной трубки в эмбриогенезе и, как следствие, к развитию такого порока, как расщепление позвоночника (*spina bifida*) [105]. В целом, сбои на каждом из этапов биогенеза рибосом — от транскрипции рДНК до образования иницирующего трансляцию комплекса — приводят к развитию нейропатологий на любой стадии развития и рассматриваются как один из ключевых факторов нейродегенеративных заболеваний [106].

ЭКСПРЕССИЯ РИБОСОМНЫХ ГЕНОВ В ОПУХОЛЕВЫХ КЛЕТКАХ

Наблюдение за изменениями экспрессии рДНК, структурой рибосомных генов и морфологией ядрышка в опухолевых клетках заслуживает отдельного внимания, поскольку они могут служить диагностическими маркерами злокачественных образований [107]. Известно, что при злокачественном перерождении клеток число копий рДНК нестабильно (может как возрастать, так и уменьшаться) [108]. Таким образом, определение копиюности рДНК может использоваться для прогноза возникновения злокачественных новообразований: обнаружена корреляция между вариациями числа копий рДНК и риском развития рака легкого у курильщиков [7]. Показано также повышение экспрессии рибосомных генов при одновременном снижении числа их копий в активно делящихся опухолевых клетках [109]. Высокий уровень продукции рРНК, наблюдаемый при относительном уменьшении количества рибосомных генов, может объясняться повышенной активностью РНК-полимеразы I, а также тем, что изменения затрагивают процессинг рРНК в опухолях [110, 111]. Кроме того, в опухолевых клетках часто встречается повышенная экспрессия транскрипционных факторов, управляющих синтезом других необходимых для биогенеза рибосом белков и регулирующих таким образом пролиферативную активность клеток [112]. Транскрипционные факторы, в свою очередь, подвергаются регуляции путем ковалентных модификаций, которые осуществляются белками-продуктами онкогенов и генов-супрессоров опухолевого роста, поэтому изменение активности последних при злокачественном перерождении клеток не может не отражаться на синтезе рРНК. Например, pRB — онкосупрессор, экспрессия которого снижена во многих опухолях, подавляет транскрипцию рибосомных генов, препятствуя формированию преинициаторного комплекса на промоторе рДНК [113]. Напротив, продукт онкогена c-Myc, чья активность в опухолевых клетках обычно повышена, стабилизирует преинициаторный комплекс и стимулирует транскрипцию рибосомных генов [114]. Поскольку транскрипция рДНК может регулироваться и эпигенетически, альтернативное метилирование CpG-островков и модификация гистонов в этой области в опухолевых клетках ассоциированы с увеличением продукции рРНК [115]. В свете того, что высокий уровень пролиферации опухолей связан с увеличением биосинтеза белка в клетках, ядрышко и ассоциированные с ним молекулы, включая рибосомные и регуляторные белки, рДНК и рРНК, рассматриваются в качестве многообещающих потенциальных мишеней в направленной противоопухолевой терапии [3]. Опухолевые клеточные линии, используемые в

лабораториях как для изучения отдельных аспектов канцерогенеза, так и в фундаментальных исследованиях, могут иметь отличную от нормы копийность рибосомных генов, профиль метилирования, активность экспрессии и т.д., что необходимо учитывать при выборе той или иной культуры в качестве модельного объекта.

K562 — первая иммортализованная клеточная линия хронического миелоидного лейкоза — содержит так называемую филаделфийскую хромосому и другие хромосомные перестройки, затрагивающие, в частности, хромосому 17 [116]. Промоторная область рДНК в этой клеточной линии обогащена модификациями гистонов, характерными для активно транскрибируемых областей генома, в ней также показан синтез промоторной РНК [43]. В клетках K562 снижен уровень метилирования ДНК в области промотора рибосомных генов и в области, кодирующей 28S, что обычно связано с повышенной транскрипционной активностью [117].

Из клеток аденокарциномы альвеолярного базального эпителия человека получена клеточная линия A549 [118], которая широко используется в качестве модельного объекта для изучения рака легкого, разработки терапевтических подходов и испытания противоопухолевых препаратов. Данные секвенирования и иммунопреципитации хроматина, а также результаты CAGE-анализа свидетельствуют о том, что область 28–32 т.п.н. рМГС и предпромоторная область консервативны во многих линиях опухолевых клеток человека, в том числе, в линиях K562 и A549. Эти области транскрибируются с образованием нкРНК, получивших названия IGS₂₈RNA и промоторная РНК [65].

MCF7 (Michigan Cancer Foundation-7) — клеточная линия рака молочной железы [119], служит модельным объектом для изучения чувствительности и резистентности опухолевых клеток к эстрогенам. Благодаря нестабильности генома (вариативности числа хромосом), клетки MCF7 обладают высокой способностью к адаптации и эволюции [120]. Показано, что в линиях опухолевых клеток молочной железы, в том числе в MCF7, ядерный белок BRCA1 ассоциирован с промоторными областями рибосомных генов, а также со спейсерными промоторами, расположенными в рМГС, и, стимулируя синтез рРНК, принимает участие в регуляции активности РНК-полимеразы I [121].

Клеточная линия PC3 (Prostate cancer 3), полученная из клеток метастаз рака предстательной железы IV степени злокачественности в костную ткань [122], характеризуется повышенной активностью РНК-полимеразы I, обусловленной, по-видимому, увеличением экспрессии онкогена *MYC*, принимающего участие в транскрипции

рибосомных генов [123]. В клетках PC3 повышена экспрессия днРНК, которая, как предполагается, ассоциирована с путями процессинга рРНК и биогенеза рибосом [124].

Изучение нарушения механизмов синтеза рибосом, а также изменения доменов ядерного генома, вовлеченных в этот процесс, особенно рДНК, является многообещающим подходом к анализу и диагностике таких злокачественных новообразований, как глиобластома. Показано, что уменьшение количества копий рибосомных генов в глиобластомах, как и в других видах опухолей, и изменение нуклеотидной последовательности рДНК коррелируют с повышенной пролиферацией клеток и высоким уровнем продукции рРНК и рибосом [15]. С одной стороны, это может происходить из-за увеличения активности РНК-полимеразы I, что наблюдается в первичных культурах глиобластомы при пассировании [110], а с другой, из-за повышения экспрессии рРНК в результате изменения профиля метилирования рДНК [125].

ЗАКЛЮЧЕНИЕ

Рибосомные гены не только кодируют рРНК, они участвуют также в образовании гетерохроматина в процессе дифференцировки и старения клеток, отвечают за ее судьбу в условиях стресса [126]. Эти последовательности, включающие несколько миллионов пар нуклеотидов, кодируют множество нкРНК, огромное количество ферментов и белков, формируют ядрышко — сложную динамическую структуру внутри ядра [127]. Нарушение транскрипции и процессинга рРНК, образования гетерохроматина, строения и функционирования ядрышка так или иначе отражается на работе всех систем органов, проявляясь в виде тяжелейших заболеваний нервной системы или возникновения опухолей [100, 101].

нкРНК, транскрибируемые в области спейсера рибосомных генов, открыты сравнительно недавно, но уже совершенно ясно, что изучение этих молекул имеет огромное значение. В области рМГС синтезируются как смысловые, так и антисмысловые днРНК [77]. При этом генерация длинных антисмысловых транскриптов усиливается в стрессовых ситуациях [128]. Взаимодействуя с различными белками посредством специфических последовательностей или вторичных структур, эти молекулы привлекают их к определенным областям рДНК, осуществляя тонкую настройку как транскрипции, так и временной или постоянной упаковки этих участков генома в неактивный гетерохроматин [69].

Исследование транскриптов, синтезируемых в области рМГС человека, и определение их роли в поддержании физиологического состояния кле-

ток и в развитии патологических процессов позволит уточнить и дополнить знания о процессах регуляции экспрессии генов, формирования гетерохроматина и дифференцировки клеток. Принимая во внимание важность постоянства структуры и функционирования рДНК для нормального существования клетки и всего организма, а также тот факт, что последовательность рДНК в опухолевых клетках часто изменена или экспрессия рРНК отличается от экспрессии в здоровых клетках, представляется особенно важным обратиться к изучению свойств транскриптов, происходящих из рМГС — некодирующей области рибосомных генов.

Данная работа финансирована Министерством науки и высшего образования Российской Федерации, грант № 075-15-2020-809 (13.1902.21.0030).

Статья не содержит каких-либо исследований с участием людей или животных в качестве объектов исследований.

Авторы заявляют об отсутствии конфликта интересов.

СПИСОК ЛИТЕРАТУРЫ

- McStay B. (2016) Nucleolar organizer regions: genomic 'dark matter' requiring illumination. *Genes Dev.* **30**(14), 1598–1610.
- Paredes S., Maggert K. (2009) Ribosomal DNA contributes to global chromatin regulation. *Proc. Natl. Acad. Sci. USA.* **106**(42), 17829–17834.
- Stępiński D. (2018) The nucleolus, an ally, and an enemy of cancer cells. *Histochem. Cell Biol.* **150**(6), 607–629.
- Rajesh Y., Pal I., Banik P., Chakraborty S., Borkar S., Dey G., Mukherjee A., Mandal M. (2017) Insights into molecular therapy of glioma: current challenges and next generation blueprint. *Acta Pharmacol. Sinica.* **38**(5), 591–613.
- Moore L., Kivinen V., Liu Y., Annala M., Cogdell D., Liu X., Liu C., Sawaya R., Yli-Harja O., Shmulevich I., Fuller G.N., Zhang W., Nykter M. (2013) Transcriptome and small RNA deep sequencing reveals deregulation of miRNA biogenesis in human glioma. *J. Pathol.* **229**(3), 449–459.
- Brower J., Clark P., Lyon W., Kuo J. (2014) MicroRNAs in cancer: glioblastoma and glioblastoma cancer stem cells. *Neurochem. Internat.* **77**, 68–77.
- Hosgood H., Hu W., Rothman N., Klugman M., Weinstein S., Virtamo J., Albanes D., Cawthon R., Lan Q. (2019) Variation in ribosomal DNA copy number is associated with lung cancer risk in a prospective cohort study. *Carcinogenesis.* **40**(8), 975–978.
- Santoro R., Grummt I. (2005) Epigenetic mechanism of rRNA gene silencing: temporal order of NoRC-mediated histone modification, chromatin remodeling, and DNA methylation. *Mol. Cell. Biol.* **25**(7), 2539–2546.
- Savić N., Bär D., Leone S., Frommel S., Weber F., Vollenweider E., Ferrari E., Ziegler U., Kaech A., Shakhova O., Cinelli P., Santoro R. (2014) LncRNA maturation to initiate heterochromatin formation in the nucleolus is required for exit from pluripotency in eSCs. *Cell Stem Cell.* **15**(6), 720–734.
- McConkey E., Hopkins J. (1964) The relationship of the nucleolus to the synthesis of ribosomal RNA in HeLa cells. *Proc. Natl. Acad. Sci. USA.* **51**(6), 1197–1204.
- Schwarzacher H., Wachtler F. (1993) The nucleolus. *Anat. Embryol.* **188**(6), 515–536.
- Gonzalez I., Sylvester J. (1995) Complete sequence of the 43-kb human ribosomal DNA repeat: analysis of the intergenic spacer. *Genomics.* **27**(2), 320–328.
- Gibbons J., Branco A., Godinho S., Yu S., Lemos B. (2015) Concerted copy number variation balances ribosomal DNA dosage in human and mouse genomes. *Proc. Natl. Acad. Sci. USA.* **112**(8), 2485–2490.
- Hall A., Turner T., Queitsch C. (2021) Thousands of high-quality sequencing samples fail to show meaningful correlation between 5S and 45S ribosomal DNA arrays in humans. *Sci. Rep.* **11**(1), 449.
- Wang M., Lemos B. (2017) Ribosomal DNA copy number amplification and loss in human cancers is linked to tumor genetic context, nucleolus activity, and proliferation. *PLoS Genet.* **13**(9), e1006994.
- Smirnov E., Chmúrčíková N., Liška F., Bažantová P., Cmarko D. (2021) Variability of human rDNA. *Cells.* **10**(2), 196.
- Erickson J., Schmickel R. (1985) A molecular basis for discrete size variation in human ribosomal DNA. *Am. J. Hum. Genet.* **37**(2), 311–325.
- Akamatsu Y., Kobayashi T. (2015) The human RNA polymerase I transcription terminator complex acts as a replication fork barrier that coordinates the progress of replication with rRNA transcription activity. *Mol. Cell. Biol.* **35**(10), 1871–1881.
- Jacob M., Audas T., Mullineux S., Lee S. (2012) Where no RNA polymerase has gone before: novel functional transcripts derived from the ribosomal intergenic spacer. *Nucleus.* **3**(4), 315–319.
- Guett C., Santoro R. (2012) Formation of nuclear heterochromatin: the nucleolar point of view. *Epigenetics.* **7**(8), 811–814.
- van Koningsbruggen S., Gierlinski M., Schofield P., Martin D., Barton G., Ariyurek Y., den Dunnen J., Lamond A. (2010) High-resolution whole-genome sequencing reveals that specific chromatin domains from most human chromosomes associate with nucleoli. *Mol. Biol. Cell.* **21**(21), 3735–3748.
- Németh A., Conesa A., Santoyo-Lopez J., Medina I., Montaner D., Péterfia B., Solovei I., Cremer T., Dopazo J., Längst G. (2010) Initial genomics of the human nucleolus. *PLoS Genet.* **6**(3), e1000889.
- Yu S., Lemos B. (2018) The long-range interaction map of ribosomal DNA arrays. *PLoS Genet.* **14**(3), e1007258.
- Strohner R., Németh A., Jansa P., Hofmann-Rohrer U., Santoro R., Längst G., Grummt I. (2001) NoRC — a novel member of mammalian ISWI-containing chromatin remodeling machines. *EMBO J.* **20**(17), 4892–4900.
- Guett C., Lienemann P., Sirri V., Grummt I., Hernandez-Verdun D., Hottiger M., Fussenegger M., Santoro R. (2010) The NoRC complex mediates the heterochromatin formation and stability of silent rRNA genes and centromeric repeats. *EMBO J.* **29**(13), 2135–2146.

26. Postepska-Igielska A., Kronic D., Schmitt N., Greulich-Bode K., Boukamp P., Grummt I. (2013) The chromatin remodelling complex NoRC safeguards genome stability by heterochromatin formation at telomeres and centromeres. *EMBO Rep.* **14**(8), 704–710.
27. Guetg C., Scheifele F., Rosenthal F., Hottiger M., Santoro R. (2012) Inheritance of silent rDNA chromatin is mediated by PARP1 via noncoding RNA. *Mol. Cell.* **45**(6), 790–800.
28. Krishnakumar R., Kraus W. (2010) PARP-1 regulates chromatin structure and transcription through a KDM5B-dependent pathway. *Mol. Cell.* **39**(5), 736–749.
29. Roper S., Chrysanthou, S., Senner C., Sienerth A., Gnan S., Murray A., Masutani M., Latos P., Hemberger M. (2014) ADP-ribosyltransferases Parp1 and Parp7 safeguard pluripotency of ES cells. *Nucl. Acids Res.* **42**(14), 8914–8927.
30. Schreiber V., Dantzer F., Amé J.C., de Murcia G. (2006) Poly(ADP-ribose): novel functions for an old molecule. *Nat. Rev.* **7**(7), 517–528.
31. Nalabothula N., Al-jumaily T., Eteleeb A., Flight R., Xiaorong S., Moseley H., Rouchka E., Fondufe-Mittendorf Y. (2015) Genome-wide profiling of PARP1 reveals an interplay with gene regulatory regions and DNA methylation. *PLoS One.* **10**(8), e0135410.
32. Mansuroglu Z., Benhelli-Mokrani H., Marcato V., Sultan A., Violet M., Chauderlier A., Delattre L., Loyens A., Talahari S., Bégard S., Nesslany F., Colin M., Souès S., Lefebvre B., Buée L., Galas M.C., Bonnefoy E. (2016) Loss of Tau protein affects the structure, transcription and repair of neuronal pericentromeric heterochromatin. *Sci. Rep.* **6**, 33047.
33. Maina M., Bailey L., Wagih S., Biasetti L., Pollack S., Quinn J., Thorpe J., Doherty A., Serpell L. (2018) The involvement of Tau in nucleolar transcription and the stress response. *Acta Neuropathol. Commun.* **6**(1), 70.
34. Kuhn A., Grummt I. (1992) Dual role of the nucleolar transcription factor UBF: trans-activator and antirepressor. *Proc. Natl. Acad. Sci. USA.* **89**(16), 7340–7344.
35. Sanij E., Poortinga G., Sharkey K., Hung S., Holloway T., Quin J., Robb E., Wong L., Thomas W., Stefanovsky V., Moss T., Rothblum L., Hannan K.M., McArthur G.A., Pearson R.B., Hannan R.D. (2008) UBF levels determine the number of active ribosomal RNA genes in mammals. *J. Cell. Biol.* **183**(7), 1259–1274.
36. Hannan K., Hannan R., Rothblum L. (1998) Transcription by RNA polymerase I. *Front. Biosci.* **3**, d376–d398.
37. Grummt I. (1999) Regulation of mammalian ribosomal gene transcription by RNA polymerase I. *Prog. Nucl. Acids Res. Mol. Biol.* **62**, 109–154.
38. Grummt I. (2010) Wisely chosen paths – regulation of rRNA synthesis. *FEBS J.* **277**(22), 4626–4639.
39. Tanaka Y., Tsuneoka M. (2018) Control of ribosomal RNA transcription by nutrients. In: *Gene Expression and Regulation in Mammalian Cells – Transcription Toward the Establishment of Novel Therapeutics*. Ed. Uchiumi F. London: IntechOpen. Ch. 2.
40. Matthews D., Olson M. (2006) What is new in the nucleolus?: workshop on the nucleolus: new perspectives. *EMBO Rep.* **7**(9), 870–873.
41. Zhao Z., Senturk N., Song C., Grummt I. (2018) IncRNA PAPAS tethered to the rDNA enhancer recruits hypophosphorylated CHD4/NuRD to repress rRNA synthesis at elevated temperatures. *Genes Dev.* **32**(11–12), 836–848.
42. Grummt I., Ladurner A. (2008) A metabolic throttle regulates the epigenetic state of rDNA. *Cell.* **133**(4), 577–580.
43. Zentner G., Saiakhova A., Manaenkov P., Adams M., Scacheri P. (2011) Integrative genomic analysis of human ribosomal DNA. *Nucl. Acids Res.* **39**(12), 4949–4960.
44. Srivastava R., Srivastava R., Ahn S. (2016) The epigenetic pathways to ribosomal DNA silencing. *Microbiol. Mol. Biol. Rev.* **80**(3), 545–563.
45. Li J., Längst G., Grummt I. (2006) NoRC-dependent nucleosome positioning silences rRNA genes. *EMBO J.* **25**(24), 5735–5741.
46. Li J., Santoro R., Koberna K., Grummt I. (2005) The chromatin remodeling complex NoRC controls replication timing of rRNA genes. *EMBO J.* **24**(1), 120–127.
47. Wang M., Lemos B. (2019) Ribosomal DNA harbors an evolutionarily conserved clock of biological aging. *Genome Res.* **29**(3), 325–333.
48. Németh A., Guibert S., Tiwari V., Ohlsson R., Längst G. (2008) Epigenetic regulation of TTF-I-mediated promoter–terminator interactions of rRNA genes. *EMBO J.* **27**(8), 1255–1265.
49. Xie W., Ling T., Zhou Y., Feng W., Zhu Q., Stunnenberg H., Grummt I., Tao W. (2012) The chromatin remodeling complex NuRD establishes the poised state of rRNA genes characterized by bivalent histone modifications and altered nucleosome positions. *Proc. Natl. Acad. Sci. USA.* **109**(21), 8161–8166.
50. Yuan X., Feng W., Imhof A., Grummt I., Zhou Y. (2007) Activation of RNA polymerase I transcription by cockayne syndrome group B protein and histone methyltransferase G9a. *Mol. Cell.* **27**(4), 585–595.
51. Salifou K., Ray S., Verrier L., Aguirrebengoa M., Trouche D., Panov K., Vandromme M. (2016) The histone demethylase JMJD2A/KDM4A links ribosomal RNA transcription to nutrients and growth factors availability. *Nat. Commun.* **7**, 10174.
52. Murayama A., Ohmori K., Fujimura A., Minami H., Yasuzawa-Tanaka K., Kuroda T., Oie S., Daitoku H., Okuwaki M., Nagata K., Fukamizu A., Kimura K., Shimizu T., Yanagisawa J. (2008) Epigenetic control of rDNA loci in response to intracellular energy status. *Cell.* **133**(4), 627–639.
53. Kumazawa T., Nishimura K., Kuroda T., Ono W., Yamaguchi C., Katagiri N., Tsuchiya M., Masumoto H., Nakajima Y., Murayama A., Kimura K., Yanagisawa J. (2011) Novel nucleolar pathway connecting intracellular energy status with p53 activation. *J. Biol. Chem.* **286**(23), 20861–20869.
54. Yan Q., Zhu C., Guang S., Feng X. (2019) The functions of non-coding RNAs in rRNA regulation. *Front. Genet.* **10**, 290.
55. Jacob M., Audas T., Uniacke J., Trinkle-Mulcahy L., Lee S. (2013) Environmental cues induce a long non-coding RNA-dependent remodeling of the nucleolus. *Mol. Biol. Cell.* **24**(18), 2943–2953.
56. Bierhoff H., Schmitz K., Maass F., Ye J., Grummt I. (2010) Noncoding transcripts in sense and antisense

- orientation regulate the epigenetic state of ribosomal RNA genes. *Cold Spring Harb. Symp. Quant. Biol.* **75**, 357–364.
57. Schmitz K., Mayer C., Postepska A., Grummt I. (2010) Interaction of noncoding RNA with the rDNA promoter mediates recruitment of DNMT3b and silencing of rRNA genes. *Genes Dev.* **24**(20), 2264–2269.
 58. Mayer C., Neubert M., Grummt I. (2008) The structure of NoRC-associated RNA is crucial for targeting the chromatin remodelling complex NoRC to the nucleolus. *EMBO Rep.* **9**(8), 774–780.
 59. Bierhoff H., Dammert M., Brocks D., Dambacher S., Schotta G., Grummt I. (2014) Quiescence-induced LncRNAs trigger H4K20 trimethylation and transcriptional silencing. *Mol. Cell.* **54**(4), 675–682.
 60. Li D., Zhang J., Wang M., Li X., Gong H., Tang H., Chen L., Wan L., Liu Q. (2018) Activity dependent LoNA regulates translation by coordinating rRNA transcription and methylation. *Nat. Commun.* **9**(1), 1726–1739.
 61. Caudron-Herger M., Pankert T., Seiler J., Németh A., Voit R., Grummt I., Rippe K. (2015) Alu element-containing RNAs maintain nucleolar structure and function. *EMBO J.* **34**(22), 2758–2774.
 62. Xing Y., Yao R., Zhang Y., Guo C., Jiang S., Xu G., Dong R., Yang L., Chen L. (2017) SLERT regulates DDX21 rings associated with Pol I transcription. *Cell.* **169**(4), 664–678.
 63. Morgan G., Reeder R., Bakken A. (1983) Transcription in cloned spacers of *Xenopus laevis* ribosomal DNA. *Proc. Natl. Acad. Sci. USA.* **80**(21), 6490–6494.
 64. Kuhn A., Grummt I. (1987) A novel promoter in the mouse rDNA spacer is active *in vivo* and *in vitro*. *EMBO J.* **6**(11), 3487–3492.
 65. Agrawa S., Ganley A. (2018) The conservation landscape of the human ribosomal RNA gene repeats. *PLoS One.* **13**(12), e0207531.
 66. Li Y., Wang H., Wan F., Liu F., Liu J., Zhang N., Jin S., Li J. (2012) Deep sequencing analysis of small non-coding RNAs reveals the diversity of microRNAs and piRNAs in the human epididymis. *Gene.* **497**(2), 330–335.
 67. Ma X., Liu H., Zheng Y., Dai Y., Lingling E., Zhang R., Zhang S. (2022) Genome-wide screening of different expressed genes and its potential associations with aging dental pulp stem cells. *Comb. Chem. High Throughput Screen.* <https://doi.org/10.2174/1386207325666220705120904>
 68. Pirogov S., Gvozdev V., Klenov M. (2019) Long non-coding RNAs and stress response in the nucleolus. *Cells.* **8**(7), 668.
 69. Mayer C., Schmitz K., Li J., Grummt I., Santoro R. (2006) Intergenic transcripts regulate the epigenetic state of rRNA genes. *Mol. Cell.* **22**(3), 351–361.
 70. McStay B., Grummt I. (2008) The epigenetics of rRNA genes: from molecular to chromosome biology. *Annu. Rev. Cell Dev. Biol.* **24**, 131–157.
 71. Mars J.C., Sabourin-Felix M., Tremblay M., Moss T. (2017) A deconvolution protocol for ChIP-seq reveals analogous enhancer structures on the mouse and human ribosomal RNA genes. *G3.* **8**(1), 303–314.
 72. Shiao Y., Lupascu S., Gu Y., Kasprzak W., Hwang C., Fields J., Leighty R., Quiñones O., Shapiro B., Alvord W., Anderson L. (2009) An intergenic non-coding RNA correlated with expression of the rRNA and frequency of an rRNA single nucleotide polymorphism in lung cancer cells. *PLoS One.* **4**(10), e7505.
 73. Vacík T., Kerečiče S., Raška I., Cmarko D., Smirnov E. (2019) Life time of some RNA products of rDNA in intergenic spacer in HeLa cells. *Histochem. Cell. Biol.* **152**(4), 271–280.
 74. Todd M., Huh M., Picketts D. (2016) The sub-nucleolar localization of PHF6 defines its role in rDNA transcription and early processing events. *Eur. J. Hum. Genet.* **24**(10), 1453–1459.
 75. Sadova A., Kupriyanova N., Pavlova G. (2020) Mapping and quantification of non-coding RNA originating from the rDNA in human glioma cells. *Cancers.* **12**(8), 2090.
 76. Sadova A., Pantelev D., Pavlova G. (2021) Zooming in: PAGE-northern blot helps to analyze anti-sense transcripts originating from human rIGS under transcriptional stress. *Noncoding RNA.* **7**(3), 50.
 77. Abraham K., Khosraviani N., Chan J., Gorthi A., Samman A., Zhao D.Y., Wang M., Bokros M., Vidya E., Ostrowski L.A., Oshidari R., Pietrobon V., Patel P.S., Algouneh A., Singhania R., Liu Y., Yerlici V.T., De Carvalho D.D., Ohh M., Dickson B.C., Hakem R., Greenblatt J.F., Lee S., Bishop A.J.R., Mekhail K. (2020) Nucleolar RNA polymerase II drives ribosome biogenesis. *Nature.* **585**(7824), 298–302.
 78. Warmerdam D., Wolthuis R. (2019) Keeping ribosomal DNA intact: a repeating challenge. *Chromosome Res.* **27**(1–2), 57–72.
 79. Machwe A., Orren D.K., Bohr V.A. (2000) Accelerated methylation of ribosomal RNA genes during the cellular senescence of Werner syndrome fibroblasts. *FASEB J.* **14**(12), 1715–1724.
 80. Zeng J., Libien J., Shaik F., Wolk J., Hernández A. (2016) Nucleolar PARP-1 expression is decreased in Alzheimer's disease: consequences for epigenetic regulation of rDNA and cognition. *Neural. Plasticity.* **2016**, 8987928.
 81. Pietrzak M., Rempala G., Nelson P., Zheng J., Hetman M. (2011) Epigenetic silencing of nucleolar rRNA genes in Alzheimer's disease. *PLoS One.* **6**(7), e22585.
 82. Teschler S., Gotthardt J., Dammann G., Dammann R. (2016) Aberrant DNA methylation of rDNA and PRIMA1 in borderline personality disorder. *Int. J. Mol. Sci.* **17**(1), E67.
 83. McGowan P., Sasaki A., Huang T., Unterberger A., Suderman M., Ernst C., Meaney M., Turecki G., Szyf M. (2008) Promoter-wide hypermethylation of the ribosomal RNA gene promoter in the suicide brain. *PLoS One.* **3**(5), e2085.
 84. Hallgren J., Pietrzak M., Rempala G., Nelson P., Hetman M. (2014) Neurodegeneration-associated instability of ribosomal DNA. *Biochim. Biophys. Acta.* **1842**(6), 860–868.
 85. Dastidar S., Nair D. (2022) A ribosomal perspective on neuronal local protein synthesis. *Front. Mol. Neurosci.* **15**, 823135.
 86. Allen K., Regier M., Hsieh C., Tsokas P., Barnard M., Phatarpekar S., Wolk J., Sacktor T., Fenton A., Hernández A. (2018) Learning-induced ribosomal RNA is required for memory consolidation in mice: evidence of differentially expressed rRNA variants in learning and memory. *PLoS One.* **13**(10), e020337.

87. Allen K., Gourov A., Harte C., Gao P., Lee C., Sylvain D., Splett J., Oxberry W., van de Nes P., Troy-Regier M., Wolk J., Alarcon J.M., Hernández A.I. (2014) Nucleolar integrity is required for the maintenance of long-term synaptic plasticity. *PLoS One*. **9**(8), e104364.
88. Lyapunova N., Porokhovnik L., Kosyakova N., Mandron I., Tsvetkova T. (2017) Effects of the copy number of ribosomal genes (genes for rRNA) on viability of subjects with chromosomal abnormalities. *Gene*. **611**, 47–53.
89. Ravaioli F., Zampieri M., Morandi L., Pirazzini C., Pellegrini C., De Fanti S., Gensous N., Pirazzoli G., Sambati L., Ghezzi A., Ciccarone F., Reale A., Monti D., Salvioli S., Caiafa P., Capri M., Bürkle A., Moreno-Villanueva M., Garagnani P., Franceschi C., Bacalini M.G. (2022) DNA methylation analysis of ribosomal DNA in adults with Down syndrome. *Front. Genet.* **13**, 792165.
90. Chestkov I.V., Jestkova E.M., Ershova E.S., Golimbet V.E., Lezheiko T.V., Kolesina N.Y., Porokhovnik L.N., Lyapunova N.A., Izhevskaya V.L., Kutsev S.I., Veiko N.N., Kostyuk S.V. (2018) Abundance of ribosomal RNA gene copies in the genomes of schizophrenia patients. *Schizophr. Res.* **197**, 305–314.
91. Ershova E.S., Malinovskaya E.M., Golimbet V.E., Lezheiko T.V., Zakharova N.V., Shmarina G.V., Veiko R.V., Umriukhin P.E., Kostyuk G.P., Kutsev S.I., Izhevskaya V.L., Veiko N.N., Kostyuk S.V. (2020). Copy number variations of satellite III (1q12) and ribosomal repeats in health and schizophrenia. *Schizophr. Res.* **223**, 199–212.
92. Umriukhin P., Ershova E., Filev A., Agafonova O., Martynov A., Zakharova N., Veiko R., Porokhovnik L., Kostyuk G., Kutsev S., Veiko N., Kostyuk S. (2022) The psychoemotional stress-induced changes in the abundance of SatIII (1q12) and telomere repeats, but not ribosomal DNA, in human leukocytes. *Genes (Basel)*. **13**(2), 343.
93. Кондратьева Е., Ершова Е., Воронкова А., Шмари́на Г., Красовский С., Жека́йте Е., Петро́ва Н., Мелья́новская Ю., Одинаева Н., Вейко Н., Костюк С. (2021) Вариация числа копий рибосомных генов в геномах больных муковисцидозом. *Мед. генет.* **20**(2), 49–60.
94. Вейко Н., Шубаева Н., Цветкова Т., Мандрон И., Малиновская Т., Сперанский А., Ляпунова Н. (2005) Особенности количественных характеристик комплекса рибосомных генов у пациентов с тяжелыми формами ревматоидного артрита. *Мед. генет.* **4**(4), 74.
95. Zamanpoor M., Ghaedi H., Omrani M. (2020) The genetic basis for the inverse relationship between rheumatoid arthritis and schizophrenia. *Mol. Genet. Genom. Med.* **8**(11), e1483.
96. Porokhovnik L., Lyapunova N. (2019) Dosage effects of human ribosomal genes (rDNA) in health and disease. *Chromosome Res.: Internat. J. Mol., Supramol. Evol. Aspects Chromosome Biol.* **27**(1–2), 5–17.
97. Malinovskaya E., Ershova E., Golimbet V., Porokhovnik L., Lyapunova N., Kutsev S., Veiko N., Kostyuk S. (2018) Copy number of human ribosomal genes with aging: unchanged. *Front. Genet.* **9**, 306.
98. Veiko N., Ershova E., Veiko R., Umriukhin P., Kurmyshev M., Kostyuk G., Kutsev S., Kostyuk S. (2022) Mild cognitive impairment is associated with low copy number of ribosomal genes in the genomes of elderly people. *Front. Genet.* **13**, 967448.
99. Holdt L., Stahlinger A., Sass K., Pichler G., Kulak N., Wilfert W., Kohlmaier A., Herbst A., Northoff B., Nicolaou A., Gäbel G., Beutner F., Scholz M., Thiery J., Musunuru K., Krohn K., Mann M., Teupser D. (2016) Circular non-coding RNA ANRIL modulates ribosomal RNA maturation and atherosclerosis in humans. *Nat. Commun.* **7**, 12429.
100. Narla A., Ebert B. (2010) Ribosomopathies: human disorders of ribosome dysfunction. *Blood*. **115**(16), 3196–3205.
101. Mills E., Green R. (2017) Ribosomopathies: there's strength in numbers. *Science*. **358**(6363), eaan2755.
102. Calo E., Gu B., Bowen M., Aryan F., Zalc A., Liang J., Flynn R., Swigut T., Chang H., Attardi L., Wysocka J. (2018) Tissue-selective effects of nucleolar stress and rDNA damage in developmental disorders. *Nature*. **554**(7690), 112–117.
103. Von Walden F., Gantelius S., Liu C., Borgström H., Björk L., Gremark O., Stål P., Nader G., Ponté N. (2018) Muscle contractures in patients with cerebral palsy and acquired brain injury are associated with extracellular matrix expansion, pro-inflammatory gene expression, and reduced rRNA synthesis. *Muscle Nerve*. **58**(2), 277–285.
104. Hwang Y., Han D., Kim K., Min S., Kowall N., Yang L., Lee J., Kim Y., Ryu H. (2014) ESET methylates UBF at K232/254 and regulates nucleolar heterochromatin plasticity and rDNA transcription. *Nucl. Acids Res.* **42**(3), 1628–1643.
105. Xie Q., Li C., Song X., Wu L., Jiang Q., Qiu Z., Cao H., Yu K., Wan C., Li J., Yang F., Huang Z., Niu B., Jiang Z., Zhang T. (2017) Folate deficiency facilitates recruitment of upstream binding factor to hot spots of DNA double-strand breaks of rRNA genes and promotes its transcription. *Nucl. Acids Res.* **45**(5), 2472–2489.
106. Hetman M., Slomnicki L. (2019) Ribosomal biogenesis as an emerging target of neurodevelopmental pathologies. *J. Neurochem.* **148**(3), 325–347.
107. Smirnov E., Chmúřciaková N., Cmarko D. (2021) Human rDNA and cancer. *Cells*. **10**(12), 3452.
108. Valori V., Tus K., Laukaitis C., Harris D., LeBeau L., Maggert K. (2019) Human rDNA copy number is unstable in metastatic breast cancers. *Epigenetics*. **15**(1–2), 85–106.
109. Xu B., Li H., Perry J., Singh V., Unruh J., Yu Z., Zakari M., McDowell W., Li L., Gerton J. (2017) Ribosomal DNA copy number loss and sequence variation in cancer. *PLoS Genet.* **13**(6), e1006771.
110. Baskaran S., Mayrhofer M., Kultima H., Bergström T., Elfineh L., Cavellier L., Isaksson A., Nelander S. (2018) Primary glioblastoma cells for precision medicine: a quantitative portrait of genomic (in)stability during the first 30 passages. *Neuro-Oncol.* **20**(8), 1080–1091.
111. Belin S., Beghin A., Solano-González E., Bezin L., Brunet-Manquat S., Textoris J., Prats A., Mertani H., Dumontet C., Diaz J. (2009) Dysregulation of ribosome biogenesis and translational capacity is associated with tumor progression of human breast cancer cells. *PLoS One*. **4**(9), e7147.

112. Rajput P., Shukla S., Kumar V. (2015) The HBx oncoprotein of hepatitis B virus potentiates cell transformation by inducing c-Myc-dependent expression of the RNA polymerase I transcription factor UBF. *Virology*. **12**, 62.
113. Grandori C., Gomez-Roman N., Felton-Edkins Z., Ngouenet C., Galloway D., Eisenman R., White R. (2005) c-Myc binds to human ribosomal DNA and stimulates transcription of rRNA genes by RNA polymerase I. *Nat. Cell. Biol.* **7**(3), 311–318.
114. Hannan K., Hannan R., Smith S., Jefferson L., Lun M., Rothblum L. (2000) Rb and p130 regulate RNA polymerase I transcription: Rb disrupts the interaction between UBF and SL-1. *Oncogene*. **19**(43), 4988–4999.
115. Frescas D., Guardavaccaro D., Bassermann F., Koyama-Nasu R., Pagano M. (2007) JHDM1B/FBXL10 is a nucleolar protein that represses transcription of ribosomal RNA genes. *Nature*. **450**(7167), 309–313.
116. Lozzio C., Lozzio B. (1975) Human chronic myelogenous leukemia cell-line with positive Philadelphia chromosome. *Blood*. **45**(3), 321–334.
117. Foltankova V., Legartova S., Kozubek S., Bartova E. (2012) Tumor-specific histone signature and DNA methylation in multiple myeloma and leukemia cells. *Neoplasma*. **59**(4), 450–462.
118. Giard D., Aaronson S., Todaro G., Arnstein P., Kersey J., Dosik H., Parks W. (1973) *In vitro* cultivation of human tumors: establishment of cell lines derived from a series of solid tumors. *J. Natl. Cancer Inst.* **51**(5), 1417–1423.
119. Soule H., Vazquez J., Long A., Albert S., Brennan M. (1973) A human cell line from a pleural effusion derived from a breast carcinoma. *J. Natl. Cancer Inst.* **51**(5), 1409–1416.
120. Lee A., Oesterreich S., Davidson N. (2015) MCF-7 cells – changing the course of breast cancer research and care for 45 years. *J. Natl. Cancer Inst.* **107**(7), djv073.
121. Johnston R., D'Costa Z., Ray S., Gorski J., Harkin D., Mullan P., Panov K. (2016) The identification of a novel role for BRCA1 in regulating RNA polymerase I transcription. *Oncotarget*. **7**(42), 68097–68110.
122. Kaighn M., Narayan K., Ohnuki Y., Lechner J., Jones L. (1979) Establishment and characterization of a human prostatic carcinoma cell line (PC-3). *Invest. Urol.* **17**(1), 16–23.
123. Zhang D., Park D., Zhong Y., Lu Y., Rycak K., Gong S., Chen X., Liu X., Chao H., Whitney P., Calhoun-Davis T., Takata Y., Shen J., Iyer V.R., Tang, D.G. (2016) Stem cell and neurogenic gene-expression profiles link prostate basal cells to aggressive prostate cancer. *Nat. Commun.* **7**, 10798.
124. Yan Y., Chen Z., Xiao Y., Wang X., Qian K. (2019) Long non-coding RNA SNHG6 is upregulated in prostate cancer and predicts poor prognosis. *Mol. Biol. Rep.* **46**(3), 2771–2778.
125. Holmberg Olausson K., Nister M., Lindstrom M. (2014) Loss of nucleolar histone chaperone NPM1 triggers rearrangement of heterochromatin and synergizes with a deficiency in DNA methyltransferase DNMT3a to drive ribosomal DNA transcription. *J. Biol. Chem.* **289**(50), 34601–34619.
126. Kobayashi T. (2008) A new role of the rDNA and nucleolus in the nucleus – rDNA instability maintains genome integrity. *Bioessays*. **30**(3), 267–272.
127. O'Sullivan J., Pai D., Cridge A., Engelke D., Ganley A. (2013) The nucleolus: a raft adrift in the nuclear sea or the keystone in nuclear structure? *Biomol. Concepts*. **4**(3), 277–286.
128. Audas T., Jacob M., Lee S. (2012) Immobilization of proteins in the nucleolus by ribosomal intergenic spacer noncoding RNA. *Mol. Cell*. **45**(2), 147–157.

Human rDNA Structure, Expression, and Non-Canonical Functions: the Role of Non-Coding Regions

A. A. Sadova^{1, 2, *}, D. Yu. Panteleev¹, and G. V. Pavlova^{1, 3, 4}

¹*Institute of Higher Nervous Activity and Neurophysiology, Russian Academy of Sciences, Moscow, 117485 Russia*

²*Pirogov Russian National Research Medical University, Moscow, 117997 Russia*

³*Burdenko National Medical Research Center for Neurosurgery, Moscow, 125047 Russia*

⁴*First Moscow State Medical University, Moscow, 119435 Russia*

*e-mail: 89652410866@mail.ru

The review is dedicated to analyzing and summarizing the data on the part of human genome encoding 45S rRNA. The sequences which seem evolutionary conserved on the first glance astonish one with their variability in structure and a variety of functions on closer examination. The major part of rDNA is non-coding and contains regulatory elements, protein binding sites, pseudogenes, repetitive sequences, and microRNA genes. Ribosomal intergenic spacers are not only in charge with the nucleolus morphology and functioning, namely, the rRNA expression and ribosome biogenesis, but also control nuclear chromatin formation thus mediating cell differentiation. Besides, alterations in the expression of these non-coding regions of rDNA in response to environmental stimuli underlies the keen sense of cell to various types of stressors. Malfunctioning of this process may result in a wide range of pathologies from oncology to neurodegenerative disease and mental illness. Here we observe to-date materials on the structure and transcription of the ribosomal intergenic spacer in humans and its role in rRNA expression, in-born disease development, and cancer.

Keywords: rDNA, rRNA, pRNA, PAPAS, IGS RNAs

УДК 578.24

КАК ГИСТОНДЕАЦЕТИЛАЗА 3 КОНТРОЛИРУЕТ ЭКСПРЕССИЮ ГЕПСИДИНА И РЕПЛИКАЦИЮ ВИРУСА ГЕПАТИТА С

© 2023 г. А. С. Щербакова^а, С. Н. Кочетков^а, М. В. Козлов^{а, *}^аИнститут молекулярной биологии им. В.А. Энгельгардта Российской академии наук, Москва, 119991 Россия

*e-mail: kozlovmavi@gmail.com

Поступила в редакцию 28.09.2022 г.

После доработки 25.11.2022 г.

Принята к публикации 25.11.2022 г.

Ключевую роль гистондеацетилаз (HDAC) в регуляции клеточного ответа на заражение вирусом гепатита С (HCV) впервые продемонстрировали в 2008 году. Изучая метаболизм железа в тканях печени больных хроническим гепатитом С, обнаружили, что в условиях окислительного стресса, вызванного вирусной инфекцией, в гепатоцитах заметно снижается экспрессия гена гепсидина (HAMP), гормона-регулятора экспорта железа. В схему регуляции экспрессии гепсидина оказались вовлечены HDAC, контролирующие уровень ацетилирования гистонов и транскрипционных факторов, прежде всего STAT3, ассоциированных с промотором *HAMP*. В настоящем обзоре обобщены современные данные о функционировании регуляторного контура HCV-HDAC3-STAT3-HAMP как пример хорошо охарактеризованного взаимодействия вируса и эпигенетического аппарата клетки-хозяина.

Ключевые слова: вирус гепатита С, репликация, STAT3, гепсидин, окислительный стресс, гистондеацетилаза 3, регуляция экспрессии

DOI: 10.31857/S0026898423030096, EDN: CHJNUP

ФОРМИРОВАНИЕ ПЕРВИЧНЫХ ПРЕДСТАВЛЕНИЙ О РЕГУЛЯТОРНОМ КОНТУРЕ HCV-HDAC-STAT3-HAMP

Вирус гепатита С (HCV) провоцирует развитие воспалительных процессов в организме с высвобождением провоспалительных цитокинов, в том числе одного из самых эффективных лейкоцитарных цитокинов — интерлейкина-6 (IL-6), содержание которого в сыворотке больных гепатитом С строго коррелирует с клиренсом вируса [1]. Транскрипционный фактор STAT3 служит сигнальным медиатором для IL-6, поэтому, когда в 2004 году Zhu и соавт. показали, что IL-6 ингибирует репликацию субгеномной РНК HCV в культуре клеток через активацию STAT3 и сигнальный путь ERK, это не стало большой неожиданностью [2]. В той же работе показано, что STAT3 играет важную роль в интерферон-индуцированном противови-

русном ответе. Совокупность полученных результатов убедительно свидетельствовала, что активная форма STAT3 инициирует внутриклеточные противовирусные механизмы.

Сигнальный путь IL-6/STAT3, наряду с BMP/SMAD и эритропоэтином, гликопротеиновым гормоном, стимулирующим выработку эритроцитов, контролирует содержание в крови гепсидина, известного также как LEAP-1 (Liver-Expressed Antimicrobial Peptide 1). Гепсидин, принадлежащий к семейству дефенсинов, обеспечивающих защитное действие врожденного иммунитета, секретируется клетками печени, а именно гепатоцитами, и циркулирует в плазме в комплексе с $\alpha 2$ -макроглобулином, также синтезируемом в гепатоцитах. Клиренс гепсидина происходит либо через почки, либо путем протеасомной кодеградации с ферропортином. Гепсидин — центральный регулятор гомеостаза железа — это пептид длиной 25 аминокислотных остатков, который образуется в результате протеолиза белка-предшественника и имеет структуру, жестко зафиксированную четырьмя внутримолекулярными дисульфидными связями остатков цистеина [3, 4].

В 2006 году Wrighting и соавт., используя технику формальдегидной сшивки с последующей иммунопреципитацией, впервые показали, что STAT3 контактирует с промотором *HAMP*, и это

Сокращения: BMP/SMAD — сигнальный путь, блокирующий экспрессию гепсидина; CHC — хронический гепатит С; ERK — сигнальные киназы семейства митоген-активируемых протеинкиназ (МАРК), регулирующих базовые клеточные процессы; *HAMP* — ген гепсидина, белка семейства дефенсинов; HAT — гистонацетилтрансфераза; HCV — вирус гепатита С; HDAC — гистондеацетилаза; HepG2, Huh7, Huh7.5 — клеточные линии гепатоцеллюлярной карциномы человека; IFN — интерферон; IRES — 5'-концевой фрагмент вирусной РНК, инициирующий кеп-независимую трансляцию; OS — окислительный стресс; ROS — активные формы кислорода; TI — терапевтический индекс.

состояние хроматина, путем каталитического удаления ацетильной группы с остатков ε -N-ацетилизина в молекулах гистонов и других белков-субстратов, что подробнее будет рассмотрено ниже. В частности, в условиях OS выявлено гипoaцетилирование гистона H3 в районе промотора *HAMP* с одновременным блокированием присоединения к промотору транскрипционных факторов C/EBP α и STAT3 и падением экспрессии *HAMP*. Наоборот, в присутствии антиоксидантов содержание мРНК *HAMP* в HCV-инфицированных клетках увеличивалось, как и при воздействии ингибитора трихостатина А (TSA), не избирательно блокирующего активность всех цинкзависимых HDAC (пан-HDAC-ингибирование). В присутствии TSA наблюдалось накопление ацетилированной формы гистона H3 (H3K9ac) и восстановление связывания C/EBP α и STAT3 с промотором гепсидина [9]. К сожалению, высокая токсичность TSA в отношении клеток гепатомы маскировала его антивирусное действие. Только через 5 лет в качестве агента, более селективно, чем TSA, блокирующего репликацию HCV в клеточной системе, был изучен другой мультипотентный (широкого спектра действия) ингибитор HDAC – вориностат (SAHA) [10]. Детально механизмы влияния OS на регуляцию экспрессии гепсидина рассмотрены ниже.

ВЗАИМОСВЯЗЬ СИГНАЛЬНОГО ПУТИ JAK/STAT3 И ПРОЛИФЕРАЦИИ HCV

Как и следовало ожидать, механизм антивирусной активности IL-6 тесно связан с его сигнальной функцией. Известно, что взаимодействие этого цитокина с рецепторными белками IL-6R α и gp130 приводит к образованию гексамерного трансмембранного комплекса (IL-6/IL-6R α /gp130)₂. На следующей стадии к комплексу со стороны цитозоля присоединяются две молекулы киназы JAK1/2, при этом происходит их сближение и трансактивация. Затем они фосфорилируют остатки тирозина в gp130 в составе комплекса, что вызывает присоединение молекул SHP-2 и STAT. Их последующее фосфорилирование киназами JAK1/2 запускает два основных gp130-зависимых сигнальных пути: SHP-2/ERK MAPK-каскад и STAT3 [12]. Чтобы оценить роль STAT3 в антивирусном ответе, Zhu и соавт. применили индуцибельную систему, в которой активация сигнального пути STAT3 происходила независимым от цитокинов и рецепторов образом [2]. Авторы использовали созданную ранее конструкцию слитого белка STAT3ER, в которой кодирующая последовательность STAT3 ассоциирована с последовательностью лиганд-связывающего домена рецептора эстрогена ER-LBD [13]. Такая конструкция позволяла активировать сигнальную функцию STAT3ER без участия киназ, но путем добавления синтети-

ческого лиганда 4-гидрокситамоксифена (4-НТ), после чего белок переставал регулироваться цитокинами и факторами роста [14].

Данный подход реализован на клетках линии FCA1, поддерживающих размножение репликаона HCV (родительская линия клеток Huh7) [15]. После трансфекции и отбора по устойчивости к пуромицину получили стабильную линию клеток FCA1-STAT3ER, экспрессирующую соответствующий слитый белок. Добавление 4-НТ к этим клеткам приводило к подавлению репликации вируса, тогда как в клетках исходной линии FCA1 уровень вирусной РНК оставался неизменным [2]. Работоспособность белка STAT3ER подтверждалась его димеризацией, необходимой для активации в ответ на добавление 4-НТ, а также повышением активности репортерной люциферазы, экспрессия которой находилась под контролем STAT3-связывающего мотива репортерной плазмиды. Кроме того, в клетках линии FCA1-STAT3ER наблюдалась прямая корреляция между STAT3-зависимой экспрессией мембранного IFN-чувствительного белка IFITM3/1-8U и антивирусным действием STAT3. Обнаруженная корреляция имела характер причинно-следственной связи, поскольку ранее показали, что сверхэкспрессия белка IFITM3/1-8U в клетках линии FCA1, трансфицированных соответствующим вектором, приводила к двукратному снижению уровня РНК HCV в сравнении с контролем [15].

В настоящее время описано несколько молекулярных механизмов антивирусного действия IFITM3/1-8U, наиболее важные из которых: (i) ингибирование IRES-опосредованной трансляции HCV [16]; (ii) нарушение клеточного гомеостаза холестерина, эфиры которого входят в состав липидных капель, формирующих вирусные реплизомы (рис. 1) [17]; и (iii) блокирование проникновения вируса в клетку [18]. Нельзя также исключить возможность ингибирования репликации HCV путем ранее установленного взаимодействия IFITM3/1-8U и трансмембранного белка VAPA [17, 19]. Дело в том, что гетеродимер двух родственных белков VAPA и VAPB играет важную роль в пролиферации HCV, принимая участие в позиционировании вирусного комплекса репликации на мембране липидного рафта клетки-хозяина за счет взаимодействия с неструктурными белками NS5A и NS5B [11, 20]. Следовательно, повышение экспрессии IFITM3/1-8U гипотетически может иметь следствием вытеснение VAPA из состава комплекса репликации на поверхность двухмембранных везикул реплизомы (рис. 1). Это тем более возможно, поскольку и IFITM3/1-8U, и VAPA изначально локализованы на мембране ER, где одинаково ориентированы в направлении цитозоля [18, 21]. Подобный механизм ингибирования репликации HCV описан

ранее у основного белка антивирусного клеточного ответа — вилерина [22].

На биологическую значимость STAT3-зависимых противовирусных механизмов в инфицированных гепатоцитах указывает также то, что вирус выработал эффективную стратегию уклонения от иммунного ответа, вызванного инфицированием или терапевтическим применением IFN- α . Так, еще в 2003 году Hosui и соавт. показали, что экспрессия кор-белка HCV в клетках печени мышей предотвращает фосфорилирование JAK1/2 и STAT3 и тем самым ингибирует STAT3-опосредованную транскрипцию, стимулированную IL-6 [23]. Оказалось, что кор-белок HCV способен связываться с JAK1/2, причем важную роль в этом процессе играет последовательность 79PGYPWP84 в кор-белке. Единичная замена P82A в кор-белке восстанавливала активирующее действие IL-6 на JAK/STAT-зависимую экспрессию репортерного гена люциферазы, что подтверждало роль взаимодействия кор-белка и JAK1/2 как фактора, определяющего блокирование работы сигнального пути. Позже аналогичные результаты получили на клетках гепатокарциномы человека, инфицированных клоном HCV генотипа 2a (J6/JFH1), продуцирующим вирусные частицы с высоким выходом [24]. Двойная замена остатков пролина в JAK-связывающем мотиве вирусного кор-белка (79AG-YAWP84) резко снижала инфекционность вируса, несмотря на эффективную репликацию его геномной РНК. Эти результаты убедительно свидетельствуют о том, что взаимодействие между кор-белком HCV и белком JAK необходимо для эффективного воспроизводства инфекционных вирусов.

Описан и другой механизм, согласно которому HCV блокирует JAK/STAT3-зависимый сигнал. Обнаружено, что содержание STAT3 в печени больных хроническим гепатитом С (СНС) значительно ниже контрольных значений. Такой же результат получен в клетках Huh7, трансфицированных полноразмерным вирусным репликоном. Показано, что сохранение эффективной репликации HCV в культуре клеток в присутствии иммуномодуляторов и цитокинов связано с нарушением передачи JAK/STAT3-зависимых сигналов [25]. Позднее были опубликованы данные о блокировании сигнального пути JAK/STAT3 как в первичных иммунных клетках, так и в гепатоцитах больных СНС, не прошедших стандартный курс лечения IFN- α /рибавирином [26]. Оказалось, что хотя у HCV-инфицированных пациентов содержание мРНК STAT1 и STAT3 во всех основных популяциях клеток оставалось на уровне контроля, содержание обоих факторов в виде белка было снижено. На примере инфекционной клеточной системы HCVcc, поддерживающей репликацию вируса штамма JFH-1 генотипа 2a на высоком уровне, удалось показать, что в HCV-инфицированных клетках суммарное количество

убиквитинированного белка выше, чем в контроле, имитирующем инфекцию. Оказалось, что в этих клетках STAT3 по большей части также убиквитинирован и подвергается протеасомной деградации, что, вероятно, и позволяет HCV формировать резистентность к терапии IFN- α .

УЧАСТИЕ ГЕПСИДИНА В КЛЕТОЧНОМ ОТВЕТЕ НА HCV-ИНФЕКЦИЮ

Как сказано выше, гепсидин — это центральный регулятор гомеостаза железа, причем единственной мишенью, связываясь с которой гепсидин контролирует этот процесс, является ферропортин (Fpn). Этот белок обеспечивает экспорт железа, располагаясь в мембранах абсорбтивных энтероцитов кишечника, макрофагов, гепатоцитов и плацентарных клеток. Fpn транспортирует Fe²⁺ совместно с медьзависимыми феррооксидазами (гефестин и церулоплазмин), которые облегчают экстракцию железа из ферропортинового канала и последующую нагрузку Fe³⁺ на трансферрин плазмы [3, 4]. Для изучения взаимодействия ферропортина с экзогенным гепсидином Nemeth и соавт. создали стабильную клеточную линию (HEK293-Fpn), в которой экспрессия химерного Fpn-GFP мыши находилась под контролем стероид-индуцируемого промотора [27]. Полученные данные указывали на то, что за связывание с ферропортином отвечает N-концевой фрагмент гепсидина, после чего происходит интернализация Fpn в цитоплазму клетки, его убиквитинирование и деградация в лизосомах. В результате выход Fe²⁺ в среду становится невозможным и наблюдается увеличение содержания железа внутри клетки. Позднее уточнили, что интернализация ферропортина возможна только в виде димерного комплекса (Fpn-HAMP)₂, в составе которого после присоединения и трансактивации двух киназ JAK2 происходит фосфорилирование остатков тирозина Y302 и Y303 в каждом из двух белков Fpn [28]. Важно отметить, что в культуре гепатоцитов, активно экспрессирующих Fpn, его деградацию равно вызывали и синтетический, и эндогенно секретируемый гепсидин, что позволило ясно идентифицировать гепатоциты как клеточную мишень данного дефенсина [29].

У 30–40% больных СНС в плазме крови повышено содержание железа, ферритина, а также увеличен уровень насыщения железом трансферрина [30]. В этой связи активно изучали механизмы, посредством которых HCV-инфекция провоцирует избыточное накопление железа в организме больного. В настоящее время признано, что воздействие вируса на метаболизм железа осуществляется путем снижения экспрессии гепсидина, что, в свою очередь, приводит к возрастанию активности Fpn в абсорбтивных энтероцитах ки-

щечника [31]. Однако имеет ли повышение уровня железа в организме больного какое-либо значение для репликации вируса или это просто одно из негативных последствий персистирующей инфекции, до сих пор остается невыясненным, поскольку данные на этот счет весьма противоречивы [30, 31]. Так, например, добавление экзогенных комплексов соединений Fe^{2+} ингибирует активность рекомбинантной вирусной РНК-зависимой РНК-полимеразы *in vitro*, а также резко подавляет размножение высокопродуктивного HCV генотипа 2a (JFH-1) в клеточной инфекционной системе [32, 33]. С другой стороны, анализ образцов биопсии печени пациентов с СНС, инфицированных вирусом генотипов 1b и 2b, выявил положительную корреляцию между уровнем железа, с одной стороны, и РНК HCV и мРНК фактора eIF3, необходимого для IRES-зависимой трансляции вирусного генома, с другой [34].

Неожиданным оказался тот факт, что не только размножение вируса подавляет экспрессию гепсидина, но и наоборот — добавление экзогенного гепсидина может снижать экспрессию вирусных белков и РНК [35]. Так, в клетках Huh7.5, инфицированных вирусом изолята JFH1 или химерой последнего — JC1, наблюдался сильный противовирусный эффект гепсидина, связанный с активацией STAT3, предположительно, через JAK-зависимое фосфорилирование. Дополнительные эксперименты по нокдауну STAT3 подтвердили важную роль этого фактора в индуцированной гепсидином противовирусной активности. Интересно, что инкубация с гепсидином вызывала значительное повышение его собственной мРНК в клетках Huh7.5, указывая на существование положительной обратной связи, возникающей за счет STAT3-зависимой активации экспрессии *HAMP* [5, 6]. Клинические исследования подтверждают строгую отрицательную корреляцию между содержанием гепсидина в крови больных СНС и вирусной нагрузкой. Такие результаты в свете всего сказанного можно трактовать как свидетельство анти-HCV-активности гепсидина *in vivo* [36].

Механизм, посредством которого гепсидин запускает фосфорилирование STAT3, необходимое для активации фактора, подробно изучен на примере макрофагов мыши, в которых с помощью цитрата Fe^{3+} индуцировали экспрессию *Fpn* [37]. Так, добавление гепсидина к макрофагам, активированным цитратом железа, вызывало ассоциацию *Fpn* с киназой JAK2, с ее последующей активацией, как описано ранее для клеток HEK293-*Fpn* [28]. Комплекс (JAK2-*Fpn*-*HAMP*)₂ рекрутировал в свой состав STAT3, что подтверждено совместной коиммунопреципитацией JAK2 и STAT3 с *Fpn* в присутствии гепсидина. Более того, фосфорилирование STAT3, связанного с комплексом по

остатку Y705, предполагает переход данного фактора в активное состояние. Таким образом, мембраносвязанный белок *Fpn* можно рассматривать как рецептор гормона гепсидина, при этом в результате связывания рецептора и гормона происходит активация сигнального пути STAT3 [3]. Действительно, глобальный анализ мРНК в макрофагах, экспрессирующих ферропортин, после обработки гепсидином выявил изменение уровня экспрессии широкого спектра генов, в первую очередь регулирующих цитокин-индуцированные воспалительные реакции, контролируемые STAT3.

РОЛЬ HDAC3 В ЭПИГЕНЕТИЧЕСКОЙ РЕГУЛЯЦИИ ЭКСПРЕССИИ *HAMP*

HDAC образуют обширное семейство ферментов, осуществляющих эпигенетическую регуляцию экспрессии генов и контролирующих структурно-функциональную реорганизацию хроматина и цитоскелета. Выделяют цинкзависимые HDAC классов I, II и IV, которые гидролизуют амидные связи в ϵ -N-ацетиллизининовых остатках белков-субстратов, а также HDAC класса III, обозначаемые как SIRT (сиртуины), использующие для деацетилирования кофермент NAD^+ [37]. Цинкзависимые HDAC могут локализоваться либо в ядре (HDAC1/2/3 класса I), либо в ядре и цитоплазме (HDAC8 класса I и HDAC4/5/7/9 класс IIa), либо в цитоплазме (HDAC6/10 класса IIb и HDAC11 класса IV). При этом HDAC класса IIa не обладают деацетилирующей активностью, но при перемещении в ядро связываются с HDAC3 в составе регуляторных комплексов, влияя таким образом на эффективность деацетилирования гистонов и факторов транскрипции [39].

В регуляции экспрессии гепсидина принимает участие HDAC3 (рис. 2), что впервые показали Pasricha и соавт. в 2017 году [40]. Первоначальное тестирование шести ингибиторов HDAC различной селективности обнаружило усиление экспрессии *HAMP* в клетках Huh7 только в присутствии селективного ингибитора HDAC3 (RGFP966), причем активность сигнального пути BMP/SMAD, также участвующего в регуляции гепсидина, оставалась на уровне контроля. К аналогичному результату приводили и встречные эксперименты по нокдауну HDAC3. Наоборот, в клетках, сверхэкспрессирующих HDAC3, наблюдалось снижение экспрессии *HAMP*, но не генов, регулируемых по BMP/SMAD-зависимому пути. Кроме того, ChIP-seq-анализ материала из печени мыши показал, что содержание HDAC3 в локусе гепсидина в районе старта транскрипции значительно выше, чем в области отрицательного контроля и является примерно таким же, как и в локусе *CDKN1*, взятом в качестве положительного контроля, учи-

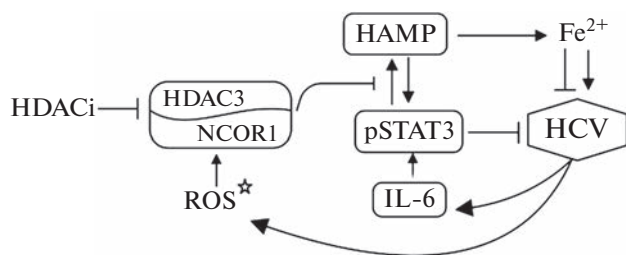


Рис. 2. Схематическое представление регуляторного контура HCV-HDAC3-STAT3-HAMP. Острая HCV-инфекция индуцирует IL-6/STAT3-зависимый противовирусный ответ. Активированный pSTAT3 регулирует экспрессию гепсидина (HAMP) по механизму положительной обратной связи ($\uparrow\downarrow$). На стадии хронической стабилизации комплекса, образованного HDAC3 и корепрессором ядерных рецепторов 1 (NCoR1), что приводит к подавлению экспрессии *HAMP* и содействует персистентной вирусной инфекции. В свою очередь, фармакологическое ингибирование HDAC3 (HDACi) восстанавливает экспрессию *HAMP* и подавляет репликацию HCV. Влияние Fe^{2+} на размножение вируса имеет, вероятно, двунаправленный характер ($\uparrow\downarrow$) и его роль в репликации HCV ясна не до конца.

таявая прямое участие HDAC3 в регуляции экспрессии p21 [41].

Изучена также роль корепрессора NCoR1 в регуляции экспрессии *HAMP* в клетках Nuh7. Нокдаун и *NCoR1*, и *HDAC3* по отдельности повышал экспрессию гепсидина, а одновременный нокдаун обоих генов давал эффект синергии [41]. Таким образом, HDAC3 необходим, но недостаточен для эффективной репрессии транскрипции гена *HAMP* гепсидина. Совместное присутствие HDAC3 и NCoR1 в составе репрессорных комплексов описано еще в 2000 году [42]. Значение этого наблюдения возросло после того, как было показано, что взаимодействие HDAC3 с DAD-

меном NCoR1 резко повышает деацетилазную активность фермента *in vivo*, что во многом гарантирует эффективность репрессии транскрипции (рис. 2) [43].

Сложная структурно-доменная организация NCoR1 объясняется его центральной ролью в сборке и функционировании репрессорного комплекса. Так, за доменом DAD следует родственный ему домен HID, который непосредственно связывает N-концевые последовательности гистонов, предпочтительно деацетилированных, что дополнительно увеличивает силу репрессии [44]. Кроме того, в составе NCoR1 выделяют три домена RD, ответственных за взаимодействие с регуляторными белками, и три изолейцин-богатых домена RID, отвечающих за взаимодействие с ядерными рецепторами и факторами транскрипции, включая STAT3. Как правило, RID-домены корепрессоров присоединяются к той же области ядерных рецепторов, что и коактиваторы, что усиливает антагонизм их действия (рис. 3) [45].

В 2012 году Watson и соавт. показали, что инозитол-1,4,5,6-тетрафосфат (IP4) выполняет роль кофермента HDAC3 [46]. В ходе уточнения кристаллической структуры комплекса HDAC3-DAD, полученного в результате транзитной трансфекции в клетках HEK-293F, идентифицирована молекула IP4, внедренная в стык между HDAC3 и DAD-доменом. Связывание IP4 не только стабилизировало взаимодействие двух белков в качестве “межмолекулярного клея”, но и облегчало доступ ацетилированного субстрата к активному центру и повышало активность фермента. Мутации, разрушающие комплекс HDAC3-DAD, инактивировали HDAC3, подтверждая необходимость связывания с DAD-доменом для ферментативной активности. Однако мутации некоторых остатков DAD-домена, расположенных в районе сайта связывания IP4, существенно инактивиро-

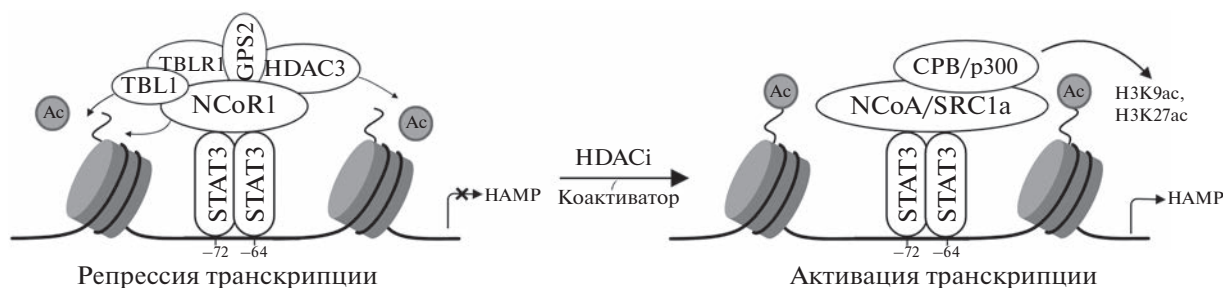


Рис. 3. Упрощенная схема активации экспрессии гепсидина (HAMP) при ингибировании активности гистондеацетилазы 3 (HDAC3). HDAC3 в комплексе с корепрессором NCoR1 связывается с димером STAT3 и способствует подавлению транскрипции гена *HAMP* за счет деацетилирования гистонов. Белки TBL1/TBLR1, стабильно присутствующие в комплексе, участвуют в репрессии транскрипции, в обмене корепрессора на коактиватор и в убиквитинировании и деградации NCoR1, в то время как молекулярная функция регуляторного белка GPS2 в составе комплекса не ясна. Связывание ингибитора фермента (HDACi) с HDAC3 дестабилизирует репрессорный комплекс, что облегчает присоединение коактиватора (NCoA/SRC1a) и гистон-ацетилазы (CPB/p300) к димеру STAT3 и обеспечивает тем самым транскрипцию гена *HAMP* в том числе и за счет ацетилирования гистонов.

вали HDAC3, не нарушая при этом белок-белковое взаимодействие. Эти результаты указывали на то, что присутствие IP4 способствует усилению деацетилазной активности. Дополнительные эксперименты *in vitro* с использованием рекомбинантных белков HDAC3 и SMRT (аналог NCoR1) показали, что IP4 увеличивает каталитическую эффективность комплекса HDAC3-SMRT против ацетилированного фрагмента гистона H4 AK12-18(K16ac) приблизительно в 2.5 раза (значение k_{cat}/K_m) [47].

ОКИСЛИТЕЛЬНЫЙ СТРЕСС РЕГУЛИРУЕТ АКТИВНОСТЬ HDAC3

В ответ на развитие OS включается дополнительный механизм HDAC-зависимой регуляции экспрессии гепсидина, детали которого до сих пор не определены [9]. Однако мы полагаем, что этот механизм можно реконструировать с большой достоверностью, опираясь на современные данные. Дело в том, что взаимодействие HDAC3 с DAD-доменом NCoR1 необходимо, но недостаточно для максимальной активации деацетилирования, и конечная настройка работы комплекса происходит уже после его сборки.

В ответ на HCV-индуцированный OS происходит падение экспрессии *HAMP*, что соответствует повышению активности HDAC3 [9], однако мы не нашли никаких указаний на положительную корреляцию между уровнем ROS и содержанием IP4. С другой стороны, известно, что кофермент NADPH является ключевым кофактором антиоксидантной защиты. Он служит донором восстановительного потенциала глутатионредуктаз (GR) и тиоредоксинредуктаз (TrxR), превращающих окисленные формы глутатиона и тиоредоксинов в их исходное восстановленное состояние [48, 49]. Недавно обнаружена независимая от клеточного метаболизма функция NADPH в модуляции эпигенетического статуса и транскрипции. Оказалось, что снижение клеточного уровня NADPH вследствие нокдауна *ME1* или *ME2* (гены цитозольной или митохондриальной малатдегидрогеназы) снижает уровень общего ацетилирования гистонов и транскрипции в адипоцитах и в опухолевых клетках [50]. При этом введение в клетки NADPH с помощью системы X-tremeGENE, предназначенной для трансфекции ДНК, вызывало дозозависимое накопление ацетилированной формы гистонов H2B, H3 и H4.

Причиной наблюдаемых закономерностей была способность NADPH непосредственно взаимодействовать с HDAC3 и разрушать комплекс HDAC3-NCoR1 или HDAC3-SMRT, препятствуя, тем самым активации HDAC3. Существенно, что NADPH и IP4, по-видимому, связывались с одними и теми же доменами HDAC3, причем NADPH имел более высокое сродство к HDAC3,

чем IP4. Таким образом, в то время как IP4 способствовал образованию комплекса HDAC3-NCoR1, NADPH ингибировал его. В экспериментах с меченной Flag HDAC3, выделенной из клеток HEK293T, показано, что и IP4, и NADPH влияют на активность HDAC3 дозозависимым образом, при этом IP4 увеличивает активность, а NADPH — уменьшает. Ни одна из других форм кофермента ($NADP^+$, $NADH$ или NAD^+) не влияла на ферментативную активность HDAC3, так как все они обладали в 100–200 раз меньшей аффинностью к ферменту, чем NADPH.

В ходе HCV-инфекции NADPH играет двоякую роль в регуляции промотированного вирусом OS. С одной стороны, кофермент, как обычно, обеспечивает восстановительную активность глутатион- и тиоредоксинредуктаз, а с другой, участвует в постоянном генерировании ROS NADPH-оксидазой 4 (NOX4), экспрессию которой индуцирует кор-белок HCV [51, 52]. Все вместе это может означать, что уровень NADPH в инфицированной клетке снижается, поскольку при осуществлении как антиоксидантной, так и прооксидантной активности кофермента, NADPH расходуется за счет окисления. Действительно, многочисленные клинические исследования показывают, что в тканях печени пациентов с СНС многократно снижено соотношение глутатиона и глутатионредуктазы, которое, сверх того, коррелирует с усилением перекисного окисления липидов и накоплением железа [53–55]. Таким образом, у нас есть все основания полагать, что в условиях OS снижение клеточной концентрации NADPH имеет результатом присоединение IP4 к вакантному сайту общего связывания с комплексом HDAC3-NCoR1 и его стабилизацию. В свою очередь, это приводит к усилению деацетилазной активности фермента и, следовательно, к подавлению экспрессии *HAMP*.

РОЛЬ HDAC3 В РАВНОВЕСИИ КОРЕПРЕССОР/КОАКТИВАТОР

Укоренившийся взгляд на то, что деацетилирование гистонов регулирует экспрессию генов прежде всего за счет изменения трехмерной структуры хроматина и компактизации ДНК, весьма односторонен и требует существенного уточнения. Такой механизм регуляции необходим в качестве своеобразного “слагбаума”, перекрывающего доступ транскрипционного аппарата к протяженным кластерам молчащих генов гетерохроматина. Однако для регуляции экспрессии активных генов эухроматина HDAC должны реагировать на поступающие сигналы быстро и строго в границах соответствующих регуляторных районов или даже промоторов. Указанные требования обеспечиваются присутствием HDAC в составе репрессорных комплексов, и, наоборот, гистон-

ацетилтрансферазы (НАТ) рекрутируются в комплексы активации транскрипции [56, 57]. Ниже будут рассмотрены возможные механизмы переключения экспрессии *HAMP* из режима репрессии в режим активации в ответ на действие ингибиторов HDAC3 (рис. 3).

Самый универсальный механизм — присоединение к корепрессорному комплексу дестабилизирующих лигандов, в число которых входят гормоны, метаболиты или химические агенты со свойствами лигандов ядерных рецепторов [58, 59]. Как упоминалось выше, транскрипционный фактор STAT3 после активации путем фосфорилирования и димеризации связывается с соответствующим мотивом промотора *HAMP* [5, 6] и в зависимости от клеточного сигнала рекрутирует либо корепрессор NCoR1, либо коактиватор NCoA/SRC1a (рис. 3) [60]. Стоит заметить, что описанное в предыдущей главе угнетение ассоциации NCoR1 с HDAC3 в присутствии NADPH [50] можно рассматривать как действие лиганда-метаболита, назначение которого — активировать экспрессию целевого гена *HAMP* в условиях нормоксии.

Могут ли ингибиторы деацетилазной активности HDAC3 играть роль лигандов, дестабилизирующих корепрессорный комплекс HDAC3-NCoR1 — точно не известно [58]. Однако в случае фармакологического ингибирования HDAC4 достоверно показано, что связывание ингибитора (TFMK) с активным сайтом белка-мишени меняет его сродство к белкам-партнерам. Гетеродимер HDAC3-SMRT, действуя как каркас для сборки мультибелкового репрессорного комплекса, присоединял к себе HDAC4 в качестве регуляторного элемента. Данные кристаллографического анализа и встречного мутагенеза показывают, что RD3-домен SMRT взаимодействует с районом, примыкающим к активному центру HDAC4 [61, 62]. Как и следовало ожидать, после внедрения ингибитора в канал активного центра HDAC4 в этом сегменте белка происходили существенные конформационные изменения, которые снижали эффективность взаимодействия с RD3-SMRT [63].

Необходимо отметить, что ингибиторы HDAC прямо влияют на транскрипцию STAT-регулируемых генов, поскольку большинство факторов транскрипции этого семейства служат субстратами HDAC [64]. В 2005 году Yuan и соавт. показали, что под действием цитокинов в клетках рака молочной железы происходит опосредованное НАТ р300-ацетилирование STAT3 по остатку Lys685. Эта модификация была обратимой благодаря активности локализованных в ядре цинкзависимых HDAC класса I, главным образом HDAC3 [65]. В экспериментах с клеточной линией PC-3 рака предстательной железы человека, в которой экспрессировался мутантный STAT3, содержащий замену Lys685Arg, обнаружили, что ацетилирова-

ние Lys685 критично для формирования стабильного гомодимера STAT3 и усиления транскрипции генов прохождения клеточного цикла. Более того, TSA, пан-ингибитор HDAC, способствовал накоплению STAT3K685ac и индуцировал экспрессию циклина D1, но только в клетках, экспрессирующих STAT3 дикого типа.

Описан еще один механизм подавления сборки корепрессорного комплекса в зависимости от уровня ацетилирования гистонов. Как минимум, три белка в составе комплекса — NCoR1, TBL1 и TBLR1 — считывают гистоновый код, преимущественно связываясь со свободными остатками лизинов гистоновых “хвостов” (рис. 3). HD-домен NCoR1 или SMRT, как показано, выполняет несколько важных функций: (i) увеличивает аффинность части комплекса HDAC3-DAD к гистоновому субстрату, способствуя его деацетилированию; (ii) увеличивает силу репрессии за счет высокого сродства к деацетилированным гистонам; (iii) блокирует НАТ-зависимое ацетилирование гистонов [66]. Домены RD1 и RD4 в составе NCoR1/SMRT взаимодействуют с N-концевой и C-концевой последовательностями белков TBL1 и TBLR1. Оба эти белка связывают гистоны H2B и H4, причем эффективность репрессии положительно коррелирует с силой их взаимодействия с гистонами [67]. Предложена модель, в которой деацетилирование, осуществляемое первоначальным набором нестабильных комплексов на базе NCoR1/SMRT через их взаимодействие с фактором транскрипции, например, рецептором тиреоидных гормонов, генерирует гистоновый код, который служит для сборки и стабилизации конечного репрессорного комплекса [68]. Очевидно, что фармакологическое ингибирование HDAC3 препятствует всем перечисленным взаимодействиям и, как следствие, не образуется стабильный репрессорный комплекс.

При нарушении сборки корепрессорного комплекса происходит конкурентное взаимодействие ДНК-связанного димера STAT3 с белком коактиватором NCoA/SRC1a и двумя паралогичными НАТ CPB/p300 (рис. 3). Анализ взаимодействия делеционных мутантов позволил установить, что стимулирующую транскрипционную активность коактиваторного комплекса обеспечивает НАТ CPB, контактирующая с центральным AD1-доменом NCoA/SRC1a [69]. Далее первичный коактиваторный комплекс рекрутирует многочисленные эпигенетические модуляторы, в том числе аргинин/лизинметилтрансферазы. Совместное действие рекрутированных ферментов приводит к ремоделированию структуры хроматина, что делает возможной сборку на промоторе общих факторов транскрипции и РНК-полимеразы II в единый мультибелковый функционально-активный транскрипционный комплекс [70].

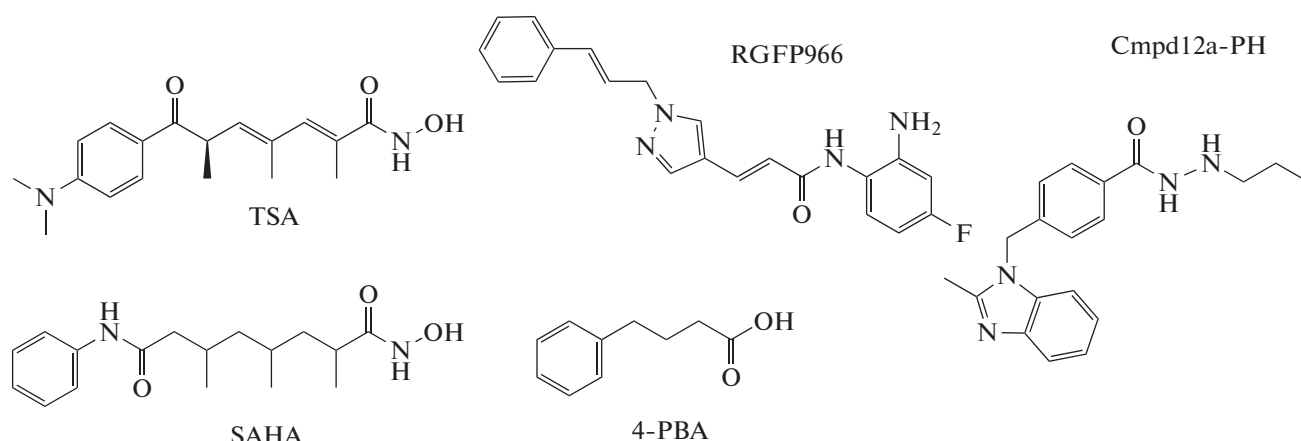


Рис. 4. Структура ингибиторов гистондеацетилаз — блокаторов репликации HCV — гидроксамовых кислот TSA и SAHA, орто-аминоанилида RGFP966, карбоновой кислоты 4-PBA и *N'*-пропилгидразида Cmpd12a-PH.

ИНГИБИТОРЫ HDAC3 КАК ПРОТИВОВИРУСНЫЕ АГЕНТЫ

Эпигенетические механизмы, по которым фармакологическое ингибирование HDAC может приводить к подавлению репликации HCV, впервые изучили на примере двух ингибиторов — TSA и вориностата (SAHA, рис. 4) [10]. ChIP-seq анализ клеток гепатомы человека после инкубации с SAHA указал на два ключевых гена, регулируемых HDAC. Выявлена сильная положительная регуляция гена остеопонтина (*OPN*) и сильная отрицательная — аполипопротеина A1 (*APOA1*). Важность белковых продуктов этих генов в репликации HCV подтверждена с помощью сверхэкспрессии *OPN* и нокаута *APOA1* — в обоих случаях репликация вируса тормозилась, о чем свидетельствовало резкое снижение люциферазного сигнала репликона. Авторы исключили ингибирующий механизм, связанный со специфической для печени микроРНК-122, которая в норме способствует пролиферации вируса [71], показав, что SAHA не влияет на ее экспрессию.

Через пять лет подробно разобрали механизм анти-HCV действия SAHA и селективного ингибитора HDAC3 RGFP966 [70]. Оба препарата эффективно подавляли пролиферацию HCV в системе клеточного репликона. Были подтверждены данные о влиянии ингибирования HDAC на экспрессию генов гепсидина (*HAMP*) и аполипопротеина A1 (*APOA1*) [9, 10]. Кроме того, впервые показано накопление ацетилированной формы STAT3 при инкубировании клеток в присутствии RGFP966. Основываясь на данных об активации с помощью RGFP966 экспрессии *HAMP in vivo* [40], на животных моделях установлено значение HDAC3 как перспективной мишени антивирусной терапии. С этой целью использовали восприимчивую к вирусу линию ген-модифицированных мышей с введенными генами человека CD81

(*TSPAN28*) и окклюдина (*OCLN*). Лечение HCV-инфицированных животных с помощью SAHA и RGFP966 показало высокую, приблизительно равную, эффективность обоих препаратов.

Другой пан-ингибитор HDAC, фармакологически ориентированный на гепсидин, — 4-фенилбутановая кислота (4-PBA, рис. 4) [73]. Интересно, что 4-PBA усиливала экспрессию *HAMP*, вызывая в соответствующем локусе модификацию гистона H3K27ac, а не H3K9ac, как описано ранее для TSA и RGFP966. Возможно, что это отличие имело следствием неожиданно сильную активацию сигнала IFN- α . Дополнительный вклад в антивирусное действие 4-PBA вносило ее деструктивное влияние на формирование HCV-реписом — двухмембранных везикул (DMV), ассоциированных с вирусным комплексом репликации (рис. 1). Интересно, что противовирусная активность 4-PBA и в клеточной системе, и на животных моделях показана на одних и тех же клетках гепатомы человека Huh7.5-Con1, поддерживаемых в адгезионной культуре или пересаженных в печень мышей с иммунодефицитом.

Селективным ингибитором HDAC класса I и одновременно эффективным блокатором репликации HCV в клеточной системе оказался *N'*-пропилгидразид 4-замещенной бензойной кислоты (Cmpd12a-PH, рис. 4) [74]. Однако между сильной противовирусной активностью Cmpd12a-PH (EC_{50} = 25 нМ и TI = 400) и накоплением ацетилированных форм гистона H3K9/K14ac наблюдался значительный концентрационный разрыв. Нельзя исключить, что в присутствии Cmpd12a-PH, как и 4-PBA [73], накапливается H3K27ac, что и могло стать причиной наблюдаемого несоответствия. Это можно также объяснить существованием другой высокочувствительной антивирусной мишени. Следует заметить, что ингибиторы HDAC, в структуру которых в качестве

цинк-хелатирующей группы входит *N*-пропил-гидразидный остаток, проявляют максимальную селективность ингибирования в отношении HDAC3. Кроме того, они обладают повышенной устойчивостью к биотрансформации в отличие от гидроксаматов и *орто*-аминоанилидов (TSA, SAHA и RGFP966 соответственно) [75].

ЗАКЛЮЧЕНИЕ

Если исходить из общих представлений, то удивителен сам факт эффективного подавления вирусной репликации ингибиторами HDAC при относительно низкой токсичности препаратов. Учитывая ключевую роль HDAC в регуляции экспрессии генов, интуитивно можно ожидать самых серьезных негативных последствий для организма пациентов. Однако в реальности это не происходит или, точнее, происходит не всегда.

Дело в том, что вирусная инфекция, в частности HCV, разнонаправленно меняет экспрессию многих генов, что позволяет вирусу репрограммировать гепатоциты в состояние перmissивности, причем экспрессия некоторых генов меняется многократно. Ингибиторы HDAC3 также действуют разнонаправленно, эффективно подавляя сверхэкспрессию “провирусных” генов (*APOA1*) и восстанавливая до нормы гипоэкспрессию “антивирусных” генов (*HAMP* и *OPN*). Заметим, что ингибиторы HDAC в качестве противоопухолевых препаратов аналогичным образом воздействуют на экспрессию онкогенов и генов-супрессоров соответственно.

Очевидно, что роль HDAC3 в устойчивой репликации HCV не ограничена модуляцией экспрессии трех генов, указанных выше. Вследствие того, что HDAC3 входит в состав корепрессорных комплексов с самыми разнообразными ядерными рецепторами, регулирующими, например, метаболизм печени или активность циркадных генов [57], закономерно ожидать появление новых генов-мишеней для анти-HCV-терапии с помощью подходов ингибиторов HDAC3. Кроме того, HDAC3 рассматривается в качестве перспективной мишени и для лечения поздних стадий СНС, когда требуется воздействовать одновременно на сложный патологический процесс в тканях печени – вирусную инфекцию, клеточное старение и злокачественную трансформацию [76].

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (грант 20-04-00504) и Российского научного фонда (грант 23-24-00542).

Настоящая статья не содержит каких-либо исследований с участием людей или животных в качестве объектов исследований.

Авторы подтверждают отсутствие конфликта интересов.

СПИСОК ЛИТЕРАТУРЫ

1. Malaguarnera M., Di Fazio I., Romeo M.A., Restuccia S., Laurino A., Trovato B.A. (1997) Elevation of interleukin 6 levels in patients with chronic hepatitis due to hepatitis C virus. *J. Gastroenterol.* **32**, 211–215.
2. Zhu H., Shang X., Terada N., Liu C. (2004) STAT3 induces anti-hepatitis C viral activity in liver cells. *Biochem. Biophys. Res. Commun.* **324**, 518–528.
3. Hentze M.W., Muckenthaler M.U., Galy B., Camaschella C. (2010) Two to tango: regulation of mammalian iron metabolism. *Cell.* **142**, 24–38.
4. Roth M.P., Meynard D., Coppin H. (2019) Regulators of hepcidin expression. *Vitam. Horm.* **110**, 101–129.
5. Wrighting D.M., Andrews N.C. (2006) Interleukin-6 induces hepcidin expression through STAT3. *Blood.* **108**, 3204–3209.
6. Falzacappa V.M.V., Spasic V.M., Kessler R., Stolte J., Hentze M.W., Muckenthaler M.U. (2007) STAT3 mediates hepatic hepcidin expression and its inflammatory stimulation. *Blood.* **109**, 353–358.
7. Choi S.O., Cho Y.S., Kim H.L., Park J.W. (2007) ROS mediate the hypoxic repression of the hepcidin gene by inhibiting C/EBP α and STAT-3. *Biochem. Biophys. Res. Commun.* **356**, 312–317.
8. Shi D., Xie F., Zhai C., Stern J.S., Liu Y., Liu S. (2009) The role of cellular oxidative stress in regulating glycolysis energy metabolism in hepatoma cells. *Mol. Cancer.* **8**(1), 32.
9. Miura K., Taura K., Kodama Y., Schnabl B., Brenner D.A. (2008) Hepatitis C virus-induced oxidative stress suppresses hepcidin expression through increased histone deacetylase activity. *Hepatology.* **48**, 1420–1429.
10. Sato A., Saito Y., Sugiyama K., Sakasegawa N., Muramatsu T., Fukuda S., Yoneya M., Kimura M., Ebinuma H., Hibi T., Saito H. (2013) Suppressive effect of the histone deacetylase inhibitor suberoylanilide hydroxamic acid (SAHA) on hepatitis C virus replication. *J. Cell. Biochem.* **114**, 1987–1996.
11. Lee J.Y., Cortese M., Haselmann U., Tabata K., Romero-Brey I., Funaya C., Schieber N.L., Qiang Y., Bartenschlager M., Kallis S., Ritter C., Rohr K., Schwab Y., Ruggieri A., Bartenschlager R. (2019) Spatiotemporal coupling of the hepatitis C virus replication cycle by creating a lipid droplet-proximal membranous replication compartment. *Cell Rep.* **27**, 3602–3617.
12. Kamimura D., Ishihara K., Hirano T. (2003) IL-6 signal transduction and its physiological roles: the signal orchestration model. *Rev. Physiol. Biochem. Pharmacol.* **149**, 1–38.
13. Littlewood T.D., Hancock D.C., Danielian P.S., Parker M.G., Evan G.I. (1995) A modified estrogen receptor ligand-binding domain as an improved switch for the regulation of heterologous proteins. *Nucl. Acids Res.* **23**, 1686–1690.
14. Matsuda T., Nakamura T., Nakao K., Arai T., Katsuki M., Heike T., Yokota T. (1999) STAT3 activation is sufficient to maintain an undifferentiated state of mouse embryonic stem cells. *EMBO J.* **18**, 4261–4269.
15. Zhu H., Liu C. (2003) Interleukin-1 inhibits hepatitis C virus subgenomic RNA replication by activation of extracellular regulated kinase pathway. *J. Virol.* **77**, 5493–5498.

16. Yao L., Dong H., Zhu H., Nelson D., Liu C., Lambiase L., Li X. (2011) Identification of the *IFITM3* gene as an inhibitor of hepatitis C viral translation in a stable STAT1 cell line. *J. Viral Hepat.* **18**(10), e523–e529.
17. Amini-Bavil-Olyae S., Choi Y.J., Lee J.H., Shi M., Huang I.C., Farzan M., Jung J.U. (2013) The antiviral effector IFITM3 disrupts intracellular cholesterol homeostasis to block viral entry. *Cell Host Microbe.* **13**, 452–464.
18. Narayana S.K., Helbig K.J., McCartney E.M., Eyre N.S., Bull R.A., Eltahla A., Lloyd A.R., Beard M.R. (2015) The interferon-induced transmembrane proteins, IFITM1, IFITM2, and IFITM3 inhibit hepatitis C virus entry. *J. Biol. Chem.* **290**, 25946–25959.
19. Neefjes J., Cabukusta B. (2021) What the VAP: The expanded VAP family of proteins interacting with FFAT and FFAT-related motifs for interorganelle contact. *Contact.* **4**, 1–11.
20. Paul D., Hoppe S., Saher G., Krijnse-Locker J., Bartenschlager R. (2013) Morphological and biochemical characterization of the membranous hepatitis C virus replication compartment. *J. Virol.* **87**, 10612–10627.
21. Bailey C.C., Kondur H.R., Huang I.-C., Farzan M. (2013) Interferon-induced transmembrane protein 3 is a type II transmembrane protein. *J. Biol. Chem.* **288**, 32184–32193.
22. Helbig K.J., Beard M.R. (2014) The role of viperin in the innate antiviral response. *J. Mol. Biol.* **426**, 1210–1219.
23. Hosui A., Ohkawa K., Ishida H., Sato A., Nakanishi F., Ueda K., Takehara T., Kasahara A., Sasaki Y., Hori M., Hayashi N. (2003) Hepatitis C virus core protein differentially regulates the JAK-STAT signaling pathway under interleukin-6 and interferon- γ stimuli. *J. Biol. Chem.* **278**, 28562–28571.
24. Lee C. (2013) Interaction of hepatitis C virus core protein with Janus kinase is required for efficient production of infectious viruses. *Biomol. Ther.* **21**, 97–106.
25. Larrea E., Aldabe R., Molano E., Fernandez-Rodriguez C., Ametzazurra A., Civeira M., Prieto J. (2006) Altered expression and activation of signal transducers and activators of transcription (STATs) in hepatitis C virus infection: *in vivo* and *in vitro* studies. *Gut.* **55**, 1188–1196.
26. Stevenson N.J., Bourke N.M., Ryan E.J., Binder M., Fanning L., Johnston J.A., John E. Hegarty J.E., Long A., O'Farrelly C. (2013) Hepatitis C virus targets the interferon- α JAK/STAT pathway by promoting proteasomal degradation in immune cells and hepatocytes. *FEBS Lett.* **587**, 1571–1578.
27. Nemeth E., Tuttle M.S., Powelson J., Vaughn M.B., Donovan A., Ward D.M. (2004) Hepcidin regulates cellular iron efflux by binding to ferroportin and inducing its internalization. *Science.* **306**, 2090–2093.
28. De Domenico I., Lo E., Ward D.M., Kaplan J. (2009) Hepcidin-induced internalization of ferroportin requires binding and cooperative interaction with Jak2. *Proc. Natl. Acad. Sci. USA.* **106**, 3800–3805.
29. Ramey G., Deschemin J.C., Durel B., Canonne-Hergaux F., Nicolas G., Vaulont S. (2009) Hepcidin targets ferroportin for degradation in hepatocytes. *Haematologica.* **95**, 501–504.
30. Georgopoulou U., Dimitriadis A., Foka P., Karamichali E., Mamalaki A. (2014) Hepcidin and the iron enigma in HCV infection. *Virulence.* **5**, 465–476.
31. Zou D.M., Sun W.L. (2017) Relationship between hepatitis C virus infection and iron overload. *Chin. Med. J.* **130**, 866–871.
32. Fillebeen C., Rivas-Estilla A.M., Bisaillon M., Ponka P., Muckenthaler M., Hentze M.W., Koromilas A.E., Pantopoulos K. (2005) Iron inactivates the RNA polymerase NS5B and suppresses subgenomic replication of hepatitis C virus. *J. Biol. Chem.* **280**, 9049–9057.
33. Fillebeen C., Pantopoulos K. (2010) Iron inhibits replication of infectious hepatitis C virus in permissive Huh7.5.1 cells. *J. Hepatol.* **53**, 995–999.
34. Theurl I., Zoller H., Obrist P., Datz C., Bachmann F., Elliott R. M., Weiss G. (2004) Iron regulates hepatitis C virus translation via stimulation of expression of translation initiation factor 3. *J. Infect. Dis.* **190**, 819–825.
35. Liu H., Trinh T.L., Dong H., Keith R., Nelson D., Liu C. (2012) Iron regulator hepcidin exhibits antiviral activity against hepatitis C virus. *PLoS One.* **7**(10), e46631.
36. Wang R., Zhao S., Du J., Qiao J., Li W., Nan Y. (2021) A correlation analysis of the serum hepcidin concentrations and viral loads in HCV-infected patients. *Am. J. Transl. Res.* **13**, 6297–6304.
37. De Domenico I., Zhang T.Y., Koenig C.L., Branch R.W., London N., Lo E., Daynes R.A., Kushner J.P., Li D., Ward D.M., Kaplan J. (2010) Hepcidin mediates transcriptional changes that modulate acute cytokine-induced inflammatory responses in mice. *J. Clin. Invest.* **120**, 2395–2405.
38. Seto E., Yoshida M. (2014) Erasers of histone acetylation: the histone deacetylase enzymes. *Cold Spring Harb. Perspect. Biol.* **6**, a018713.
39. Martin M., Kettmann R., Dequiedt F. (2007) Class IIa histone deacetylases: regulating the regulators. *Oncogen.* **26**, 5450–5467.
40. Pasricha S.R., Lim P.J., Duarte T.L., Casu C., Oosterhuis D., Mleczko-Sanecka K., Suciu M., Da Silva A.R., Al-Hourani K., Arezes J., McHugh K., Gooding S., Frost J.N., Wray K., Santos A., Porto G., Repapi E., Gray N., Draper S.J., Ashley N., Soilleux E., Olinga P., Muckenthaler M.U., Hughes J.R., Rivella S., Milne T.A., Armitage A.E., Drakesmith H. (2017) Hepcidin is regulated by promoter-associated histone acetylation and HDAC3. *Nat. Commun.* **8**(1), 403.
41. Wilson A.J., Byun D.S., Popova N., Murray L.B., L'Italien K., Sowa Y., Arango D., Velcich A., Leonard H., Augenlicht L.H., Mariadason J.M. (2006) Histone deacetylase 3 (HDAC3) and other class I HDACs regulate colon cell maturation and p21 expression and are deregulated in human colon cancer. *J. Biol. Chem.* **281**, 13548–13558.
42. Li J., Wang J., Wang J.X., Nawaz Z., Liu J.M., Qin J. (2000) Both corepressor proteins SMRT and N-CoR exist in large protein complexes containing HDAC3. *EMBO J.* **19**, 4342–4350.
43. You S.H., Lim H.W., Sun Z., Broache M., Won K.J., Lazar M.A. (2013) Nuclear receptor co-repressors are required for the histone-deacetylase activity of HDAC3 *in vivo*. *Nat. Struct. Mol. Biol.* **20**, 182–187.

44. Yu J., Li Y., Ishizuka T., Guenther M.G., Lazar M.A. (2003) A SANT motif in the SMRT corepressor interprets the histone code and promotes histone deacetylation. *EMBO J.* **22**, 3403–3410.
45. Mottis A., Mouchiroud L., Auwerx J. (2013) Emerging roles of the corepressors NCoR1 and SMRT in homeostasis. *Genes Dev.* **27**, 819–835.
46. Watson P.J., Fairall L., Santos G.M., Schwabe J.W.R. (2012) Structure of HDAC3 bound to co-repressor and inositol tetrakisphosphate. *Nature*. **481**, 335–340.
47. Watson P.J., Millard C.J., Riley A.M., Robertson N.S., Wright L.C., Godage H.Y., Cowley S.M., Jamieson A.G., Potter B.V., Schwabe J.W. (2016) Insights into the activation mechanism of class I HDAC complexes by inositol phosphates. *Nat. Commun.* **7**, 11262.
48. Fernandez-Marcos P.J., Nóbrega-Pereira S. (2016) NADPH: new oxygen for the ROS theory of aging. *Onco-target.* **7**, 50814–50815.
49. Moreno-Sánchez R., Marín-Hernández Á., Gallardo-Pérez J.C., Vázquez C., Rodríguez-Enríquez S., Saavedra E. (2018) Control of the NADPH supply and GSH recycling for oxidative stress management in hepatoma and liver mitochondria. *Biochim. Biophys. Acta – Bioenerg.* **1859**, 1138–1150.
50. Li W., Kou J., Qin J., Li L., Zhang Z., Pan Y., Xue Y., Du W. (2021) NADPH levels affect cellular epigenetic state by inhibiting HDAC3–Ncor complex. *Nat. Metab.* **3**, 75–89.
51. Boudreau H.E., Emerson S.U., Korzeniowska A., Jendrysyk M.A., Leto T.L. (2009) Hepatitis C virus (HCV) proteins induce NADPH oxidase 4 expression in a transforming growth factor dependent manner: a new contributor to HCV-induced oxidative stress. *J. Virol.* **83**, 12934–12946.
52. Smirnova O.A., Ivanova O.N., Bartosch B., Valuev-El-liston V.T., Mukhtarov F., Kochetkov S.N., Ivanov A.V. (2016) Hepatitis C virus NS5A protein triggers oxidative stress by inducing NADPH oxidases 1 and 4 and cytochrome P450 2E1. *Oxid. Med. Cell Longev.* **2016**, 1–10.
53. Barbaro G., Lorenzo G.D., Ribersani M., Soldini M., Giancaspro G., Bellomo G., Belloni G., Grisorio B., Barbarini G. (1999) Serum ferritin and hepatic glutathione concentrations in chronic hepatitis C patients related to the hepatitis C virus genotype. *J. Hepatol.* **30**, 774–782.
54. Venturini D., Simão A.N.C., Barbosa D.S., Lavado E.L., Narciso V.E.S., Dichi I., Dichi J.B. (2009) Increased oxidative stress, decreased total antioxidant capacity, and iron overload in untreated patients with chronic hepatitis C. *Dig. Dis. Sci.* **55**, 1120–1127.
55. Razzaq Z., Malik A. (2014) Viral load is associated with abnormal serum levels of micronutrients and glutathione and glutathione-dependent enzymes in genotype 3 HCV patients. *Biochim. Biophys. Acta Clin.* **2**, 72–78.
56. Müller L., Hainberger D., Stolz V., Ellmeier W. (2018) NCOR1—a new player on the field of T cell development. *J. Leukoc. Biol.* **104**, 1061–1068.
57. Sun Z., Xu Y. (2020) Nuclear receptor coactivators (NCOAs) and corepressors (NCORs) in the brain. *Endocrinology*. **161**, 1–12.
58. Kong Y., Zhou W., Sun Z. (2020) Nuclear receptor corepressors in intellectual disability and autism. *Mol. Psychiatry*. **25**, 2220–2236.
59. Emmett M.J., Lazar M.A. (2019) Integrative regulation of physiology by histone deacetylase 3. *Nat. Rev. Mol. Cell Biol.* **20**, 102–115.
60. Wingelhofer B., Neubauer H.A., Valent P., Han X., Constantinescu S.N., Gunning P.T., Müller M., Moriggl R. (2018) Implications of STAT3 and STAT5 signaling on gene regulation and chromatin remodeling in hematopoietic cancer. *Leukemia*. **32**, 1713–1726.
61. Bottomley M.J., Lo Surdo P., Di Giovine P., Cirillo A., Scarpelli R., Ferrigno F., Jones P., Neddermann P., De Francesco R., Steinkuhler C., Gallinari P., Carfi A. (2008) Structural and functional analysis of the human HDAC4 catalytic domain reveals a regulatory structural zinc-binding domain. *J. Biol. Chem.* **283**, 26694–26704.
62. Park S.Y., Kim G.S., Hwang H.J., Nam T.H., Park H.S., Song J., Jang T.H., Lee Y.C., Kim J.S. (2018) Structural basis of the specific interaction of SMRT corepressor with histone deacetylase 4. *Nucl. Acids Res.* **46**, 11776–11788.
63. Park S.Y., Kim J.S. (2020) A short guide to histone deacetylases including recent progress on class II enzymes. *Exp. Mol. Med.* **52**, 204–212.
64. Zhuang S. (2013) Regulation of STAT signaling by acetylation. *Cell. Signal.* **25**, 1924–1931.
65. Yuan Z., Guan Y., Chatterjee D., Chin E. (2005) STAT3 dimerization regulated by reversible acetylation of a single lysine residue. *Science*. **307**, 269–273.
66. Yu J., Li Y., Ishizuka T., Guenther M.G., Lazar M.A. (2003) A SANT motif in the SMRT corepressor interprets the histone code and promotes histone deacetylation. *EMBO J.* **22**, 3403–3410.
67. Yoon H.G., Chan D.W., Huang Z.Q., Li J., Fondell J.D., Jun Qin J., Wong J. (2003) Purification and functional characterization of the human N-CoR complex: the roles of HDAC3, TBL1 and TBLR1. *EMBO J.* **22**, 1336–1346.
68. Yoon H.G., Choi Y., Cole P.A., Wong J. (2005) Reading and function of a histone code involved in targeting corepressor complexes for repression. *Mol. Cell. Biol.* **25**, 324–335.
69. Giraud S., Bienvenu F., Avril S., Gascan H., Heery D.M., Coqueret O. (2001) Functional interaction of STAT3 transcription factor with the coactivator NcoA/SRC1a. *J. Biol. Chem.* **277**, 8004–8011.
70. Xu J., Wu R.C., O'Malle B.W. (2009) Normal and cancer-related functions of the p160 steroid receptor co-activator (SRC) family. *Nat. Rev. Cancer*. **9**, 615–630.
71. Jopling C.L. (2008) Regulation of hepatitis C virus by microRNA-122. *Biochem. Soc. Trans.* **36**, 1220–1223.
72. Zhou Y., Wang Q., Yang Q., Tang J., Xu C., Gai D., Chen X., Chen J. (2018) Histone deacetylase 3 inhibitor suppresses hepatitis C virus replication by regulating Apo-A1 and LEAP-1 expression. *Virol. Sin.* **33**, 418–428.
73. Kim K., Lee Y., Jeong S., Kim D., Chon S., Pak Y. K., Kim S., Ha J., Kang I., Choe W. (2020) A small molecule, 4-phenylbutyric acid, suppresses HCV replication

- via epigenetically induced hepatic hepcidin. *Int. J. Mol. Sci.* **21**(15), 5516.
74. Kozlov M.V., Konduktorov K.A., Shcherbakova A.S., Kochetkov S.N. (2019) Synthesis of N'-propylhydrazide analogs of hydroxamic inhibitors of histone deacetylases (HDACs) and evaluation of their impact on activities of HDACs and replication of hepatitis C virus (HCV). *Bioorg. Med. Chem. Lett.* **29**, 2369–2374.
 75. McClure J.J., Zhang C., Inks E.S., Peterson Y.K., Li J., Chou C.J. (2016) Development of allosteric hydrazide-containing class I histone deacetylase inhibitors for use in acute myeloid leukemia. *J. Med. Chem.* **59**, 9942–9959.
 76. Malikova A.Z., Shcherbakova A.S., Konduktorov K.A., Zemskaya A.S., Dalina A.A., Popenko V.I., Leonova O.G., Morozov A.V., Kurochkin N.N., Smirnova O.A., Kochetkov S.N., Kozlov M.V. (2021) Pre-senescence induction in hepatoma cells favors hepatitis C virus replication and can be used in exploring antiviral potential of histone deacetylase inhibitors. *Int. J. Mol. Sci.* **22**, 4559.

How Histone Deacetylase 3 Controls Hepcidin Expression and Hepatitis C Virus Replication

A. S. Shcherbakova¹, S. N. Kochetkov¹, and M. V. Kozlov^{1, *}

¹Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Moscow, 119991 Russia

*e-mail: kozlovmavi@gmail.com

The key role of histone deacetylases (HDACs) in the regulation of the cellular response to infection with the hepatitis C virus (HCV) was first demonstrated in 2008. Studying the metabolism of iron in the liver tissues of patients with chronic hepatitis C, the authors found that the expression of the hepcidin gene (*HAMP*), a hormone regulator of iron export, is markedly reduced in hepatocytes under conditions of oxidative stress caused by viral infection. HDACs were involved in the regulation of hepcidin expression through the control of acetylation level of histones and transcription factors, primarily STAT3, associated with the *HAMP* promoter. The purpose of this review is to summarize current data on the functioning of the HCV-HDAC3-STAT3-*HAMP* regulatory circuit as an example of a well-characterized interaction between the virus and the epigenetic apparatus of the host cell.

Keywords: hepatitis C virus, replication, STAT3, hepcidin, oxidative stress, histone deacetylase 3, expression regulation

УДК 57.088

МЕТОДЫ ОБОГАЩЕНИЯ ДНК-ФРАГМЕНТОВ ДЛЯ ВЫСОКОПРОИЗВОДИТЕЛЬНОГО СЕКВЕНИРОВАНИЯ

© 2023 г. А. Н. Синяков^a, *, Е. В. Костина^a

^aИнститут химической биологии и фундаментальной медицины, Сибирское отделение Российской академии наук, Новосибирск, 630090 Россия

*e-mail: sinyakov@niboch.nsc.ru

Поступила в редакцию 08.11.2022 г.

После доработки 16.12.2022 г.

Принята к публикации 16.12.2022 г.

В обзоре описано использование олигонуклеотидов, в основном полученных с помощью синтезаторов ДНК нового поколения (микрочиповых синтезаторов ДНК), для обогащения целевых геномных фрагментов. Рассмотрены способы обогащения целевых мишеней с использованием методов молекулярной гибридизации, ПЦР и системы CRISPR-Cas9. Приведены примеры практического применения разработанных методов для исследовательских и диагностических целей.

Ключевые слова: высокопроизводительное секвенирование, олигонуклеотиды, микрочипы, амплификация, гибридизация, диагностика, экзом, CRISPR-Cas9

DOI: 10.31857/S0026898423030126, **EDN:** CHPOFS

ВВЕДЕНИЕ

Несмотря на стремительное развитие технологии ДНК-секвенирования нового поколения (NGS), позволяющей определять полную нуклеотидную последовательность геномов, ее использование для прочтения определенных областей генома затруднительно. Это касается исследований, связанных с определением однонуклеотидных полиморфизмов (SNP), анализом перестройки ДНК, а также при изучении подмножеств участков генома, а не всего генома, например кодирующих белок экзонов. Для выполнения таких задач необходимо проводить “целевое обогащение” анализируемой ДНК и создание специализированных библиотек целевых последовательностей для последующего их секвенирования на выбранной платформе NGS.

Удаление той части геномной ДНК, которая не представляет интереса для проводимого исследования, позволяет обеспечить большую глубину покрытия при секвенировании целевых районов генома, увеличить число мишеней, а также значительно сэкономить время и стоимость эксперимента, в том числе за счет упрощения биоинформатического анализа результатов секвенирования.

Методы обогащения мишеней можно разделить на несколько категорий в зависимости от способа захвата целевого участка ДНК. В одних способах обогащения мишеней используют метод гибридизации нуклеиновых кислот. В этом

случае денатурированную геномную ДНК “отжигают” с одноцепочечными (оц) ДНК- или РНК-зондами к целевым участкам. Связавшиеся с зондами геномную ДНК выделяют и секвенируют. Другие способы обогащения мишеней основаны на методе ПЦР. Используют мультиплексную ПЦР с десятками тысяч праймеров. Эмульсионная ПЦР позволяет одновременно проводить миллионы цепных реакций и получать миллионы мишеней для секвенирования. В последние годы появились методы обогащения целевых фрагментов ДНК, основанные на использовании РНК-программируемого комплекса CRISPR-Cas9.

СПОСОБЫ ОБОГАЩЕНИЯ ЦЕЛЕВЫХ ФРАГМЕНТОВ ДНК С ПОМОЩЬЮ МЕТОДА МОЛЕКУЛЯРНОЙ ГИБРИДИЗАЦИИ

На первом этапе проводят фрагментацию (физическую или ферментативную) геномной ДНК с последующей ферментативной репарацией концов полученных ДНК-дуплексов и пришивкой к ним специфичных для секвенирующей платформы адаптеров. Адаптеры содержат в своем составе “штрих-код” — уникальную последовательность для каждого образца. После секвенирования биоинформатический анализ позволяет с помощью этого штрих-кода группировать данные, относящиеся к одному образцу. Производители наборов для целевого обогащения ДНК поставляют от 96

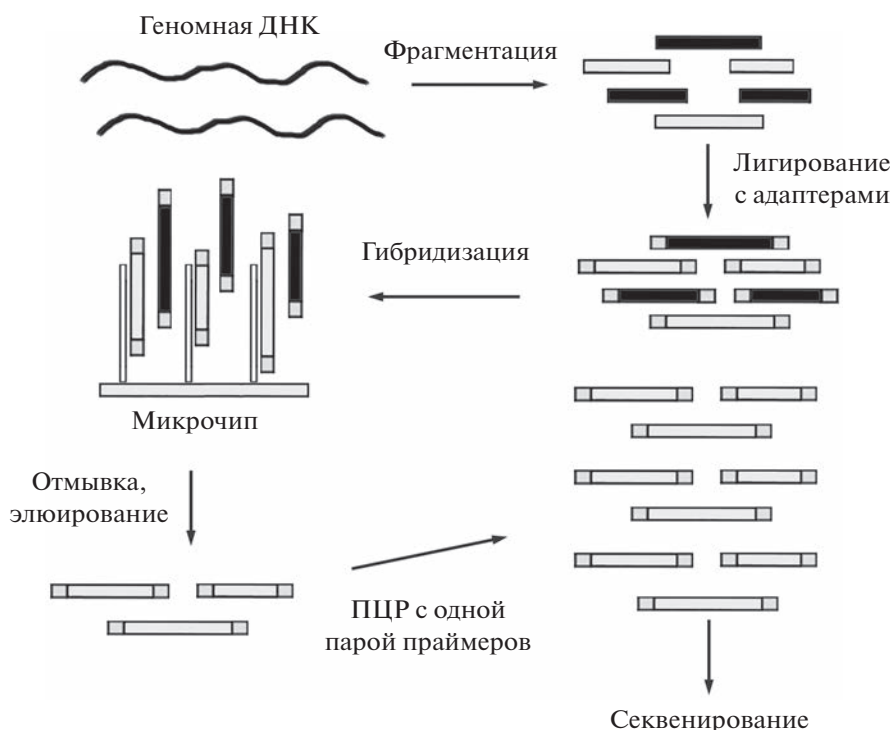


Рис. 1. Схема обогащения геномных фрагментов ДНК с помощью слайдов, содержащих олигонуклеотидные зонды (метод MGS).

до 384 различных адаптеров, содержащих различные штрих-коды. Использование штрих-кодов обеспечивает одновременное секвенирование нескольких образцов и тем самым значительное сокращение временных и финансовых затрат.

Кроме штрих-кода в адаптерах содержатся последовательности праймеров для амплификации. С помощью ПЦР нарабатывается достаточное количество ДНК для последующей стадии гибридизации с одноцепочечными зондами, комплементарными целевым областям генома. Как правило, это ДНК-зонды. Фирма “Agilent” (США) использует в своих наборах SureSelect гибридационные РНК-зонды. Основное преимущество РНК-зондов — большая стабильность РНК:ДНК-гибридов по сравнению с ДНК:ДНК-дуплексами [1].

Гибридизация на твердой поверхности

М. Zwick с сотрудниками [2] первыми использовали слайды “Affymetrix Chip Design Group” (США) с ковалентно связанными олигонуклеотидами для селекции геномных фрагментов (метод MGS, microarray-based genomic selection).

Протокол MGS включает пять основных этапов (рис. 1):

1) расщепление геномной ДНК на случайные фрагменты длиной 300 п.н;

2) репарация концевых фрагментов ДНК с добавлением 3'-выступающих остатков аденина и присоединение уникальных адаптеров с 5'-выступающим тимидином;

3) гибридизация фрагментов ДНК с комплементарными олигонуклеотидами на микрочипе высокой плотности;

4) элюирование фрагментов ДНК, связанных с зондами;

5) амплификация выделенных фрагментов с помощью праймеров, интегрированных в адаптеры.

В методе использовали перекрывающиеся гибридационные зонды к обеим цепям дуплекса, при этом каждая цепь перекрывалась зондами от 1.5 до 4 раз. На одном слайде содержалось 385 тыс. гибридационных зондов длиной от 50 до 93 нуклеотидов, позволяющих проводить изотермическую гибридизацию. Зонды покрывали фрагмент генома размером от 4 до 5 млн.п.н. Позднее их число увеличили до 2.1 млн. Эти зонды потенциально могут связаться с 34 млн.п.н. геномной последовательности [3, 4].

Аналогичным методом [5] выделяли более 200 тыс. экзонов для секвенирования. В качестве гибридационных зондов были выбраны олигонуклеотиды к кодирующей последовательности экзона длиной >60 н., перекрывающие последовательность экзона со сдвигом 20 н. Всего использовали в работе 7 слайдов (“Roche NimbleGen”, США), каж-

дый из которых содержал 385 тыс. зондов. В зависимости от применяемого протокола от 55 до 85% выделенных фрагментов ДНК содержали целевые последовательности. Метод позволяет секвенировать до 98% всех целевых экзонов.

Т. Albert и соавт. [6] ввели в схему дополнительную процедуру — амплификацию фрагментов ДНК перед гибридизацией на микрочипе. Секвенирование всего экзона с использованием гибридизации на микрочиповых слайдах доказало свою полезность для выявления редких вариантов генов и мутаций, вызывающих заболевания [7, 8]. Для обогащения 1000 генов из транскриптома хлопка A. Salmon и др. [9] использовали 135 тыс. зондов, содержащихся на слайдах “Roche NimbleGen”.

Nelson с сотрудниками [10] удалось повысить эффективность обогащения целевых мишеней при блокировании побочной гибридизации адаптерных частей зондов с геномной ДНК и дублировании стадии гибридизационного обогащения целевых мишеней.

Для выделения последовательностей генов *BRCA1* и *TP53* в микрофлюидных кассетах микрочипового ДНК-синтезатора Geniom (“Geniom”, Германия) синтезировали 50-членные олигонуклеотиды [11, 12]. На основе этого синтезатора разработали автоматизированный метод HybSelect для подготовки образцов для высокопроизводительного секвенирования. Метод позволяет параллельно проводить обогащение ≤8 образцов размером от 125 т.п.н. до 1 млн.п.н. Метод был апробирован на анализе 115 генов, ассоциированных с развитием злокачественных опухолей [13].

Несмотря на удовлетворительную производительность, разработанный метод HybSelect не лишен существенных недостатков. Во-первых, для проведения селективного выделения целевых дуплексов требуется много (от 10 до 15 мкг) исходного материала ДНК. Во-вторых, необходимость дорогостоящего специализированного оборудования (станции для проведения гибридизации). В-третьих, метод хорошо работает при обогащении геномных фрагментов размером ~500 п.н., но значительно менее эффективен в случае коротких экзонов (~120 п.н.) [5]. В-четвертых, HybSelect сложно масштабировать и автоматизировать.

Гибридизация в растворе

Из-за недостатков обогащения целевых фрагментов ДНК на твердой поверхности “Roche NimbleGen” и другие компании перешли на обогащение мишеней в растворе, где гибридизация идет с избытком гибридизационных зондов. Фирма “Roche NimbleGen” прекратила выпуск слайдов для гибридизации геномных фрагментов непосредственно на микрочипе.

Компания “Agilent” была первой, разработавшей коммерческий продукт, использующий гибридизацию олигонуклеотидов с геномными фрагментами ДНК в растворе [14]. В этом методе применяют синтезированные на микрочиповом ДНК-синтезаторе 200-членные олигонуклеотиды, содержащие 170-членную мишеньспецифическую последовательность, фланкированную двумя праймерами для амплификации. Набор синтезированных олигонуклеотидов отщепляют от слайда и подвергают двукратной ПЦР. В первом раунде используют интегрированные в последовательность олигонуклеотидов праймеры, во втором — в мишеньспецифические дуплексы вводят последовательность промотора фага T7. Затем, используя транскрипцию *in vitro* в присутствии биотин-УТР, получают биотинилированные оцРНК-зонды для выделения целевых геномных фрагментов. Геномную ДНК случайным образом расщепляют на фрагменты длиной ~250 п.н., к которым присоединяют адаптеры. После 12 циклов ПЦР полученные ампликоны гибридизуют с РНК-зондами в растворе. Биотинилированные РНК, содержащие и не содержащие комплементарную геномную ДНК, выделяют из реакционной смеси с помощью магнитных шариков с иммобилизованным стрептавидином. ДНК:РНК-гибриды разрушают и полученную ДНК амплифицируют с помощью ранее введенных в состав фрагментов геномной ДНК адаптеров (рис. 2).

Разработанный метод не требует специального оборудования и совместим с различными секвенирующими платформами. Он был апробирован на секвенировании 15 тыс. экзонов (2.5 млн.п.н.) человеческого генома.

К основным достоинствам метода относятся:

- 1) высокая концентрация оцРНК-зондов, что позволяет повысить эффективность процесса гибридизации с целевыми ДНК-фрагментами;
- 2) относительно небольшое количество геномной ДНК (0.5–3.0 мкг) для секвенирования;
- 3) проведение гибридизации в растворе, что позволяет масштабировать и автоматизировать процесс в отличие от твердофазной гибридизации на слайдах [2];
- 4) нивелирование различий в гибридизации разных аллелей благодаря использованию протяженных зондов;
- 5) возможность готовить и характеризовать большие партии РНК-зондов, что позволяет иметь стандартизованный материал для использования в большом числе экспериментов;
- 6) эффективность при таргетном секвенировании как множества несмежных геномных фрагментов, так и протяженных геномных районов;
- 7) высокая специфичность — успешно определяют 85–90% целевых последовательностей.

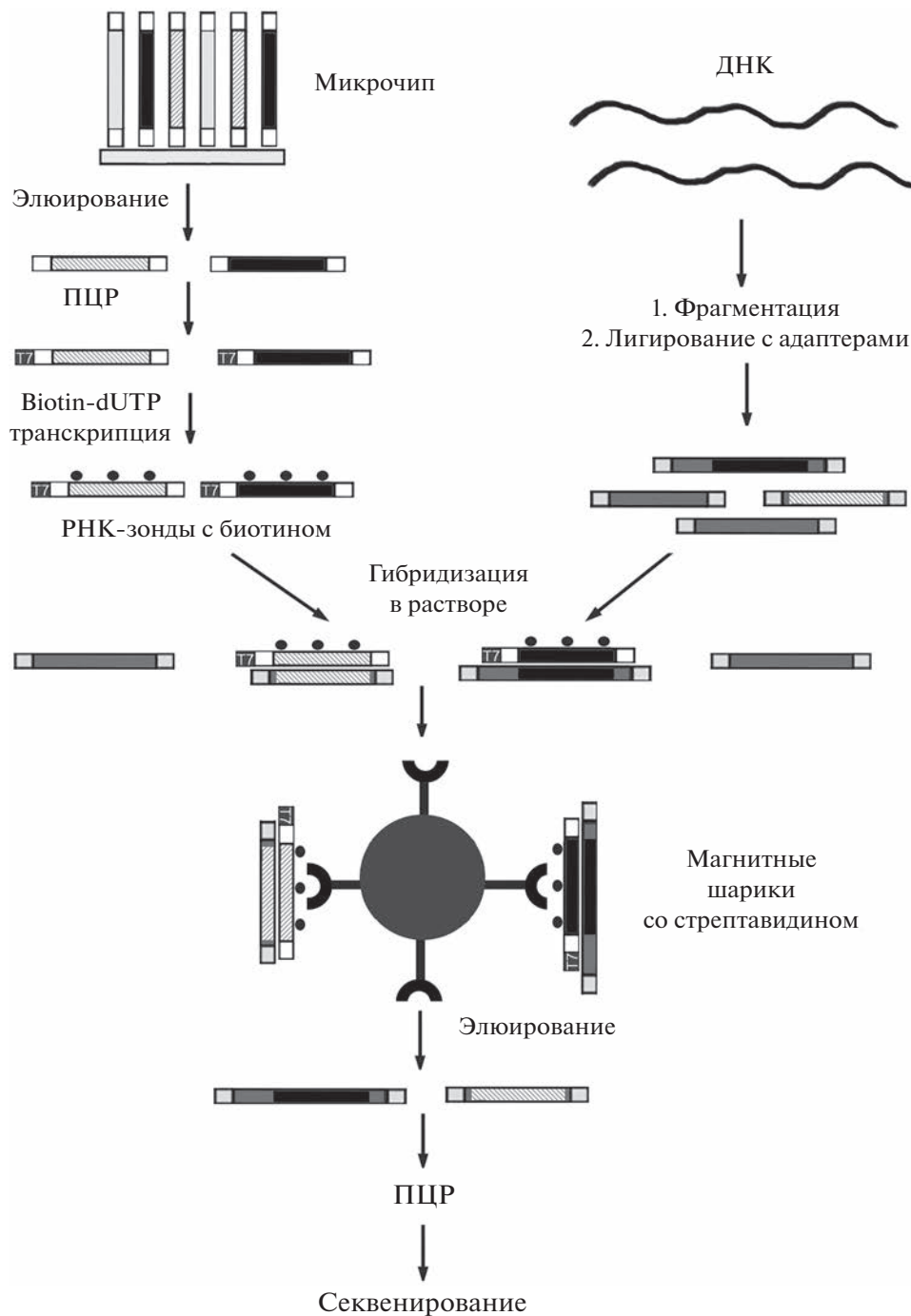


Рис. 2. Схема обогащения геномных фрагментов ДНК с помощью олигонуклеотидных зондов в растворе, применяемая в наборах Agilent.

Созданы коммерческие наборы, основанные на различных вариантах обогащения последовательностей ДНК в растворе: например, Agilent SureSelect (“Agilent”) [15], Illumina TruSeq (“Illumina”, США) [16], Roche NimbleGen SeqCap EZ (“Roche NimbleGen”) [17]. Принципиальная разница между “Agilent” и другими коммерческими наборами заключается в природе гибридизационных

зондов: “Agilent” использует 120–170-членные РНК-пробы, в то время как “Roche NimbleGen” 60–90-членные, а “Illumina” 95-членные ДНК-зонды [3, 16]. Для процедуры проведения обогащения геномной ДНК в растворе применяются стандартные 96-луночные планшеты и термоциклер, то есть специализированного оборудования в этом случае не требуется.

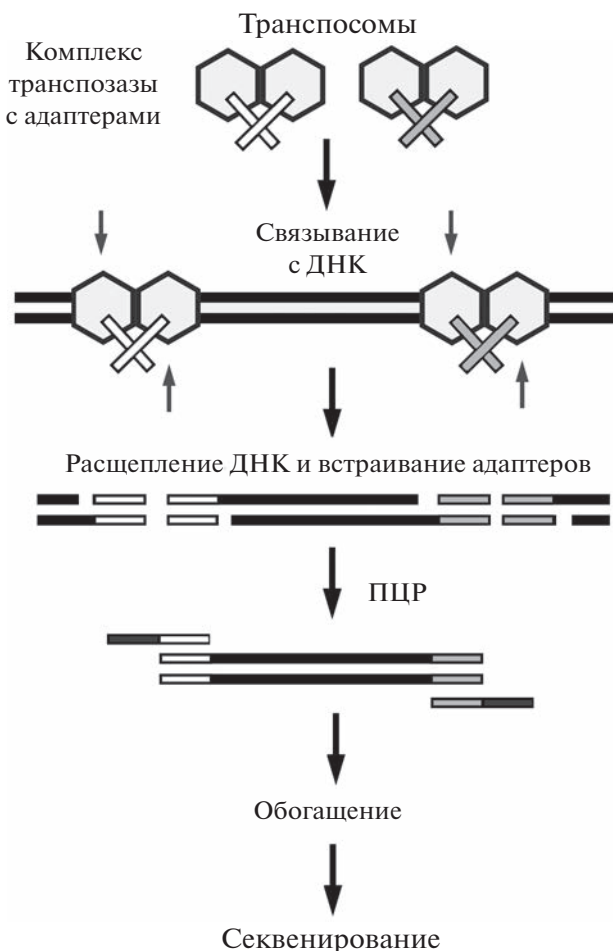


Рис. 3. Схема использования транспозоноопосредованной фрагментации анализируемой ДНК. Комплекс транспозазы с двухцепочечными олигонуклеотидами, содержащими адаптерные участки, связывается с геномной ДНК. Транспозаза расщепляет двухцепочечную геномную ДНК и встраивает олигонуклеотиды на участке разрезания. Полученные молекулы ДНК используют в ПЦР с праймерами, необходимыми для секвенирования.

Модификация метода обогащения целевых фрагментов путем гибридизации олигонуклеотидов с геномными фрагментами ДНК в растворе — транспозоноопосредованная фрагментация анализируемой ДНК (рис. 3).

Данная модификация значительно упрощает процесс получения фрагментов анализируемой ДНК, поскольку фрагментация и присоединение адаптеров происходит за одну стадию с применением фермента транспозазы [18, 19]. Транспозомы (комплекс фермента транспозазы с двухцепочечными олигонуклеотидами, содержащими участки связывания с транспозазой и часть последовательности праймера для последующей амплификации) обладают способностью случайным образом связываться с целевыми последовательностями

ми двухцепочечной геномной ДНК. Транспозазы в транспозоме расщепляют ДНК и одновременно встраивают на участке разрезания олигонуклеотиды из комплекса.

Разработанные методы обогащения экзома оказались эффективными для обнаружения мутаций, служащих причиной редких менделевских заболеваний [20–23], сложных расстройств [24–27], митохондриальных нарушений [28–30]. Также их применяли при скрининге потенциальных генетических мутаций, связанных с раковыми заболеваниями [31–36].

Метод гибридизационного обогащения успешно используют при анализе древней ДНК человека (ancient DNA; аДНК), сохранившейся в человеческих останках. Первым генетическим маркером, проанализированным в палеогенетических исследованиях человека, была митохондриальная ДНК (мтДНК), так как в клетках ее копияемость гораздо выше, чем ядерной ДНК. Для гибридизации использовали биотинилированные ДНК или РНК-зонды, направленные на два гипервариабельных сегмента контрольной области мтДНК [37–43].

Селективное обогащение фрагментов Y-хромосомы аДНК проводили как на твердой поверхности [44], так и в растворе [45].

Carpenter и соавт. [46] разработали метод полногеномного обогащения (whole-genome in-solution capture, WISC) аДНК, используя современную ДНК человека в качестве зондов.

В настоящее время разработаны коммерческие наборы, ориентированные на мтДНК, индивидуальные локусы или целые геномы, такие как myBaitsR³ (“Daicel Arbor Biosciences”, США), которые используют для секвенирования аДНК [47, 48].

Применение смеси олигонуклеотидов, полученных с помощью микрочипового синтезатора, содержащей 962438 зондов, позволило секвенировать экзоны бурундуков [49]. Из целевых 11975 экзонов обогатить удалось более 99%. Зонды были рассчитаны на основе известной последовательности генома бурундука *Tamias alpinus*. Однако эти же зонды были успешно использованы для селективного выделения экзонов родственных видов: *Tamias amoenus*, *Tamias ruficaudus* и *Tamias striatus*.

Аналогичным образом, используя зонды для экзонов генов домашней коровы (*Bos taurus*), T. Cosart и др. [50] секвенировали 16131 экзон быка (*Bos indicus*) и дикого бизона (*Bison bison*). Предложенный метод селективного обогащения геномной ДНК применяли для секвенирования ряда хромосом и митохондриального генома западных шимпанзе (*Pan troglodytes verus*) [51]. R. Tethley и соавт. [52] для гибридизации использовали перекрывающиеся РНК-зонды длиной 120 нуклеотидов. На мишени размером 3.9 млн.п.н. они

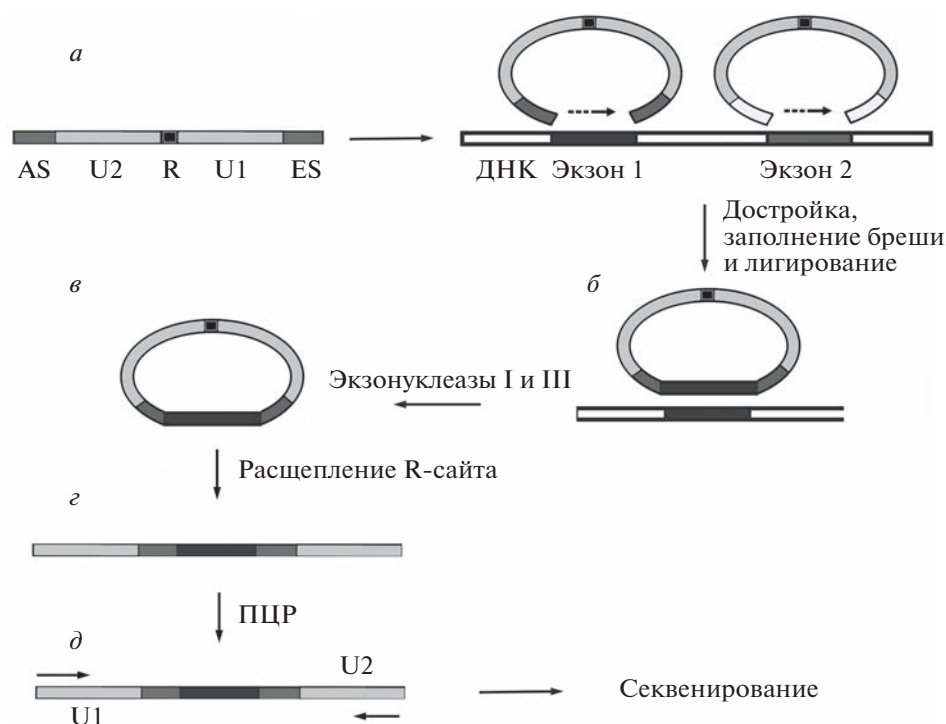


Рис. 4. Схема обогащения целевых фрагментов ДНК с использованием инверсионных зондов CIPer. а — Структура CIPer. AS и ES — 20-звенные олигонуклеотиды, гомологичные целевым фрагментам геномной ДНК; R — сайт рестрикции; U1 и U2 — универсальные праймеры для амплификации инвертированного зонда. б — Циклизация CIPer путем заполнения бреши ДНК-полимеразой и лигирования. в — Обогащение кольцевой ДНК за счет деградации линейной формы ДНК экзонуклеазами I и III. г — Линейаризация кольцевой формы ДНК путем расщепления сайта рестрикции. д — Амплификация полученной линейной формы ДНК с праймерами U1 и U2.

картировали 93% мутаций с точностью более 99%. Подход Agilent был успешно применен также при проведении секвенирования мутаций в генах, ответственных за потерю слуха [53].

С помощью гибридизации в растворе с ДНК-олигонуклеотидами были получены для секвенирования образцы ретровирусов, интегрированные в геном человека [54].

Добавление коммерческого препарата *Cot1* DNA при проведении гибридизации, как на поверхности, так и в растворе, увеличивает ее специфичность [14, 55]. *Cot1* DNA состоит из коротких фрагментов (50–300 п.н.) плацентарной ДНК человека, обогащенной повторяющимися последовательностями. При гибридизации на слайде или в растворе *Cot1* DNA обычно добавляют соответственно в 5- и 20-кратном избытке по отношению к геномной ДНК.

СПОСОБЫ ОБОГАЩЕНИЯ ЦЕЛЕВЫХ ФРАГМЕНТОВ ДНК, ОСНОВАННЫЕ НА ИСПОЛЬЗОВАНИИ МЕТОДА МОЛЕКУЛЯРНОЙ ИНВЕРСИИ

Разработанные М. Nilsson с соавт. [56] инверсионные зонды в сочетании с мультиплексной ПЦР нашли широкое применение при детектировании

точечных мутаций. Так, М. Akhras и др. [57] использовали их для обогащения мишеней NGS. Схема применения таких зондов (Connector Inversion Probe; CIPer) приведена на рис. 4. Разработанный метод был успешно апробирован авторами при генотипировании вируса папилломы человека (HPV), ассоциированных с онкологическими заболеваниями вирусов, скрининге мутаций антибиотикорезистентности в геноме гонококка (*Neisseria gonorrhoeae*).

Метод коммерциализирован под названием HaloPlex (“Halo Genomics”, Швеция). В настоящее время набор производит компания “Agilent Technologies”. Их использование позволяет выделять 21 тыс. генов общим размером в 37 млн.п.н. Зонды такого типа получили общее название “молекулярные инверсионные пробы” (molecular inversion probes, MIP).

Способ обогащения целевой ДНК с использованием MIP был применен для исследования большой группы пациентов с дистонией на наличие патогенных и редких вариантов гена *ATM* [58], для анализа мутаций гена *GNAS*, вызывающих внутримышечную миксому [59].

Применение MIP в совокупности с таргетным РНК-секвенированием нового поколения позволило авторам создать независимую от гистологии

молекулярную диагностику опухолей головного мозга [60]. Аналогичный подход использовали для анализа профилей экспрессии генов в случае почечно-клеточного рака [61].

Применение МІР для амплификации вариабельных областей V1, V2, V3, V6 и V7 гена 16S рРНК и последующее NGS полученных ампликонов позволило создать метод диагностики бактериальных возбудителей [62]. МІР оказались эффективными при создании автоматизированного процесса анализа мутаций генов *BRCA1* и *BRCA2* с целью диагностики рака молочной железы [63]. Использование МІР при секвенировании гена *SCN1A* позволило создать метод диагностики родительского мозаицизма при синдроме Даве [64]. Аналогично был проведен анализ мозаичности гена *PIK3CA* при нарушении развития [65].

Y. Wada и соавт. [66], используя МІР, выявляли в генах семейства PPAR/RXR мутации, ассоциированные с развитием шизофрении. NGS с использованием МІР применяли при анализе мутаций генов ионных каналов при диабетической периферической нейропатии [67] и для анализа мутаций в генах *QRDR* [68], определяющих устойчивость к ципрофлоксацину в бактериях, вызывающих сибирскую язву, чуму и туляремию. Разработана методика NGS с применением МІР, предназначенная для исследований в области фармакогеномики [69].

По сравнению с другими способами обогащения мишеней методы на основе МІР имеют ряд преимуществ. Они высокоспецифичны к мишеням, работают с мизерным количеством (≤ 5 пг) геномной ДНК и не требуют этапов предварительной обработки, таких как фрагментация ДНК. Кроме того, при использовании для NGS ограниченного числа МІР можно применять для их синтеза традиционные колоночные и планшетные синтезаторы.

Для создания тысяч МІР, захватывающих большое число целей одновременно, обычно используют микрочиповые синтезаторы. При получении одноцепочечных олигонуклеотидов с помощью микрочиповых синтезаторов необходимо провести ряд последовательных стадий, таких как расщепление эндо- или экзонуклеазами, очистку полученных продуктов в геле и амплификацию. D. Bang с сотрудниками [70] разработали значительно более простой и эффективный протокол применения МІР из синтезированных на микрочиповых синтезаторах олигонуклеотидов — с использованием только одного фермента и двухцепочечных МІР-зондов. Метод был успешно апробирован на 11 510 дуплексных МІР, предназначенных для выделения 3554 экзонов 228 генов в образце геномной ДНК.

СПОСОБЫ ОБОГАЩЕНИЯ ЦЕЛЕВЫХ ФРАГМЕНТОВ ДНК С ПОМОЩЬЮ ПЦР

Моноплексная ПЦР

ПЦР с момента появления в 1980 годах предназначалась для наработки целевого фрагмента ДНК с помощью праймеров для амплификации [71]. Именно поэтому технология ПЦР была применена для NGS. В моноплексной ПЦР в одной пробирке с помощью одной пары праймеров нарабатывают только один ДНК-дуплекс. После амплификации всех целевых фрагментов их объединяют в эквимольных количествах для создания библиотеки для секвенирования. Получение ампликонов с помощью моноплексной ПЦР может быть очень трудоемким процессом, поэтому эта стадия была автоматизирована. Например, для упрощения процедуры получения ампликонов компания “Fluidigm” (Канада) разработала микрофлюидное устройство Access Array system, позволяющее автоматизировать процесс получения целевых ампликонов. Одновременно можно получить 480 различных ампликонов [72].

Эмульсионная технология RainDance (“Bio-Rad”, США) [73], использующая моноплексную ПЦР, значительно более производительна по сравнению с описанной выше. В ней генерируют миллионы капель двух типов: один содержит геномную ДНК (1 молекула геномной ДНК, дезоксинуклеозидтрифосфаты, ДНК-полимераза в буфере), а другой только молекулы одной пары праймеров. Затем эти типы капель смешивают в соотношении 1 : 1. Примерно 1.5 млн капель ПЦР собирают в одну пробирку для ПЦР объемом 0.2 мл, амплифицируют в стандартном термоциклере с последующим разрушением эмульсии с высвобождением ампликонов в раствор. После очистки смесь секвенируют.

Технологии как RainDance, так и Fluidigm совместимы с большинством коммерческих секвенаторов и требуют только несколько наногرامмов исходного материала для получения ампликонов для секвенирования. Однако технология RainDance позволяет получать значительно больше ампликонов одновременно: до 20 тыс. на образец по сравнению с 480 для Fluidigm. В продуктах компании “Bio-Rad” есть панели ампликонов, связанных с мутациями в генах, ассоциированных с онкологическими заболеваниями и аутизмом.

Мультиплексная ПЦР

Мультиплексная ПЦР — реакция, в которой несколько пар праймеров одновременно генерируют несколько ампликонов из одного и того же исходного материала. До недавнего времени использование мультиплексной ПЦР было ограничено — из-за высокого уровня неспецифической

амплификации. Однако в последние годы несколько компаний разработали алгоритмы, позволяющие значительно повысить селективность мультиплексной ПЦР.

Компания “Illumina” разработала вариант мультиплексной ПЦР, включающий следующие стадии:

- 1) гибридизацию олигонуклеотидов к их комплементарным последовательностям в интересующей области;
- 2) достраивание одного из пары олигонуклеотидов;
- 3) сшивание достроенного фрагмента со вторым олигонуклеотидом;
- 4) ПЦР-амплификацию полученного фрагмента ДНК с использованием универсальных праймеров, содержащих индексные последовательности.

Разработанный подход позволяет создавать разнообразные панели, в том числе и нестандартные. Пользовательские панели могут быть созданы для 1536 мишеней в геномах человека, мыши, крысы и коровы. Сконструированы панели TruSeq Amplicon Cancer Panel (“Illumina”), нацеленные на гены, связанные с развитием рака. Одна из них создана для секвенирования 212 ампликонов 48 ассоциированных со злокачественными опухолями генов. TruSight Tumor Panel (“Illumina”) разработана для получения 174 ампликонов, покрывающих сайты мутаций в 26 ассоциированных с развитием рака генах [74].

Компания “Thermo Fisher Scientific” (США) разработали технологию Ion AmpliSeq, позволяющую проводить до 24 тыс. амплификаций в одной пробирке (рис. 5).

В протоколе этого метода предусмотрено использование всего 1 нг ДНК и РНК. Технология AmpliSeq получила широкое распространение во всем мире. Так, ее использовали в исследованиях злокачественных опухолей [75], наследственных заболеваний [76] и бактерий [77].

На основе технологии AmpliSeq разработаны наборы для исследования либо определенных генов (например, Ion AmpliSeq *TP53* Panel, Ion AmpliSeq *BRCA1* и *BRCA2* Panel), либо заболеваний (например, муковисцидоза, деменции, рака толстой кишки и легкого). Создан набор для полноэкзомного секвенирования, содержащий 294 тыс. пар праймеров.

СПОСОБЫ ОБОГАЩЕНИЯ ЦЕЛЕВЫХ ФРАГМЕНТОВ ДНК С ПОМОЩЬЮ СИСТЕМЫ CRISPR-Cas9

В последние годы стали появляться методы обогащения целевых фрагментов ДНК, основанные на использовании РНК-программируемого ком-

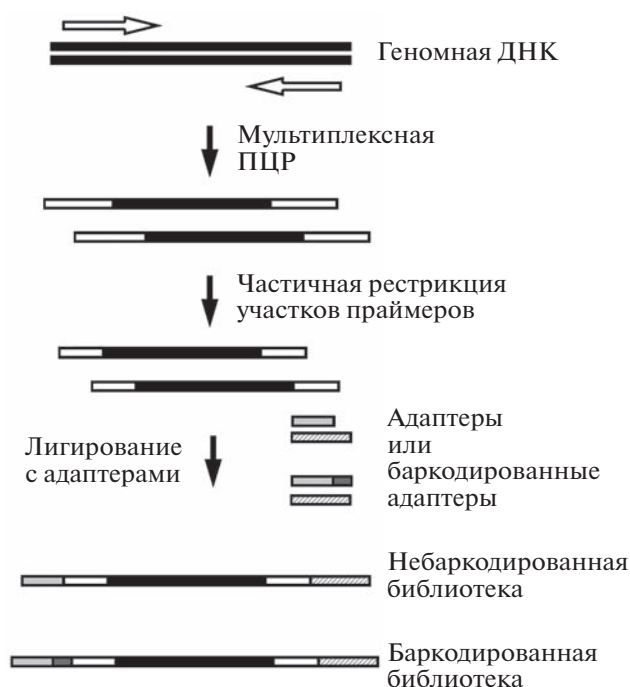


Рис. 5. Схема подготовки библиотек для секвенирования по технологии AmpliSeq.

плекса CRISPR-Cas9 [78, 79]. Системы CRISPR-Cas обеспечивают бактериям приобретенный иммунитет к вирусам. CRISPR-ассоциированный белок Cas9 — это эндонуклеаза, которая использует направляющую последовательность (РНК-гид), чтобы сформировать комплекс с ДНК-мишенями, что позволяет Cas9 произвести сайтспецифичный двухцепочечный разрыв в ДНК (рис. 6). Эта система направленно действует на любую последовательность геномной ДНК.

Истощение фоновых ДНК

Прежде всего, система CRISPR-Cas была использована для обогащения целевых ДНК путем истощения нежелательных. Метод обогащения получил название DASH (Depletion of Abundant Sequences by Hybridization) [80]. На первом этапе конструируют библиотеку фрагментов ДНК для секвенирования, содержащую в частности праймеры для ее амплификации. Разрушение ненужных для секвенирования последовательностей в полученной библиотеке проводят с помощью РНК-гидов, направленных на неинформативные последовательности. После амплификации дуплексов, содержащих в своем составе оба амплификационных праймера, библиотеку секвенируют. Используя этот метод, J. DeRisi с сотрудниками [80] удалось снизить содержание митохондриальной рРНК в клетках HeLa на два порядка и обогатить содержание целевых последовательностей

ются фрагменты ДНК произвольного размера, что может приводить к проблемам при секвенировании, связанным с низким уровнем “покрытия”, неравномерным “покрытием” и ложными мутациям. Подход, названный CRISPR-DS, успешно использован для секвенирования экзонных областей *TP53* [85]. Целевые фрагменты ДНК для секвенирования могут быть выделены с помощью простого фракционирования по молекулярному весу, что приводит к обогащению примерно в 49 тыс. раз. Метод позволяет снизить на один–два порядка массу исследуемой ДНК.

На основе системы CRISPR-Cas быстро растет разнообразие методов обогащения целевых последовательностей ДНК, однако до сих пор ни один из них не достиг эффективности таковых с применением гибридизации нуклеиновых кислот.

ПРИМЕРЫ ИСПОЛЬЗОВАНИЯ ТАРГЕТНОГО СЕКВЕНИРОВАНИЯ В МЕДИЦИНСКОЙ ДИАГНОСТИКЕ

Методы NGS быстро развиваются и стремительно внедряются в практическую медицину. Многие заболевания человека имеют генетическую основу. Некоторые из них обусловлены отсутствием или дисфункцией определенного белка из-за мутаций в кодирующем гене. Так обстоит дело с заболеваниями менделевской наследственности, такими как болезнь Хантингтона, талассемия и около тысячи других наследственных редких заболеваний [97]. Все больше генетических вариантов и полиморфизмов идентифицируют в качестве факторов риска орфанных заболеваний [98]. Часто развитие злокачественных новообразований тоже опосредовано генетически: мутацией одного или нескольких генов, которые либо повышают риск злокачественной трансформации клеток (например, мутации зародышевой линии), либо способствуют развитию опухоли (онкогены), либо нарушают клеточные механизмы, контролирующие пролиферацию клеток (гены-супрессоры), как это происходит при соматических мутациях [99]. Таргетное или полноэкзомное секвенирование генома пациента позволяют поставить диагноз и выработать адекватную тактику лечения.

Использование в медицинской диагностике NGS не ограничено генетическими заболеваниями. Высокопроизводительное секвенирование применяют и в исследованиях инфекционных заболеваний человека [100]. Эту технологию эффективно используют для выявления малярийных плазмодиев (*Plasmodium spp.*) и возбудителей микозов. Так, для обнаружения генома *Plasmodium falciparum* на фоне геномной ДНК человека применяли метод гибридизационного обогащения в растворе [101, 102]. Предложенный подход был распространен на анализ *Plasmodium vivax* [103] и

Candida albicans [104]. Предварительное обогащение генома бактерий с последующим секвенированием может быть очень полезным в клинической практике. А. Brown с соавт. [105] использовали NGS для определения устойчивости *Mycobacterium tuberculosis* H37Rv к противомикробным препаратам.

Применив для полногеномного обогащения РНК-зонды в растворе, М. Christiansen и др. [106] описали быстрый способ обнаружения и анализа генома *Chlamydia trachomatis*, превышающий на порядок по чувствительности другие известные методы. Аналогичным образом предлагают проводить мониторинг некультивируемого менингококка *Neisseria meningitidis* в клинических образцах [107].

В вирусологических исследованиях технология NGS находит применение при выявлении лекарственной устойчивости патогенов, разработке новых лекарственных средств и вакцин [108].

Первым технологию целевого обогащения вирусных геномов из клинических образцов использовали D. Depledge с соавт. [109]. С помощью 120-членных РНК-зондов авторы провели обогащение геномов вирусов семейства *Herpesviridae*: варицелла-зостер, Эпштейна–Барр и герпесвируса, ассоциированного с саркомой Капоши. Полноразмерные геномы этих вирусов были реконструированы и использованы в работах по исследованию структуры и разнообразию вирусной популяции. Применяя предложенный подход, многие авторы исследовали геномы вируса Эпштейна–Барр [110], вируса Зика [111], норовируса [112], вируса Ласса [113], вируса гепатита С [114], герпесвируса-7 человека [115], вирусов простого герпеса-1 и -2 [116].

Т. Wylie и др. [117] создали гибридизационную панель ViroCap, предназначенную для обогащения нуклеиновой кислоты ДНК- и РНК-содержащих вирусов 34 семейств, которые инфицируют позвоночных и включают 190 вирусных родов и 337 видов. Панель была создана на основе анализа 1×10^9 геномных последовательностей, а ее применение позволило увеличить число прочтенных вирусных геномов в 296–674 раз. Разработанная панель может быть востребована как в фундаментальных, так и в прикладных исследованиях.

Т. Briese и др. [118] разработали платформу ViroCapSeq-VERT для секвенирования виroma. В систему входит около 2 млн зондов, которые охватывают геномы вирусов, поражающих позвоночных, включая человека. Список всех родов вирусов, которые инфицируют позвоночных, был составлен из списка основных видов Международного комитета по таксономии вирусов (International Committee on Taxonomy of Viruses). Всего для выбора зондов было использовано 342438 белоккодирующих вирусных последовательностей.

Библиотека биотинилированных олигонуклеотидов была синтезирована на основе массивов данных NimbleGen (NimbleGen Systems Inc., США) и использована для захвата в растворе вирусных нуклеиновых кислот. Использование VirCapSeq-VERT привело к 100–10000-кратному увеличению числа прочтений вирусного генома в образцах крови и гомогенатов тканей по сравнению с секвенированием Illumina. Для технологии VirCapSeq-VERT предел обнаружения сравним со специфической ПЦР в реальном времени в сыворотке, крови и экстрактах тканей. Кроме того, этот метод позволяет идентифицировать новые вирусы, геномы которых примерно на 40% отличаются от известных вирусов, использованных для разработки библиотеки зондов.

В 2018 году был разработан аналогичный подход — ViroFind [119]. Библиотека ViroFind включает 165433 вирусных зонда, которые охватывают геномы 535 отобранных ДНК- и РНК-вирусов, инфицирующих человека или способных вызывать зооноз. ViroFind использовали для обнаружения и анализа всех вирусных популяций в головном мозге пяти пациентов с прогрессирующей многоочаговой лейкоэнцефалопатией. По сравнению с прямым глубоким секвенированием, с помощью ViroFind удалось обогатить вирусные последовательности, присутствующие в клинических образцах, до 127 раз. Предложенный подход был распространен на анализ ретровирусных геномов (вирус Т-клеточного лейкоза типа 1 человека и вирус иммунодефицита типа 1 человека), интегрированных в геном человека [54]. При использовании вирусспецифичных зондов обогащение геномных последовательностей вирусов в сотни и тысячи раз превосходило этот показатель при прямом секвенировании.

ЗАКЛЮЧЕНИЕ

Наиболее эффективными методами таргетного обогащения в настоящее время считаются основанные на гибридизации в растворе, мультиплексной ПЦР и молекулярной инверсии. Разработаны эффективные компьютерные программы, позволяющие проводить подбор олигонуклеотидов для таргетного и полногеномного секвенирования. Ряд фирм производит наборы для полноэкзомного обогащения, которые при этом постоянно совершенствуются. Эти методы основаны на гибридизации в растворе. Проведены сравнительные анализы специфичности и производительности таких наборов [120–123].

В целом, все платформы работают хорошо, но между ними выявляют и небольшие различия. Так, РНК-зонды образуют более прочные комплексы с мишенью по сравнению с ДНК-зондами. При сверхглубоком секвенировании РНК-зонды, направленные на обе цепи мишени, работают во

всех случаях лучше, чем РНК-зонды направленные только на одну цепь мишени. ДНК-зонды более эффективны для анализа GC-богатых участков, а РНК-зонды — для AT-областей.

С развитием методов таргетного и полноэкзомного секвенирования появляются новые методы медицинской диагностики. Для развития высокоэффективных современных методов диагностики, основанных на использовании NGS, требуется не только инструментальная база (секвенаторы ДНК), но и современная индустрия для получения массива синтетических олигонуклеотидов (сотен тысяч и миллионов праймеров для амплификации ДНК и зондов для гибридизации). Традиционные колоночные и планшетные синтезаторы ДНК малоприспособлены для получения таких массивов олигонуклеотидов. Для такого рода задач необходимы микрочиповые синтезаторы ДНК. Экспорт высокопроизводительных зарубежных синтезаторов в нашу страну был запрещен с момента их создания. Учитывая исключительную важность таких приборов для современной фундаментальной и прикладной науки, с целью создания макета микрочипового синтезатора олигонуклеотидов Сибирским отделением Российской академии наук был организован консорциум, в который вошли следующие институты: Институт химической биологии и фундаментальной медицины, Институт физики полупроводников, Институт органической химии, Институт автоматизации и электрометрии. В настоящее время эта работа успешно завершена [124, 125]. Разработанный макет позволяет синтезировать до 12 тыс. олигонуклеотидов в одном эксперименте. Полученный прибор открывает возможность проведения исследований с использованием таргетного секвенирования и развивать синтетическую биологию в нашей стране на мировом уровне.

Написание обзора не потребовало специального финансирования.

Настоящая статья не содержит каких-либо исследований с участием людей или животных в качестве объектов исследований.

Авторы заявляют об отсутствии конфликта интересов.

СПИСОК ЛИТЕРАТУРЫ

1. Lesnik E.A., Freier S.M. (1995) Relative thermodynamic stability of DNA, RNA, and DNA:RNA hybrid duplexes: relationship with base composition and structure. *Biochemistry*. **34**, 10807–10815.
2. Okou D.T., Steinberg K.M., Middle C., Cutler D.J., Albert T.J., Zwick M.E. (2007) Microarray-based genomic selection for high-throughput resequencing. *Nat. Methods*. **4**(11), 907–909.
3. Mamanova L., Coffey A.J., Scott C.E., Kozarewa I., Turner E.H., Kumar A., Howard E., Shendure J., Turner D.J. (2010) Target-enrichment strategies for

- next-generation sequencing. *Nat. Methods*. **7**, 111–118.
4. Kozarewa I., Armisen J., Gardner A.F., Slatko B.E., Hendrickson C.L. (2015) Overview of target enrichment strategies. *Curr. Protoc. Mol. Biol.* **112**, 7.21.1–7.21.23.
 5. Hodges E., Xuan Z., Balija V., Kramer M., Molla M.N., Smith S.W., Middle C.M., Rodesch M.J., Albert T.J., Hannon G.J., McCombie W.R. (2008) Genome-wide *in situ* exon capture for selective resequencing. *Nat. Genet.* **39**(12), 1522–1527.
 6. Albert T.J., Molla M.N., Muzny D.M., Nazareth L., Wheeler D., Song X., Richmond T.A., Middle C.M., Rodesch M.J., J Packard C.J., Weinstock G.M., Gibbs R.A. (2007) Direct selection of human genomic loci by microarray hybridization. *Nat. Methods*. **4**(11), 903–905.
 7. Choi M., Scholl U.I., Ji W., Liu T., Tikhonova I.R., Zumbo P., Nayir A., Lu A.B., Ozen S., Sanjad S., Nelson-Williams C., Farhi A., Mane S., Lifton R.P. (2009) Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. *Proc. Natl. Acad. Sci. USA*. **106**(45), 19096–19101.
 8. Ng S.B., Turner E.H., Robertson P.D., Flygare S.D., Bigham A.W., Lee C., Shaffer T., Wong M., Bhattacharjee A., Eichler E.E., Bamshad M., Nickerson D.A., Shendure J. (2009) Targeted capture and massively parallel sequencing of 12 human exomes. *Nature*. **461**, 272–276.
 9. Salmon A., Udall J.A., Jeddell J.A., Wendel J. (2012) Targeted capture of homoeologous coding and noncoding sequence in polyploid cotton. *G3* (Bethesda). **2**(8), 921–930.
 10. Lee H., O'Connor B.D., Merriman B., Funari V.A., Homer N., Chen Z., Cohn D.H., Nelson S.F. (2009) Improving the efficiency of genomic loci capture using oligonucleotide arrays for high throughput resequencing. *BMC Genomics*. **10**, 646.
 11. Summerer D., Wu H., Haase B., Cheng Y., Schracke N., Stahler C.F., Chee M.S., Stahler P.F., Beier M. (2009) Microarray-based multicycle-enrichment of genomic subsets for targeted next-generation sequencing. *Genome Res.* **19**(9), 1616–1621.
 12. Bau S., Schracke N., Kränzle M., Wu H., Stähler P.F., Hoheisel J.D., Beier M., Summerer D. (2009) Targeted next-generation sequencing by specific capture of multiple genomic loci using low-volume microfluidic DNA arrays. *Anal. Bioanal. Chem.* **393**, 171–175.
 13. Summerer D., Schracke N., Wu H., Cheng Y., Bau S., Stähler C.F., Stähler P.F., Beier M. (2010) Targeted high throughput sequencing of a cancer-related exome subset by specific sequence capture with a fully automated microarray platform. *Genomics*. **95**, 241–246.
 14. Gnirke A., Melnikov A., Maguire J., Rogov P., LeProust E.M., Brockman W., Fennell T., Giannoukos G., Fisher S., Russ C., Gabriel S., Jaffe D.B., Lander E.S., Nusbaum C. (2009) Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. *Nat. Biotechnol.* **27**, 182–189.
 15. Chen R., Im H., Snyder M. (2015) Whole-exome enrichment with the Agilent SureSelect Human All Exon platform. *Cold Spring Harb. Protoc.* **7**, 626–633.
 16. Chen R., Im H., Snyder M. (2015) Whole-exome enrichment with the Illumina TruSeq exome enrichment platform. *Cold Spring Harb. Protoc.* **7**, 642–648.
 17. Chen R., Im H., Snyder M. (2015) Whole-exome enrichment with the roche NimbleGen SeqCap EZ exome library SR platform. *Cold Spring Harb. Protoc.* **7**, 634–641.
 18. Caruccio N. (2011) Preparation of next-generation sequencing libraries using Nextera™ technology: simultaneous DNA fragmentation and adaptor tagging by *in vitro* transposition. *Methods Mol. Biol.* **733**, 241–255.
 19. Marine R., Polson S.W., Ravel J., Hatfull G., Russell D., Sullivan M., Syed F., Dumas M., Wommack K.E. (2011) Evaluation of a transposase protocol for rapid generation of shotgun high-throughput sequencing libraries from nanogram quantities of DNA. *Appl. Environ. Microbiol.* **77**(22), 8071–8079.
 20. Shearer A.E., DeLuca A.P., Hildebrand M.S., Taylor K.R., Gurrola J., Scherer S., Scheetz T.E., Smith R.J.H. (2010) Comprehensive genetic testing for hereditary hearing loss using massively parallel sequencing. *Proc. Natl. Acad. Sci. USA*. **107**, 21104–21109.
 21. Martignetti J.A., Tian L., Li D., Ramirez M.C.M., Camacho-Vanegas O., Camacho S.C., Guo Y., Zand D.J., Bernstein A.M., Masur S.K., Kim C.E., Otieno F.G., Hou C., Abdel-Magid N., Tweddle B., Metry D., Fournet J.C., Papp E., McPherson E.V., Zabel C., Vaksmann G., Morisot C., Keating B., Sleiman P.M., Cleveland J.A., Everman D.B., Zackai E., Hakonarson H. (2013) Mutations in *PDGFRB* cause autosomal-dominant infantile myofibromatosis. *Am. J. Hum. Genet.* **92**, 1001–1007.
 22. Nectoux J., de Cid R., Baulande S., Leturcq F., Urtizberea J.A., Penisson-Besnier I., Nadaj-Pakleza A., Roudaut C., Criqui A., Orhant L., Peyroulan D., Yaou R.B., Nelson I., Cobo A.M., Arné-Bes M.C., Uro-Coste E., Nitschke P., Claustres M., Bonne G., Lévy N., Chelly J., Richard I., Cossée M. (2015) Detection of *TRIM32* deletions in LGMD patients analyzed by a combined strategy of CGH array and massively parallel sequencing. *Eur. J. Hum. Genet.* **23**, 929–934.
 23. Rousseau-Nepton I., Okubo M., Grabs R., Mitchell J., Polychronakos C., Rodd C. (2015) A founder *AGL* mutation causing glycogen storage disease type IIIa in Inuit identified through whole-exome sequencing: a case series. *Can. Med. Assoc. J.* **187**, E68–E73.
 24. Poultney C.S., Goldberg A.P., Drapeau E., Kou Y., Harony-Nicolas H., Kajiwar Y., De Rubeis S., Durand S., Stevens C., Rehnström K., Palotie A., Daly M.J., Ma'ayan A., Fromer M., Buxbaum J.D. (2013) Identification of small exonic *CNV* from whole-exome sequence data and application to autism spectrum disorder. *Am. J. Hum. Genet.* **93**, 607–619.
 25. Guipponi M., Santoni F.A., Setola V., Gehrig C., Rotharmel M., Cuenca M., Guillin O., Dikeos D., Georgantopoulos G., Papadimitriou G., Curtis L., Méary A., Schürhoff F., Jamain S., Avramopoulos D.,

- Leboyer M., Rujescu D., Pulver A., Campion D., Siderovski D.P., Antonarakis S.E. (2014) Exome sequencing in 53 sporadic cases of schizophrenia identifies 18 putative candidate genes. *PLoS One*. **9**, e112745.
26. Griesi-Oliveira K., Acab A., Gupta A.R., Sunaga D.Y., Chailangkarn T., Nicol X., Nunez Y., Walker M.F., Murdoch J.D., Sanders S.J., Fernandez T.V., Ji W., Lifton R.P., Vadasz E., Dietrich A., Pradhan D., Song H., Ming G.L., Gu X., Haddad G., Marchetto M.C.N., Spitzer N., Passos-Bueno M.R., State M.W., Muotri A.R. (2015) Modeling non-syndromic autism and the impact of *TRPC6* disruption in human neurons. *Mol. Psychiatry*. **20**, 1350–1365.
 27. Pérez-Serra A., Toro R., Campuzano O., Sarquella-Brugada G., Berne P., Iglesias A., Mangas A., Brugada J., Brugada R. (2015) A novel mutation in lamin A/c causing familial dilated cardiomyopathy associated with sudden cardiac death. *J. Card. Fail.* **21**, 217–225.
 28. Calvo S.E., Compton A.G., Hershman S.G., Lim S.C., Lieber D.S., Tucker E.J., Laskowski A., Garone C., Liu S., Jaffe D.B., Christodoulou J., Fletcher J.M., Bruno D.L., Goldblatt J., Dimauro S., Thorburn D.R., Mootha V.K. (2012) Molecular diagnosis of infantile mitochondrial disease with targeted next-generation sequencing. *Sci. Transl. Med.* **4**, 118ra10.
 29. Gai X., Ghezzi D., Johnson M.A., Biagosch C.A., Shamseldin H.E., Haack T.B., Reyes A., Tsukikawa M., Sheldon C.A., Srinivasan S., Gorza M., Kremer L.S., Wieland T., Strom T.M., Polyak E., Place E., Consugar M., Ostrovsky J., Vidoni S., Robinson A.J., Wong L.-J., Sondheimer N., Salih M.A., Al-Jishi E., Raab C.P., Bean C., Furlan F., Parini R., Lamperti C., Mayr J.A., Konstantopoulou V., Huemer M., Pierce E.A., Meitinger T., Freisinger P., Sperl W., Prokisch H., Alkuraya F.S., Falk M.J., Zeviani M. (2013) Mutations in *FBXL4*, encoding a mitochondrial protein, cause early-onset mitochondrial encephalomyopathy. *Am. J. Hum. Genet.* **93**, 482–495.
 30. Sikkema-Raddatz B., Johansson L.F., de Boer E.N., Almomani R., Boven L.G., van den Berg M.P., van Spaendonck-Zwarts K.Y., van Tintelen J.P., Sijmons R.H., Jongbloed J.D.H., Sinke R.J. (2013) Targeted next-generation sequencing can replace Sanger sequencing in clinical diagnostics. *Hum. Mutat.* **34**, 1035–1042.
 31. Drilon A., Wang L., Arcila M.E., Balasubramanian S., Greenbowe J.R., Ross J.S., Stephens P., Lipson D., Miller V.A., Kris M.G., Ladanyi M., Rizvi N.A. (2015) Broad, hybrid capture-based next-generation sequencing identifies actionable genomic alterations in “driver-negative” lung adenocarcinomas. *Clin. Cancer Res.* **21**, 3631–3639.
 32. Xie J., Lu X., Wu X., Lin X., Zhang C., Huang X., Chang Z., Wang X., Wen C., Tang X., Shi M., Zhan Q., Chen H., Deng X., Peng C., Li H., Fang Y., Shao Y., Shen B. (2016) Capture-based next-generation sequencing reveals multiple actionable mutations in cancer patients failed in traditional testing. *Mol. Genet. Genomic Med.* **4**, 262–272.
 33. Rozenblum A.B., Ilouze M., Dudnik E., Dvir A., Soussan-Gutman L., Geva S., Peled N. (2017) Clinical impact of hybrid capture-based next-generation sequencing on changes in treatment decisions in lung cancer. *J. Thorac. Oncol.* **12**, 258–268.
 34. Xu M.-D., Liu S.-L., Feng Y.-Z., Liu, Q., Shen M., Zhi Q., Liu Z., Gu D.M., Yu J., Shou L.-M., Gong F.-R., Zhu Q., Duan W., Chen K., Zhang J., Wu M.-Y., Tao M., Li W. (2017) Genomic characteristics of pancreatic squamous cell carcinoma, an investigation by using high throughput sequencing after in-solution hybrid capture. *Oncotarget*. **8**, 14620–14635.
 35. Clark T.A., Chung J.H., Kennedy M., Hughes J.D., Chennagiri N., Lieber D.S., Fendler B., Young L., Zhao M., Coyne M., Breese V., Young G., Donahue A., Pavlick D., Tsiros A., Brennan T., Zhong S., Mughal T., Bailey M., He J., Roels S., Frampton G.M., Sporerke J.M., Gendreau S., Lackner M., Schleifman E., Peters E., Ross J.S., Ali S.M., Miller V.A., Gregg J.P., Stephens P.J., Welsh A., Otto G.A., Lipson D. (2018) Analytical validation of a hybrid capture-based next-generation sequencing clinical assay for genomic profiling of cell-free circulating tumor DNA. *J. Mol. Diagn.* **20**, 686–702.
 36. Schrock A.B., Pavlick D., Klempner S.J., Chung J.H., Forcier B., Welsh, A., Young L., Leyland-Jones B., Bordoni R., Carvajal R.D., Chao J., Kurzrock R., Sicklick J.K., Ross J.S., Stephens P.J., Devoe C., Braithe F., Ali S.A., Miller V.A. (2018) Hybrid capture-based genomic profiling of circulating tumor DNA from patients with advanced cancers of the gastrointestinal tract or anus. *Clin. Cancer Res.* **24**, 1881–1890.
 37. Briggs A.W., Good J.M., Green R.E., Krause J., Maricic T., Stenzel U., Lalueza-Fox C., Rudan P., Brajkovic D., Kucan Z., Gusic I., Schmitz R., Doronichev V.B., Golovanova L.V., de la Rasilla M., Fortea J., Rosas A., Pääbo S. (2009) Targeted retrieval and analysis of five Neandertal mtDNA genomes. *Science*. **325**, 318–321.
 38. Krause J., Fu Q., Good J.M., Viola B., Shunkov M.V., Derevianko A.P., Pääbo S. (2010) The complete mitochondrial DNA genome of an unknown hominin from southern Siberia. *Nature*. **464**, 894–897.
 39. Enk J., Rouillard J.-M., Poinar H. (2013) Quantitative PCR as a predictor of aligned ancient DNA read counts following targeted enrichment. *Biotechniques*. **55**, 300–309.
 40. Kihana M., Mizuno F., Sawafuji R., Wang L., Ueda S. (2013) Emulsion PCR-coupled target enrichment: an effective fishing method for high-throughput sequencing of poorly preserved ancient DNA. *Gene*. **528**, 347–351.
 41. Templeton J.E.L., Brotherton P.M., Llamas B., Soubrier J., Haak W., Cooper A., Austin J.J. (2013) DNA capture and next-generation sequencing can recover whole mitochondrial genomes from highly degraded samples for human identification. *Investig. Genet.* **4**, 26.
 42. Eduardoff M., Xavier C., Strobl C., Casas-Vargas A., Parson W. (2017) Optimized mtDNA control region primer extension capture analysis for forensically relevant samples and highly compromised mtDNA of different age and origin. *Genes*. **8**, E237.

43. Loreille O., Ratnayake S., Bazinet A.L., Stockwell T.B., Sommer D.D., Rohland N., Mallick S., Johnson P.L.F., Skoglund P., Onorato A.J., Bergman N.H., Reich D., Irwin J.A. (2018) Biological sexing of a 4000-year-old Egyptian mummy head to assess the potential of nuclear DNA recovery from the most damaged and limited forensic specimens. *Genes*. **9**, E135.
44. Fu Q., Meyer M., Gao X., Stenzel U., Burbano H. A., Kelso J., Pääbo S. (2013) DNA analysis of an early modern human from Tianyuan Cave, China. *Proc. Natl. Acad. Sci. USA*. **110**, 2223–2227.
45. Cruz Dávalos D.I., Nieves-Colón M.A., Sockell A., Poznik D.G., Schroeder H., Stone A.C., Bustamante C.D., Malaspinas A.-S., Ávila-Arcos M.C. (2017) In-solution Y-chromosome capture-enrichment on ancient DNA libraries. *BMC Genomics*. **19**, 608.
46. Carpenter M.L., Buenrostro J.D., Valdiosera C., Schroeder H., Allentoft M.E., Sikora M., Rasmussen M., Gravel S., Guillén S., Nekhrizov G., Leshtakov K., Dimitrova D., Theodossiev N., Pettener D., Luiselli D., Sandoval K., Moreno-Estrada A., Li Y., Wang J., Gilbert M.T.P., Willerslev E., Greenleaf W.J., Bustamante C.D. (2013) Pulling out the 1%: whole-genome capture for the targeted enrichment of ancient DNA sequencing libraries. *Am. J. Hum. Genet.* **93**, 852–864.
47. Enk J.M., Devault A.M., Kuch M., Murgha Y.E., Rouillard J.-M., Poinar H.N. (2014) Ancient whole genome enrichment using baits built from modern DNA. *Mol. Biol. Evol.* **31**, 1292–1294.
48. Lindo J., Achilli A., Perego U.A., Archer D., Valdiosera C., Petzelt B., Mitchell J., Worl R., Dixon E.J., Fifield T.E., Rasmussen M., Willerslev E., Cybulski J.S., Kemp B.M., DeGiorgio M., Malhi R.S. (2017) Ancient individuals from the North American Northwest Coast reveal 10000 years of regional genetic continuity. *Proc. Natl. Acad. Sci. USA*. **114**, 4093–4098.
49. Bi K., Vanderpool D., Singhal S., Linderth T., Moritz C., Good J.M. (2012) Transcriptome-based exon capture enables highly cost-effective comparative genomic data collection at moderate evolutionary scales. *BMC Genomics*. **13**, 403.
50. Cosart T., Beja-Pereira A., Chen S., Ng S.B., Shendure J., Luikart G. (2011) Exome-wide DNA capture and next generation sequencing in domestic and wild species. *BMC Genomics*. **12**, 347.
51. Perry G.H., Marioni J.C., Melsted P., Gilad Y. (2010) Genomic-scale capture and sequencing of endogenous DNA from feces. *Mol. Ecol.* **19**(24), 5332–5344.
52. Tewhey R., Nakano M., Wang X., Pabón-Peña C., Novak B., Giuffre A., Lin E., Happe S., Roberts D.N., LeProust E.M., Topol E.J., Harismendy O., Frazer K.A. (2009) Enrichment of sequencing targets from the human genome by solution hybridization. *Genome Biol.* **10**, R116.
53. Brownstein Z., Friedman L.M., Shahin H., Oron-Karni V., Kol N., Rayyan A.A., Parzefall T., Lev D., Shalev S., Frydman M., Davidov B., Shohat M., Rahile M., Lieberman S., Levy-Lahad E., Lee M.K., Shomron N., King M.-C., Walsh T., Kanaan M., Avraham K.B. (2011) Targeted genomic capture and massively parallel sequencing to identify genes for hereditary hearing loss in Middle Eastern families. *Genome Biol.* **12**, R89.
54. Miyazato P., Katsuya H., Fukuda A., Uchiyama Y., Matsuo M., Tokunaga M., Hino S., Nakao M., Satou Y. (2016) Application of targeted enrichment to next-generation sequencing of retroviruses integrated into the host human genome. *Sci. Rep.* **6**, 28324.
55. Hodges E., Rooks M., Xuan Z., Bhattacharjee A., Gordon D.B., Brizuela L., McCombie W.R., Hannon G.J. (2009) Hybrid selection of discrete genomic intervals on custom designed microarrays for massively parallel sequencing. *Nat. Protocols*. **4**, 960–974.
56. Nilsson M., Malmgren H., Samiotaki M., Kwiatkowski M., Chowdhary B.P., Landegren U. (1994) Padlock probes: circularizing oligonucleotides for localized DNA detection. *Science*. **265**, 2085–2088.
57. Akhras M.S., Unemo M., Thiagarajan S., Nyren P., Davis R.W., Fire A.Z., Pourmand N. (2007) Connector inversion probe technology: a powerful one-primer multiplex DNA amplification system for numerous scientific applications. *PLoS One*. **2**(9), e915.
58. Pogoda M., Hilke F.-J., Lohmann E., Sturm M., Lenz F., Matthes J., Muyas F., Ossowski S., Hoischen A., Faust U., Sepahi I., Casadei N., Poths S., Riess O., Schroeder C., Grundmann K. (2019) Single molecule molecular inversion probes for high throughput germline screenings in dystonia. *Front. Neurol.* **10**, 1332.
59. Bekers E.M., Eijkelenboom A., Rombout P., van Zwam P., Mol S., Ruijter E., Scheijen B., Flucke U. (2019) Identification of novel GNAS mutations in intramuscular myxoma using next-generation sequencing with single-molecule tagged molecular inversion probes. *Diagn. Pathol.* **14**, 15.
60. Lenting K., van den Heuvel C.N.A.M., van Ewijk A., ElMelik D., de Boer R., Tindall E., Wei G., Kusters B., Te Dorsthorst M., Ter Laan M., Huynen M.A., Leenders W.P. (2019) Mapping actionable pathways and mutations in brain tumours using targeted RNA next generation sequencing. *Acta Neuropathol. Commun.* **7**, 185.
61. van den Heuvel C.N.A.M., van Ewijk A., Zeelen C., de Bitter T., Huynen M., Mulders P., Oosterwijk E., Leenders W.P.J. (2019) Molecular profiling of druggable targets in clear cell renal cell carcinoma through targeted RNA sequencing. *Front. Oncol.* **9**, 117.
62. Stefan C.P., Hall A.T., Minogue T.D. (2018) Detection of 16S rRNA and KPC genes from complex matrix utilizing a molecular inversion probe assay for next-generation sequencing. *Sci. Rep.* **8**(1), 2028.
63. Neveling K., Mensenkamp A.R., Derks R., Kwint M., Ouchene H., Steehouwer M., van Lier B., Bosgoed E., Rikken A., Tychon M., Zafeiropoulou D., Castelein S., Hehir-Kwa J., Thung D.T., Hofste T., Lelieveld S.H., Bertens S.M.M., Adan I.B.J.F., Eijkelenboom A., Tops B.B., Yntema H., Stokowy T., Knappskog P.M., Hoberg-Vetti H., Steen V.M., Boyle E., Martin B., Ligtenberg M.J.L., Shendure J., Nelen M.R., Hoischen A. (2017) *BRCA* testing by single-molecule molecular inversion probes. *Clin. Chem.* **63**(2), 503–512.

64. Lange I.M., Koudijs M.J., van 't Slot R., Sonsma A.C.M., Mulder F., Carbo E.C., van Kempen M.J.A., Nijman I.J., Ernst R.F., Savelberg S.M.C., Knoers N.V.A.M., Brilstra E.H., Koelman B.P.C. (2018) Assessment of parental mosaicism in *SCN1A*-related epilepsy by single-molecule molecular inversion probes and next-generation sequencing. *J. Med. Genet.* **56**(2), 75–80.
65. Mirzaa G., Timms A.E., Conti V., Boyle E.A., Girisha K.M., Martin B., Kircher M., Olds C., Juusola J., Collins S., Park K., Carter M., Glass I., Krogeloh-Mann I., Chitayat D., Parikh A.S., Bradshaw R., Torti E., Braddock S., Burke L., Ghedia S., Stephan M., Stewart F., Prasad C., Napier M., Saitta S., Straussberg R., Gabbett M., O'Connor B.C., Keegan C.E., Yin L.J., Lai A.H.M., Martin N., McKinnon M., Addor M.-C., Boccuto L., Schwartz C.E., Lanoel A., Conway R.L., Devriendt K., Tatton-Brown K., Pierpont M.E., Painter M., Worgan L., Reggin J., Hennekam R., Tsuchiya K., Pritchard C.C., Aracena M., Gripp K.W., Cordisco M., Esch H.V., Garavelli L., Curry C., Goriely A., Kayserilli H., Shendure J., Graham J.Jr., Guerrini R., Dobyns W.B. (2016) *PIK3CA*-associated developmental disorders exhibit distinct classes of mutations with variable expression and tissue distribution. *JCI Insight.* **1**(9), e87623.
66. Wada Y., Maekawa M., Ohnishi T., Balan S., Matsuo S., Iwamoto K., Iwayama Y., Ohba H., Watanabe A., Hisano Y., Nozaki Y., Toyota T., Shimogori T., Itokawa M., Kobayashi T., Yoshikawa T. (2020) Peroxisome proliferator-activated receptor α as a novel therapeutic target for schizophrenia. *EBioMedicine.* **62**, 103130.
67. Sleczkowska M., Almomani R., Marchi M., de Greef B.T.A., Sopacua M., Hoeijmakers J.G.J., Lindsey P., Salvi E., Bönhof G.J., Ziegler D., Malik R.A., Waxman S.G., Lauria G., Faber C.G., Smeets H.J.M., Gerrits M.M. (2022) Peripheral ion channel gene screening in painful- and painless-diabetic neuropathy. *Int. J. Mol. Sci.* **23**, 7190.
68. Almomani R., Marchi M., Sopacua M., Lindsey P., Salvi E., de Koning B., Santoro S., Magri S., Smeets H.J.M., Martinelli B.F., Malik R.R., Ziegler D., Hoeijmakers J.G.J., Bönhof G., Dib-Hajj S., Waxman S.G., Merkies I.S.J., Lauria G., Faber C.G., Gerrits M.M. (2020) Evaluation of molecular inversion probe versus TruSeq® custom methods for targeted next-generation sequencing. *PLoS One.* **15**(9), e0238467.
69. Suzuki O., Dong O.M., Howard R.M., Wiltshire T. (2019) Characterizing the pharmacogenome using molecular inversion probes for targeted next-generation sequencing. *Pharmacogenomics.* **20**(14), 1005–1020.
70. Yoon J.-K., Ahn J., Kim H.S., Han S.M., Jang H., Lee M.G., Lee J.H., Bang D. (2015) microDuMIP: target-enrichment technique for microarray-based duplex molecular inversion probes. *Nucleic Acids Res.* **43**(5), e28.
71. Saiki R.K., Scharf S., Faloona F., Mullis K.B., Horn G.T., Erlich H.A., Arnheim N. (1985) Enzymatic amplification of β -globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science.* **230**, 1350–1354.
72. Halbritter J., Diaz K., Chaki M., Porath J.D., Tarrier B., Fu C., Innis J.L., Allen S.J., Lyons R.H., Stefanidis C.J., Omran H., Soliman N.A., Otto E.A. (2012) High-throughput mutation analysis in patients with a nephronophthisis-associated ciliopathy applying multiplexed barcoded array-based PCR amplification and next-generation sequencing. *J. Med. Genet.* **49**(12), 756–767.
73. Tewhey R., Warner J.B., Nakano M., Libby B., Medkova M., David P.H., Kotsopoulos S.K., Samuels M.L., Hutchison J.B., Larson J.W., Topol E.J., Weiner M.P., Harismendy O., Olson J., Link D.R., Frazer K.A. (2009) Microdroplet-based PCR enrichment for large-scale targeted sequencing. *Nat. Biotechnol.* **27**, 1025–1031.
74. Vollbrecht C., Werner R., Walter R.F.H., Christoph D.C., Heukamp L.C., Peifer M., Hirsch B., Burbat L., Mairinger T., Schmid K.W., Wohlschlaeger J., Mairinger F.D. (2015) Mutational analysis of pulmonary tumours with neuroendocrine features using targeted massive parallel sequencing: a comparison of a neglected tumour group. *Br. J. Cancer.* **113**, 1704–171.
75. Lee A., Lee S.-H., Jung C.K., Park G., Lee K.Y., Choi H.J., Min K.O., Kim T.J., Lee E.J., Lee Y.S. (2018) Use of the Ion AmpliSeq Cancer Hotspot Panel in clinical molecular pathology laboratories for analysis of solid tumours: with emphasis on validation with relevant single molecular pathology tests and the OncoPrint Focus Assay. *Pathol. Res. Pract.* **214**(5), 713–719.
76. Mustafa A.E., Faquih T., Baz B., Kattan R., Al-Issa A., Tahir A.I., Imtiaz F., Ramzan K., Al-Sayed M., Alowain M., Al-Hassan Z., Al-Zaidan H., Abouelhoda M., Al-Mubarak B.R., Al Tassan N.A. (2018) Validation of ion Torrent™ inherited disease panel with the PGM™ sequencing platform for rapid and comprehensive mutation detection. *Genes (Basel).* **9**(5), 267.
77. Cao C., Lou I., Huang C., Lee M.-Y. (2016) Metagenomic sequencing of activated sludge filamentous bacteria community using the Ion Torrent platform. *Desalin. Water Treat.* **57**(5), 2175–2183.
78. Jiang F., Doudna J.A. (2017) CRISPR-Cas9 structures and mechanisms. *Annu. Rev. Biophys.* **46**, 505–529.
79. Schultzhause Z., Wang Z., Stenger D. (2021) CRISPR-based enrichment strategies for targeted sequencing. *Biotechnol. Adv.* **46**, 107672.
80. Gu W., Crawford E.D., O'Donovan B.D., Wilson M.R., Chow E.D., Retallack H., DeRisi J.L. (2016) Depletion of Abundant Sequences by Hybridization (DASH): using Cas9 to remove unwanted high-abundance species in sequencing libraries and molecular counting applications. *Genome Biol.* **17**(1), 41.
81. Lee S.H., Yu J., Hwang G.H., Kim S., Kim H.S., Ye S., Kim K., Park J., Park D.Y., Cho Y.K., Kim J.S., Bae S. (2017) CUT-PCR: CRISPR-mediated, ultrasensitive detection of target DNA using PCR. *Oncogene.* **36**(49), 6823–6829.
82. Stevens R.C., Steele J.L., Glover W.R., Sanchez-Garcia J.F., Simpson S.D., O'Rourke D., Ramsdell J.S., MacManes M.D., Thomas W.K., Shuber A.P. (2019)

- A novel CRISPR/Cas9 associated technology for sequence-specific nucleic acid enrichment. *PLoS One*. **14**(4), e0215441.
83. Bennett-Baker P.E., Mueller J.L. (2017) CRISPR-mediated isolation of specific megabase segments of genomic DNA. *Nucleic Acids Res.* **45**(19), e165.
 84. Gabrieli T., Sharim H., Fridman D., Arbib N., Michaeli Y., Ebenstein Y. (2018) Selective nanopore sequencing of human *BRCA1* by Cas9-assisted targeting of chromosome segments (CATCH). *Nucleic Acids Res.* **46**(14), e87.
 85. Nachmanson D., Lian S., Schmidt E.K., Hipp M.J., Baker K.T., Zhang Y., Tretiakova M., Loubet-Seneor K., Kohn B.F., Salk J.J., Kennedy S.R., Risques R.A. (2018) Targeted genome fragmentation with CRISPR/Cas9 enables fast and efficient enrichment of small genomic regions and ultra-accurate sequencing with low DNA input (CRISPR-DS). *Genome Res.* **28**(10), 1589–1599.
 86. Lee J., Lim H., Jang H., Hwang B., Lee J.H., Cho J., Lee J.H., Bang D. (2019) CRISPR-Cap: multiplexed double-stranded DNA enrichment based on the CRISPR system. *Nucleic Acids Res.* **47**(1), e1.
 87. Peter B.J., Ach R.A. (2014) US Patent No. US 20140356867 A1. patents.google.com/patent/US20140356867A1/en?q=US+20140356867+A1.
 88. Carpenter M.L., Bustamante C.D. (2018) US Patent No. US 20180298421 A1. patents.google.com/patent/US20180298421A1/en?q=US+20180298421.
 89. Bang D., Ji Won Lee J.W., Lim H.S. (2016) US Patent No. US 20160244829 A1. patents.google.com/patent/US20160244829A1/en?q=US20160244829A1.
 90. Xu X., Luo T., Gao J., Lin N., Li W., Xia X., Wang J. (2020) CRISPR-assisted DNA detection: a novel dCas9-based DNA detection technique. *CRISPR J.* **3**(6), 487–502.
 91. Aalipour A., Dudley J.C., Park S.M., Murty S., Chabon J.J., Boyle E.A., Diehn M., Gambhir S.S. (2018) Deactivated CRISPR associated protein 9 for minor-allele enrichment in cell-free DNA. *Clin. Chem.* **64**(2), 307–316.
 92. Quan J., Langelier C., Kuchta A., Batson J., Teyssier N., Lyden A., Caldera S., McGeever A., Dimitrov B., King R., Wilhelm J., Murphy M., Ares L.P., Trivisano K.A., Sit R., Amato R., Mumbengegwi D.R., Smith J.L., Bennett A., Gosling R., Mourani P.M., Calfee C.S., Neff N.F., Chow E.D., Kim P.S., Greenhouse B., DeRisi J.L., Crawford E.D. (2019) FLASH: a next-generation CRISPR diagnostic for multiplexed detection of antimicrobial resistance sequences. *Nucleic Acids Res.* **47**(14), e83.
 93. Hafford-Tear N.J., Tsai Y.-C., Sadan A.N., Sanchez-Pintado B., Zarouchlioti C., Maher G.J., Liskova P., Tuft S.J., Hardcastle A.J., Clark T.A., Davidson A.E. (2019) CRISPR/Cas9-targeted enrichment and long-read sequencing of the Fuchs endothelial corneal dystrophy-associated *TCF4* triplet repeat. *Genet. Med.* **21**(9), 2092–2102.
 94. Stangl C., de Blank S., Renkens I., Westera L., Verbeek T., Valle-Inclan J.E., González R.C., Henssen A.G., van Roosmalen M.J., Stam R.W., Voest E.E., Kloosterman W.P., van Haaften G., Monroe G.R. (2020) Partner independent fusion gene detection by multiplexed CRISPR-Cas9 enrichment and long read nanopore sequencing. *Nat. Commun.* **11**(1), 2861.
 95. Watson C.M., Crinnion L.A., Hewitt S., Bates J., Robinson R., Carr I.M., Sheridan E., Adlard J., Bonthron D.T. (2019) Cas9-based enrichment and single-molecule sequencing for precise characterization of genomic duplications. *Lab. Inv.* **100**(1), 135–146.
 96. Gilpatrick T., Lee I., Graham J.E., Raimondeau E., Bowen R., Heron A., Sedlazeck F.J., Timp W. (2020) Targeted Nanopore Sequencing with Cas9 for studies of methylation, structural variants, and mutations. *Nat. Biotechnol.* **38**(4), 433–438.
 97. Kuhlensäuer G., Hullmann J., Appenzeller S. (2011) Novel genomic techniques open new avenues in the analysis of monogenic disorders. *Hum. Mutat.* **32**(2), 144–151.
 98. Kiezun A., Garimella K., Do R., Stitzel N.O., Neale B.M., McLaren P.J., Gupta N., Sklar P., Sullivan P.F., Morn J.L., Hultman C.M., Lichtenstein P., Magnusson P., Lehner T., Shugart Y.Y., Price A.L., de Bakker P.I., Purcell S.M., Sunyaev S.R. (2012) Exome sequencing and the genetic basis of complex traits. *Nat. Genet.* **44**(6), 623–630.
 99. Vogelstein B., Papadopoulos N., Velculescu V.E., Zhou S., Diaz L.A.J., Kinzler K.W. (2013) Cancer genome landscapes. *Science*. **339**(6127), 1546–1558.
 100. Gaudin M., Desnues C. (2018) Hybrid capture-based next generation sequencing and its application to human infectious diseases. *Front. Microbiol.* **9**, 2924.
 101. Melnikov A., Galinsky K., Rogov P., Fennell T., Van Tyne D., Russ C., Daniels R., Barnes K.G., Bochicchio J., Ndiaye D., Sene P.D., Wirth D.F., Nusbaum C., Volkman S.K., Birren B.W., Gnirke A., Neafsey D.E. (2011) Hybrid selection for sequencing pathogen genomes from clinical samples. *Genome Biol.* **12**, R73.
 102. Smith M., Campino S., Gu Y., Clark T.G., Otto T.D., Maslen G., Manske M., Imwong M., Dondorp A.M., Kwiatkowski D.P., Quail M.A., Swerdlow H. (2012). An in-solution hybridisation method for the isolation of pathogen DNA from human DNA-rich clinical samples for analysis by NGS. *Open Genom. J.* **5**, 18–29.
 103. Bright A.T., Tewhey R., Abeles S., Chuquiyauri R., Llanos-Cuentas A., Ferreira M.U., Schork N.J., Vinetz J.M., Winzler E.A. (2012) Whole genome sequencing analysis of *Plasmodium vivax* using whole genome capture. *BMC Genomics*. **13**, 262.
 104. Amorim-Vaz S., Tran V.D.T., Pradervand S., Pagni M., Coste A.T., Sanglard D. (2015) RNA enrichment method for quantitative transcriptional analysis of pathogens *in vivo* applied to the fungus *Candida albicans*. *mBio*. **6**(5), e00942-15.
 105. Brown A.C., Bryant J.M., Einer-Jensen K., Holdstock J., Houniet D.T., Chan J.Z.M., Depledge D.P., Nikolayevskiy V., Broda A., Stone M.J., Christiansen M.T., Williams R., McAndrew M.B., Tutill H., Brown J., Melzer M., Rosmarin C., McHugh T.D., Shorten R.J., Drobniowski F., Speight G., Breuer J. (2015) Rapid whole-genome sequencing of *Mycobac-*

- terium tuberculosis* isolates directly from clinical samples. *J. Clin. Microbiol.* **53**(7), 2230–2237.
106. Christiansen M.T., Brown A.C., Kundu S., Tutill H.J., Williams R., Brown J.R., Holdstock J., Holland M.J., Stevenson S., Dave J., Tong C.Y., Einer-Jensen K., Depledge D.P., Breuer J. (2014) Whole-genome enrichment and sequencing of *Chlamydia trachomatis* directly from clinical samples. *BMC Infect. Dis.* **14**, 591.
 107. Clark S.A., Doyle R., Lucidarme J., Borrow R., Breuer J. (2018) Targeted DNA enrichment and whole genome sequencing of *Neisseria meningitidis* directly from clinical specimens. *Int. J. Med. Microbiol.* **308**, 256–262.
 108. Houldcroft C.J., Beale M.A., Breuer J. (2017) Clinical and biological insights from viral genome sequencing. *Nat. Rev. Microbiol.* **15**, 183–192.
 109. Depledge D.P., Palser A.L., Watson S.J., Lai I.Y.-C., Gray E.R., Grant P., Kanda R.K., Leproust E., Kellam P., Breuer J. (2011) Specific capture and whole-genome sequencing of viruses from clinical samples. *PLoS One*. **6**, e27805.
 110. Kwok H., Wu C.W., Palser A.L., Kellam P., Sham P.C., Kwong D.L.W., Chiang A.K.S. (2014) Genomic diversity of Epstein–Barr virus genomes isolated from primary nasopharyngeal carcinoma biopsy samples. *J. Virol.* **88**, 10662–10672.
 111. Metsky H.C., Matranga C.B., Wohl S., Schaffner S.F., Freije C.A., Winnicki S.M., West K., Qu J., Baniecki M.L., Gladden-Young A., Lin A.E., Tomkins-Tinch C.H., Ye S.H., Park D.J., Luo C.Y., Barnes K.G., Shah R.R., Chak B., Barbosa-Lima G., Delatorre E., Vieira Y.R., Paul L.M., Tan A.L., Barcellona C.M., Porcelli M.C., Vasquez C., Cannons A.C., Cone M.R., Hogan K.N., Kopp E.W., Anzinger J.J., Garcia K.F., Parham L.A., Ramírez R.M.G., Montoya M.C.M., Rojas D.P., Brown C.M., Hennigan S., Sabina B., Scotland S., Gangavarapu K., Grubaugh N.D., Oliveira G., Robles-Sikisaka R., Rambaut A., Gehrke L., Smole S., Halloran M.E., Villar L., Mattar S., Lorenzana I., Cerbino-Neto J., Valim C., Degraeve W., Bozza P.T., Gnirke A., Andersen K.G., Isern S., Michael S.F., Bozza F.A., Souza T.M.L., Bosch I., Yozwiak N.L., MacInnis B.L., Sabeti P.C. (2017) Zika virus evolution and spread in the Americas. *Nature*. **546**, 411–415.
 112. Brown J.R., Roy S., Ruis C., Yara Romero E., Shah D., Williams R., Breuer J. (2016) Norovirus whole-genome sequencing by sureselect target enrichment: a robust and sensitive method. *J. Clin. Microbiol.* **54**, 2530–2537.
 113. Matranga C.B., Andersen K.G., Winnicki S., Busby M., Gladden A.D., Tewhey, R., Stremlau M., Berlin A., Gire S.K., England E., Moses L.M., Mikkelsen T.S., Oda I., Ehiane P.E., Folarin O., Goba A., Kahn S.H., Grant D.S., Honko A., Hensley L., Happi C., Garry R.F., Malboeuf C.M., Birren B.W., Gnirke A., Levin J.Z., Sabeti P.C. (2014) Enhanced methods for unbiased deep sequencing of Lassa and Ebola RNA viruses from clinical and biological samples. *Genome Biol.* **15**, 519.
 114. Thomson E., Ip C.L.C., Badhan A., Christiansen M.T., Adamson W., Ansari M.A., Bibby D., Breuer J., Brown A., Bowden R., Bryant J., Bonsall D., Da Silva Filipe A., Hinds C., Hudson E., Klennerman P., Lythgow K., Mbisa J.L., McLauchlan J., Myers R., Piazza P., Roy S., Trebes A., Sreenu V.B., Witteveldt J., STOP-HCV Consortium, Barnes E., Simmonds P. (2016) Comparison of next generation sequencing technologies for the comprehensive assessment of full-length hepatitis C viral genomes. *J. Clin. Microbiol.* **54**, 2470–2484.
 115. Donaldson C.D., Clark D.A., Kidd I.M., Breuer J., Depledge D.D. (2013) Genome sequence of human herpesvirus 7 strain UCL-1. *Genome Announc.* **1**, e00830-13.
 116. Greninger A.L., Roychoudhury P., Xie H., Casto A., Cent A., Pepper G., Koelle D.M., Huang M.L., Wald A., Johnston C., Jerome K.R. (2018) Ultrasensitive capture of human herpes simplex virus genomes directly from clinical samples reveals extraordinarily limited evolution in cell culture. *mSphere*. **3**, e00283-18.
 117. Wylie T.N., Wylie K.M., Herter B.N., Storch G.A. (2015) Enhanced virome sequencing using targeted sequence capture. *Genome Res.* **25**, 1910–1920.
 118. Briese T., Kapoor A., Mishra N., Jain K., Kumar A., Jabado O., Lipkin W.I. (2015) Virome capture sequencing enables sensitive viral diagnosis and comprehensive virome analysis. *mBio*. **6**, e01491-15.
 119. Chalkias S., Gorham J.M., Mazaika E., Parfenov M., Dang X., DePalma S., McKean D., Seidman C.E., Seidman J.G., Koralnik I.J. (2018) ViroFind: a novel target-enrichment deep-sequencing platform reveals a complex JC virus population in the brain of PML patients. *PLoS One*. **13**, e0186945.
 120. Chilamakuri C.S.R., Lorenz S., Madoui M.-A., Vodák D., Sun J., Hovig E., Myklebost O., Meza-Zepeda L.A. (2014) Performance comparison of four exome capture systems for deep sequencing. *BMC Genomics*. **15**(1), 449.
 121. Clark M.J., Chen R., Lam H.Y.K., Karczewski K.J., Chen R., Euskirchen G., Butte A.J., Snyder M. (2011) Performance comparison of exome DNA sequencing technologies. *Nat. Biotechnol.* **29**(10), 908–914.
 122. Shigemizu D., Momozawa Y., Abe T., Morizono T., Borojevich K.A., Takata S., Ashikawa K., Kubo M., Tsunoda T. (2015) Performance comparison of four commercial human whole-exome capture platforms. *Sci. Rep.* **5**(1), 12742.
 123. Zhou J., Zhang M., Li X., Wang Z., Pan D., Shi Y. (2021) Performance comparison of four types of target enrichment baits for exome DNA sequencing. *Hereditas*. **158**, 10.
 124. Синяков А.Н., Бессмельцев В.П., Камаев Г.Н. (2018) Микрочиповый синтезатор ДНК. II Всероссийская научно-практическая конференция “Научное приборостроение — современное состояние и перспективы развития”. Казань 4–7 июня. Сб. материалов конференции. с. 280–282.
 125. Sinyakov A.N., Ryabinin V.A., Kostina E.V., Zaytsev D.E., Chukanov N.V., Kamaev G.N. (2021) Linkers for oligonucleotide microarray synthesis. *J. Saudi Chem. Soc.* **25**, 101382.

DNA Fragment Enrichment for High-Throughput Sequencing

A. N. Sinyakov^{1, *} and E. V. Kostina¹

¹*Institute of Chemical Biology and Fundamental Medicine, Siberian Branch, Russian Academy of Sciences,
Novosibirsk, 630090 Russia*

**e-mail: sinyakov@niboch.nsc.ru*

Application of oligonucleotides, mainly obtained using new generation DNA synthesizers (microarray DNA synthesizers), for the enrichment of targeted genomic fragments are described. Methods of molecular hybridization, polymerase chain reaction and CRISPR-based methods for targets enrichment are considered. Examples of the practical use of the developed methods for research and diagnostic purposes are given.

Keywords: high-throughput sequencing, oligonucleotides, microarray, amplification, diagnostics, exome, CRISPR-Cas

ГЕНОМИКА.
ТРАНСКРИПТОМИКА

УДК 577.218:578.842.1

ТРАНСКРИПЦИОННЫЙ АНАЛИЗ ГЕНА
ПРЕДПОЛАГАЕМОГО МИРИСТОИЛИРОВАННОГО
МЕМБРАННОГО БЕЛКА (ORF458R) ВИРУСА
РАДУЖНОЙ ОКРАСКИ БЕСПОЗВОНОЧНЫХ 6 (IIV6)¹

© 2023 г. С. А. Kuz^a, Е. Ozsahin^a, R. Nalcacioglu^a, Z. Demirbag^a, *

^aDepartment of Biology, Faculty of Science, Karadeniz Technical University, Trabzon, 61080 Turkey

*e-mail: zjhni@ktu.edu.tr

Поступила в редакцию 06.07.2022 г.

После доработки 06.07.2022 г.

Принята к публикации 04.09.2022 г.

Вирус радужной окраски беспозвоночных 6 (IIV6) относится к роду *Iridovirus* в семействе Iridoviridae. Полностью секвенированный двухцепочечный ДНК-геном, состоящий из 212482 п.н., кодирует 215 потенциальных открытых рамок считывания (ORF). Предполагают, что ORF 458R кодирует миристоилированный мембранный белок. Анализ экспрессии *ORF458R* в присутствии ингибиторов репликации ДНК и синтеза белка показал, что транскрипция этого гена начинается на поздней фазе вирусного инфекционного цикла. По результатам временного анализа транскрипция гена *ORF458R* инициируется в интервале 12–24 ч после инфекции и затем ее интенсивность начинает снижаться. Транскрипция *ORF458R* инициируется на 53 нуклеотида выше сайта начала трансляции и заканчивается через 40 нуклеотидов после стоп-кодона. По результатам двойного люциферазного теста сделан вывод, что последовательность между –61 и +18 нуклеотидами необходима для активности промотора. Интересно заметное снижение активности промотора в присутствии последовательности, находящейся между –299 и –143 нуклеотидами, из чего можно предположить наличие репрессорной активности в этой области. Нами показано, что *ORF458R* – транскрипционно активная и отдельно расположенная последовательность, в восходящей области которой локализованы регуляторные участки: промоторный и репрессорный. Надеемся, что полученная нами информация по структурной организации *ORF458R* позволит лучше понять молекулярные механизмы репликации IIV6.

Ключевые слова: вирус радужной окраски беспозвоночных 6, *ORF458R*, транскрипционный анализ, вирусы насекомых

DOI: 10.31857/S0026898423030059, EDN: CHGANM

Transcriptional Analysis of the Gene Encoding the Putative Myristoylated Membrane Protein (ORF458R) of Invertebrate Iridescent Virus 6 (IIV6)

C. A. Kuz¹, E. Ozsahin¹, R. Nalcacioglu¹, and Z. Demirbag¹, *

¹Department of Biology, Faculty of Science, Karadeniz Technical University, Trabzon, 61080 Turkey

*e-mail: zjhni@ktu.edu.tr

Invertebrate iridescent virus 6 (IIV6) is a member of the genus *Iridovirus* and belongs to the Iridoviridae family. The entirely sequenced dsDNA genome, composed of 212482 bp, encodes 215 putative open reading frames (ORFs). *ORF458R* encodes a putative myristoylated membrane protein. RT-PCR analysis of *ORF458R* expression in the presence of DNA replication and protein synthesis inhibitors showed that this

¹ Статья представлена авторами на английском языке.

gene is transcribed in the late phase of the virus infection. Time course analysis showed that transcription of *ORF458R* initiates between 12 and 24 h p.i. and starts to decrease after this point. Transcription of *ORF458R* initiated 53 nucleotides upstream of the translation start site and ended 40 nucleotides after the stop codon. Dual luciferase reporter gene assay showed that sequences between -61^{st} and $+18^{\text{th}}$ nucleotides are essential for promoter activity. Interestingly, a remarkable decrease in promoter activity, in the presence of sequences between -299^{th} and -143^{rd} nucleotides, suggested a repressor activity between these regions. Our results showed that *ORF458R* is transcriptionally active, and separately located sequences at its upstream region with promoter and repressor activities regulating its expression. This information on the transcriptional analysis of *ORF458R* will contribute to our understanding of the molecular mechanisms of IIV6 replication.

Keywords: invertebrate iridescent virus 6, *ORF458R*, transcriptional analysis, insect virus

ГЕНОМИКА.
ТРАНСКРИПТОМИКА

УДК 577.29

ДИФФЕРЕНЦИАЛЬНАЯ ЭКСПРЕССИЯ ЧУЖЕРОДНОГО ГЕНА
В МИТОХОНДРИЯХ АРАБИДОПСИСА В СИСТЕМЕ *in organello*

© 2023 г. В. И. Тарасенко^{а, 1}, Т. А. Тарасенко^{а, 1, *}, И. В. Горбенко^а,
Ю. М. Константинов^а, М. В. Кулинченко^{а, б}

^аСибирский институт физиологии и биохимии растений Сибирского отделения Российской академии наук,
Иркутск, 664033 Россия

^бКазанский институт биохимии и биофизики Казанского научного центра Российской академии наук,
Казань, 420111 Россия

*e-mail: bolotova_t.a@mail.ru

Поступила в редакцию 04.08.2022 г.

После доработки 16.09.2022 г.

Принята к публикации 30.09.2022 г.

Генетическая трансформация митохондрий высших эукариот *in vivo* — одна из нерешенных и практически важных задач. Для эффективной экспрессии чужеродного генетического материала в митохондриях необходим подбор регуляторных элементов, обеспечивающих высокий уровень транскрипции и стабильность образующихся транскриптов. С использованием явления природной компетентности митохондрий растений нами изучена эффективность регуляторных элементов митохондриальных генов, фланкирующих экзогенную ДНК. С этой целью в изолированные митохондрии арабидопсиса импортировали генетические конструкции, несущие ген *GFP* под контролем промоторных областей генов *RRN26* или *COX1* и одной из двух 3'-нетранслируемых областей (3'-НТО) митохондриальных генов, с последующей транскрипцией *in organello*. Показано, что уровень экспрессии *GFP* под контролем промоторов генов *RRN26* или *COX1* *in organello* коррелирует с уровнем транскрипции этих генов, наблюдаемым *in vivo*. При этом присутствие последовательности, кодирующей тРНК^{Trp}, в 3'-НТО приводит к более высокому уровню транскрипта *GFP*, чем сайта связывания белка MTSF1 из 3'-НТО гена *NAD4*. Полученные результаты открывают перспективы для создания системы эффективной трансформации митохондриального генома.

Ключевые слова: регуляция экспрессии генов, импорт ДНК, *GFP*, промоторная область, нетранслируемая область, RPO^{Tmp}, митохондрии, *Arabidopsis thaliana*, *in organello*

DOI: 10.31857/S002689842303014X, EDN: CHUECL

ВВЕДЕНИЕ

Генетическая трансформация митохондрий растений и животных *in vivo* — одна из нерешенных и практически важных задач [1]. Доставка ДНК в митохондрии затрудняет не только малый размер этих органелл и динамичная природа митохондриального генома растений [2], но и необходимость создания генетических конструкций, содержащих весь набор регуляторных элементов митохондриальных генов, способных обеспечить эффективную экспрессию целевого гена в органеллах *in vivo*.

Известно, что интенсивность транскрипции с той или иной промоторной области митохондриальных генов в системе *in vitro* часто не совпадает с интенсивностью, наблюдаемой *in vivo* [3], что затрудняет подбор оптимальных регуляторных элементов для генетической конструкции. Как

показано ранее, изолированные митохондрии растений, как и митохондрии животных, способны импортировать ДНК [4–6]. Использование этой способности митохондрий представляется перспективным подходом, позволяющим провести предварительное тестирование элементов генетических конструкций, предназначенных для трансформации митохондриального генома *in vivo*, в условиях *in organello*.

Транскрипция митохондриальных генов растений контролируется рядом промоторов, имеющих различную эффективность и распознаваемых разными РНК-полимеразами [7]. Ранее было показано, что транскрипция гена *COX1* в митохондриях арабидопсиса осуществляется преимущественно РНК-полимеразой двойной локализации RPO^{Tmp}, в то время как ген *RRN26* транскрибируется основной митохондриальной РНК-полимеразой RPOTm [8]. Методом run-on показано, что интенсивность транскрипции гена *RRN26*

¹ Эти авторы внесли равный вклад в выполнение исследования.

арабидопсиса существенно выше, чем гена *COX1* [8]. Учитывая эти особенности, мы выбрали для исследования уровня транскриптов гена *GFP in organello* промоторные области генов *RRN26* и *COX1*: дифференциальный характер экспрессии с разных промоторных областей позволил бы судить о корректном функционировании данных регуляторных элементов.

Известно, что в регуляции экспрессии генов в митохондриях растений ключевую роль играет посттранскрипционный уровень [3, 9]. Вторичные структуры и сайты связывания белков, присутствующие в 3'-нетранслируемых областях (3'-НТО) митохондриальных транскриптов, обеспечивают их защиту от деградации митохондриальными экзорибонуклеазами, в первую очередь полинуклеотидфосфорилазой, расщепляющей любые незащищенные транскрипты в митохондриальном матриксе [10, 11], тем самым внося решающий вклад в контроль уровня РНК в органеллах [12, 13]. В митохондриях растений тРНК-подобные элементы (а иногда и последовательности функциональных тРНК) нередко играют важную роль в стабилизации 3'-концов мРНК-транскриптов. Подобные структуры обнаружены в генах *scmC* и *nad6* арабидопсиса, а гены *atp6-1* и *atp6-2* содержат последовательности, кодирующие тРНК^{Сер}, непосредственно вслед за кодирующей последовательностью [14]. В созревании 3'-концов транскриптов этих генов участвует РНКазы Р, обеспечивающая процессинг тРНК [15]. В связи с этим при создании генетических конструкций логично использовать в качестве 3'-НТО последовательность гена тРНК.

Однако тРНК-подобные элементы и другие вторичные структуры [16, 17] присутствуют на 3'-концах лишь некоторых митохондриальных транскриптов растений. Предполагается, что наиболее распространенным механизмом стабилизации 3'-концов мРНК могут служить сайты, взаимодействующие со специфическими РНК-связывающими белками. Действительно, на 3'-конце зрелого транскрипта гена *NAD4* арабидопсиса обнаружен сайт связывания PPR-белка, названного MTSF1 [12]. Позднее идентифицировали еще два подобных белка, стабилизирующих 3'-концы зрелого транскрипта *NAD2* и одного из экзонов предшественника транскрипта *NAD1* [18, 19]. Мы использовали 30-нуклеотидную последовательность 3'-НТО гена *NAD4* арабидопсиса в качестве альтернативного варианта области, потенциально способной обеспечивать стабильность транскрипта целевого гена.

В результате проделанной работы установлено, что уровни транскриптов, синтезированные *in organello* с разных генетических конструкций, импортированных в митохондрии, зависят от промоторной области (*COX1* или *RRN26*), регули-

рующей транскрипцию *GFP*. Кроме того, показано, что последовательность тРНК^{Трп} на 3'-конце образующегося транскрипта стабилизирует его более эффективно, чем 3'-НТО гена *NAD4*. Представленные в работе данные открывают перспективы использования разработанной нами системы импорта и транскрипции рекомбинантной ДНК *in organello* для изучения регуляторных последовательностей митохондриального генома растений и эффективного подбора элементов генетических конструкций для трансформации митохондриального генома *in vivo*.

ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

Растительный материал и условия выращивания. Семена *Arabidopsis thaliana* (L.) Heynh. экотип Columbia (Col-0) подвергали стратификации в течение 3 сут при 4°C, а затем выращивали при 22°C в ростовой камере KBW720 ("Binder", Германия) в горшках, наполненных смесью компост/вермикулит в соотношении 2 : 1, при плотности потока фотосинтетических фотонов 150 мкмоль · м⁻² · с⁻¹ и 16-часовом фотопериоде.

Получение генетических конструкций, содержащих регуляторные последовательности митохондриальных генов, для проведения импорта в системе *in organello*. Генетические конструкции включали в себя: (а) 5'-область, содержащую промоторные элементы митохондриальных генов арабидопсиса *RRN26* (ATMG00020 согласно базе данных TAIR) или *COX1* (ATMG01360); (б) последовательность гена *GFP* медузы *Aequoria victoria*; (в) 3'-область, представляющую собой ген тРНК^{Трп} из линейной митохондриальной плазмиды 2.3 т.п.н. кукурузы (*Zea mays*) [20], либо принадлежащая гену *NAD4* (ATMG00580) *A. thaliana*. Все генетические конструкции получены на основе плазмиды pCK/GFP/PRmt [20]. Входящий в состав плазмиды промотор гена 18S рРНК из картофеля был замещен на промоторную область гена *RRN26* (*Prn26*) или *COX1* (*Pcox1*) арабидопсиса, в результате чего получили генетические конструкции pCK-Prn26-tRNA и pCK-Pcox1-tRNA. Для амплификации фрагментов Prn26-tRNA и Pcox1-tRNA, содержащих указанные регуляторные элементы, использованы специфические праймеры, последовательности которых представлены в табл. 1.

Для получения на основе pCK-Prn26-tRNA генетической конструкции, несущей альтернативную 3'-НТО, проведена амплификация фрагмента, содержащего промоторную область гена *RRN26* и ген *GFP* без последовательности тРНК на 3'-конце, с последующим клонированием в pJET1.2 согласно инструкции производителя ("Thermo Scientific", США). Для этого использовали прямой праймер к последовательности линейной плазмиды 2.3 т.п.н. в паре с обратным

Таблица 1. Праймеры, использованные при создании генетических конструкций

Синтезируемый фрагмент ДНК	Пары праймеров	Нуклеотидная последовательность (5' → 3')	Фрагмент, п.н. (матрица)
<i>Prn26*</i>	<i>rrn26-For</i> <i>rrn26-Rev</i>	<u>GAATTCAGGCGTTATTGCTGTGCTTCC</u> <u>CCATGGTCAGGGTACACTGAACACCAC</u>	449 (мтДНК)
<i>PcoxI*</i>	<i>cox1-For</i> <i>cox1-Rev</i>	<u>GAATTCGAGTTTGACCACTGTTTCTCT</u> <u>CCATGGAAATTTGAGATTCTTCGTTATTACC</u>	525 (мтДНК)
<i>Prn26-GFP-3'UTR_NAD4**</i>	2.3(257)-For nad4-3'UTR-Rev	CCAACCACCACATACCGAAA ACCCGTTTTCTTTCTTTGCTGATCCTCTTATGTATAGTTCATCCATGCCA	2752 (pCK-Prn26- tRNA)

*При клонировании в состав конструкции pCK/GFP/PRmt промоторная область гена *RRN26* или *COX1* из митохондриального генома арабидопсиса встраивается непосредственно перед геном *GFP* по сайтам *EcoRI* (подчеркнут) и *NcoI* (выделен курсивом), фрагмент промоторной области (Y08501.2) выделен **жирным шрифтом**.

Прямой праймер подобран к последовательности линейной плазмиды кукурузы 2.3 т.п.н., находящейся в конструкции pCK-Prn26-tRNA перед последовательностями промотора и гена *GFP*, обратный праймер включает последовательность “3-НТО” митохондриального гена *NAD4* (Y08501.2) арабидопсиса (выделена **жирным шрифтом) и фрагмент гена *GFP*.

Таблица 2. Праймеры для синтеза ДНК-субстратов импорта

Пары праймеров	Нуклеотидная последовательность (5' → 3')	Фрагмент, п.н. (матрица)
2,3(257)-For 2,3(2989)-Rev	CCAACCACCACATACCGAAA ACGCTCTGTAGGATTTGAACC	2893 (pCK-Prn26-tRNA или pCK-Pcox1-tRNA)
2,3(257)-For pJET(930)-Rev	CCAACCACCACATACCGAAA GCAGCTGGCAGCAGAGTTTC	3260 (pJET-Prn26-NAD4)

праймером, содержащим 3'-НТО гена *NAD4* из митохондриального генома арабидопсиса (табл. 1). В результате получена генетическая конструкция pJET-Prn26-NAD4, которая служила матрицей для амплификации ДНК-субстрата импорта Prn26-NAD4.

Получение субстратов для импорта ДНК. ДНК амплифицировали с использованием Taq-полимеразы (“Thermo Scientific”, США) в соответствии с рекомендациями изготовителя. Генетические конструкции получали с помощью специфических пар праймеров, используя один и тот же “прямой” праймер и разные “обратные” праймеры в зависимости от ДНК-матрицы (табл. 2).

Режим амплификации: 94°C – 3 мин (1 цикл); 94°C – 30 с, 60°C – 45 с, 72°C – 180 с (30 циклов); 72°C – 5 мин. ДНК очищали на колонках GeneJET™ PCR Purification Kit (“Thermo Scientific”) согласно инструкции производителя. Качество ПЦР-продуктов оценивали электрофоретически с использованием Gel Doc XR System (“Bio-Rad”, США), количество ДНК определяли с помощью спектрофотометра NanoPhotometer NP80 (“IMPLEN”, Германия).

Выделение митохондрий. Митохондрии получали из трехнедельных растений *A. thaliana* согласно описанному ранее протоколу [21]. Гомогенизацию растительного материала проводили в среде выделения (25 мМ пирогенфосфат натрия, 10 мМ KH_2PO_4 , 300 мМ сахароза, 2 мМ EDTA, 0.5% (в/о) бычьего сывороточного альбумина (БСА), 1% (в/о) поливинилпирролидона-40, 5 мМ цистеин, 20 мМ ас-

корбиновая кислота, pH 7.6). Соотношение веса растительного материала к объему среды – 1 : 12. Для получения грубой митохондриальной фракции проведено два цикла дифференциального центрифугирования при 2500 g (5 мин) и 17400 g (20 мин) в среде промывания (1 мМ EGTA, 50 мМ MOPS, 300 мМ сахароза, 0.2% (в/о) БСА, pH 7.2). Очищенную митохондриальную фракцию получали путем разделения грубой фракции митохондрий в ступенчатом градиенте плотности перколлы (50–28–20%) в течение 40 мин при 40000 g. Суспензию митохондрий отбирали на границе слоев с 50 и 28% перколлы. Количество белка в суспензии митохондрий оценивали по методу Брэдфорда [22].

Определение дыхательного контроля и интактности митохондрий. Коэффициент дыхательного контроля митохондрий (100–150 мкг белка/мл) определяли как отношение скорости поглощения кислорода митохондриями в присутствии ADP (100–200 мкМ) и в отсутствие ADP с использованием в качестве субстратов дыхания сукцината натрия (10 мМ) и глутамата натрия (5 мМ), в среде, содержащей 0.3 М сахарозу, 10 мМ KH_2PO_4 , 10 мМ KCl, 5 мМ MgCl_2 , 0.1% БСА (pH 7.2) [23], с применением платинового кислородного электрода и ячейки Oxytherm system (“Hansatech”, Великобритания). Целостность внешней мембраны митохондрий рассчитывали по разнице скорости аскорбатзависимого стимулированного цитохромом c KCN-чувствительного поглощения кислорода в отсутствие и в присутствии 0.04% Тритона X-100 [24].

Таблица 3. Праймеры, использованные в количественном ПЦР-анализе

Пары праймеров	Нуклеотидная последовательность (5'→3')	ДНК-матрица для ПЦР-РВ, фрагмент, п.н.
NAD4-L NAD4-R	GCATTTTCAGTGGGTTGGTCTGGT AGGGATTGGCACGCTTTCGG	мтДНК (146)
GFP-L GFP-R	GATGTGGAAAACAAGACAGGGGTTT TGGTGAACCGGGCGTACTATTT	Prrn26-tRNA/-NAD4, Pcox1-tRNA (185)

Импорт рекомбинантных ДНК в митохондрии арабидопсиса и транскрипция *in organello*. ДНК импортировали согласно [25]. К 200 мкл буфера импорта (0.4 М сахара, 40 мМ фосфат калия, pH 7.0), содержащего 500 мкг ДНК, добавляли 200 мкг очищенных митохондрий, затем инкубировали при 25°C в течение 30 мин. Обработку митохондрий ДНКазой I (1 ед./мкл) ("Thermo Scientific") проводили в 100 мкл буфера импорта в присутствии 10 мМ MgCl₂ в течение 20 мин при 25°C. Затем отбирали 1/4 пробы, подвергали отмывке в среде промывания, содержащей дополнительно 10 мМ EDTA и 10 мМ EGTA, и экстрагировали мтДНК с целью дальнейшего анализа эффективности импорта ДНК. Оставшуюся часть пробы отмывали в буфере импорта, после чего митохондрии ресуспендировали в 200 мкл буфера транскрипции (330 мМ сахара, 90 мМ KCl, 10 мМ MgCl₂, 12 мМ трицин, 5 мМ KH₂PO₄, 1.2 мМ EGTA, 1 мМ GTP, 2 мМ ДТТ, 2 мМ ADP, 10 мМ сукцинат натрия, 0.15 мМ СТР, 0.15 мМ УТР, pH 7.2) [26] и инкубировали в течение 2–20 ч при легком помешивании (300 об/мин) и температуре 25°C. мтДНК экстрагировали с использованием фенола и буфера экстракции (10 мМ Трис-HCl, 1 мМ EDTA, 1% (в/о) SDS, pH 7.5) в соотношении 1 : 1, как описано ранее [25]. Митохондриальную РНК выделяли после окончания реакции транскрипции при помощи TRIzol reagent ("Ambion", США) согласно инструкции производителя.

Анализ импорта ДНК. Количество импортированной в митохондрии ДНК определяли методом ПЦР в реальном времени (ПЦР-РВ) [25] с использованием набора qPCRmix-HS SYBR ("Евроген", Россия) согласно инструкции производителя. Амплификацию проводили на CFX96 ("Bio-Rad") в следующем режиме: 95°C – 5 мин, 1 цикл; 95°C – 20 с, 60°C – 30 с, 72°C – 30 с, 35 циклов. Данные анализировали при помощи программного обеспечения CFX Manager ("Bio-Rad"). Использовали пары праймеров, специфичные к последовательностям гена *GFP* и гена *NAD4* из митохондриального генома арабидопсиса (табл. 3).

Анализ экспрессии гена *GFP* в изолированных митохондриях. Уровень транскрипта гена *GFP* оценивали с помощью обратной транскрипции с последующей ПЦР-РВ (ОТ-ПЦР-РВ). На первом этапе синтезировали кДНК с использованием в качестве матрицы РНК, экстрагированной из ми-

тохондрий после транскрипции, как описано ранее [27]. Содержание транскрипта определяли с помощью ПЦР-РВ по протоколу, приведенному выше.

Статистическая обработка результатов. Эксперименты проводили не менее чем в трех биологических повторностях. Построение диаграмм выполняли с помощью пакета программ Microsoft Excel. Степень статистической значимости различий оценивали по критерию Стьюдента при уровне значимости (P) ≤ 0.05.

РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ

Для тестирования возможности экспрессии *GFP* в митохондриях клонировали (см. раздел "Экспериментальная часть") промоторные области митохондриальных генов *RRN26* и *COX1* арабидопсиса (рис. 1а). Области для амплификации были выбраны с учетом данных о позиции промоторов, согласно которым сайт инициации транскрипции гена *RRN26* находится в положении (–893) от начала зрелого транскрипта, а промоторная область гена *COX1* содержит два сайта инициации транскрипции (–355) и (–88) [7, 28]. В полученных генетических конструкциях ген *GFP* в 3'-направлении фланкирован последовательностью гена тРНК^{Trp}, входящей в состав митохондриальной ДНК-плазмиды 2.3 т.п.н. кукурузы (рис. 1б). Создана также альтернативная генетическая конструкция, в которой последовательность гена тРНК замещена 3'-НТО гена *NAD4* арабидопсиса (рис. 1а, 1б).

ДНК-субстраты, полученные на основе описанных генетических конструкций, использовали для импорта в митохондрии, выделенные из растений арабидопсиса. Использование нами высокоэнергизованных митохондрий подтверждено величиной дыхательного контроля и результатами оценки интактности наружной митохондриальной мембраны (данные не приведены). Прежде всего, было необходимо убедиться, что используемые субстраты ДНК импортируются в изолированные митохондрии с равной эффективностью вне зависимости от их последовательности. Контролем служил фрагмент ДНК, в котором отсутствовали регуляторные последовательности, контролирующие экспрессию *GFP* (рис. 2). В первой серии экспериментов (рис. 2а) в митохондрии арабидопсиса импортировали фрагменты ДНК, в кото-

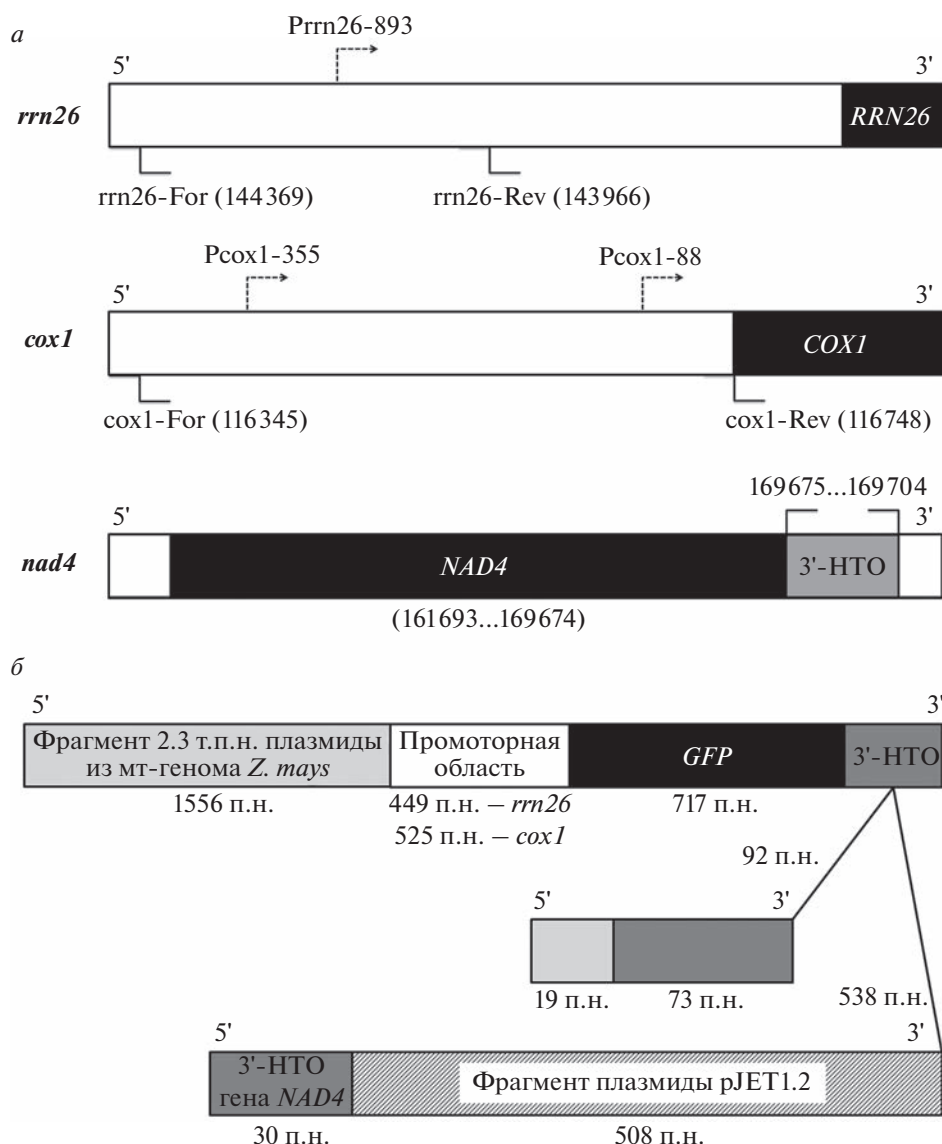


Рис. 1. Регуляторные элементы генетических конструкций, полученных на основе рСК/ГФР/РМт. **а** – Позиции праймеров в митохондриальном геноме арабидопсиса (Y08501.2) для амплификации фрагментов, содержащих промоторную область генов *RRN26* или *COX1* (сплошная линия), включавшую промоторы с сайтами инициации (прерывистая линия); альтернативная 3'-НТО из митохондриального генома арабидопсиса (Y08501.2), принадлежащая гену *NAD4*; **б** – генетические конструкции *Prn26*-GFP или *Pcox1*-GFP, включающие промоторную область генов *RRN26* или *COX1* соответственно, чужеродный ген *GFP* и один из двух вариантов 3'-НТО: последовательности тРНК линейной плазмиды 2.3 т.п.н. из митохондриального генома *Z. mays* (20) или 3'-НТО гена *NAD4* из митохондриального генома арабидопсиса (Y08501.2) с дополнительной последовательностью плазмидного вектора pJET1.2 на 3'-конце. На 5'-концах конструкций *Prn26*-GFP и *Pcox1*-GFP находится фрагмент линейной митохондриальной плазмиды 2.3 т.п.н. (1556 п.н.), являющийся частью исходной конструкции рСК/ГФР/РМт [20].

рых *GFP* был фланкирован в 5'-направлении промоторной областью *Prn26* или *Pcox1*, а в 3'-направлении – последовательностью тРНК^{Трп}. Показан приблизительно равный уровень импорта ДНК всех генетических конструкций. Сходный результат получен при анализе импорта ДНК, содержащей один и тот же промотор *Prn26*, но разные 3'-НТО (последовательности тРНК^{Трп} или 3'-НТО гена *NAD4*) (рис. 2б). Равный

уровень импорта полученных нами конструкций свидетельствовал об адекватности использования данной системы для изучения экспрессии импортированной ДНК в митохондриях арабидопсиса.

Эффективность экспрессии чужеродной ДНК в изолированных митохондриях арабидопсиса изучали в несколько этапов (рис. 3а). Для определения оптимальной продолжительности синтеза целевого транскрипта *GFP in organello* после им-

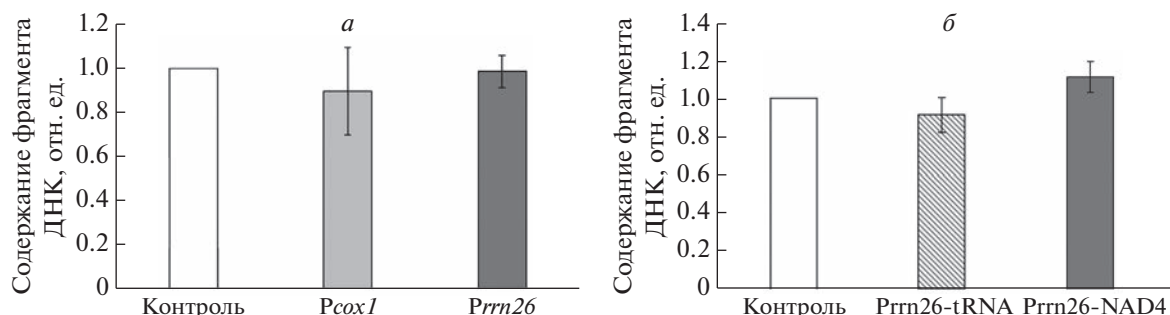


Рис. 2. Анализ эффективности импорта генетических конструкций в изолированные митохондрии арабидопсиса с использованием ПЦР-РВ: импорт ДНК, содержащей промоторы *Prnr26* или *Pcox1* и 3'-НТО из митохондриального генома *Z. mays* (последовательности тРНК) (а); импорт ДНК, включающих в свой состав промотор *Prnr26* и один из двух вариантов 3'-НТО (последовательности тРНК или 3'-НТО гена *NAD4*) (б). Уровень импорта ДНК без регуляторных элементов митохондриального генома (контроль) принят за условную единицу. Допуски на диаграммах обозначают ошибки среднего.

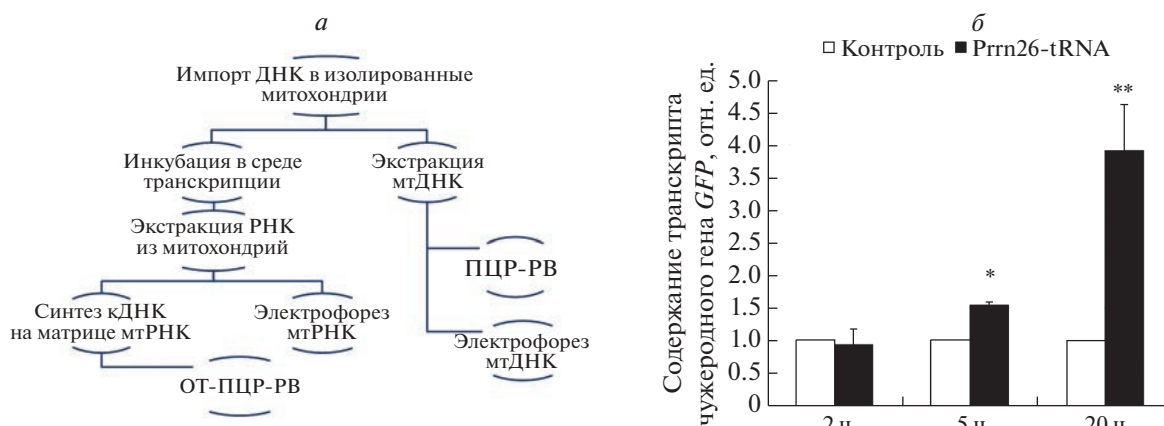


Рис. 3. Исследование уровня транскрипта гена *GFP* под контролем регуляторных элементов митохондриальных генов *in organello*. Схема стандартного эксперимента (а). Зависимость уровня транскрипта гена *GFP* под контролем промотора *Prnr26* от времени (б). Уровень неспецифической транскрипции с ДНК-матрицы без регуляторных элементов (контроль) принят за условную единицу. Допуски на диаграммах обозначают ошибки среднего; * и ** обозначают статистически значимые различия при $P \leq 0.02$ и $P \leq 0.01$ соответственно.

порта генетической конструкции *Prnr26*-tRNA в изолированные митохондрии арабидопсиса инкубацию митохондрий проводили в течение 2, 5 или 20 ч (рис. 3б). Показано, что содержание транскрипта *GFP* зависело от времени инкубации митохондрий: транскрипт детектировался во всех трех временных точках, при этом уровень сигнала возрастал с увеличением времени инкубации. В качестве контроля использовали фрагмент ДНК сопоставимой длины (около 3 т.п.н.), в составе которого отсутствовали промотор и 3'-НТО. Уровень сигнала, детектируемого в ОТ-ПЦР после транскрипции *in organello* с этой конструкции, мы расценивали как уровень неспецифической транскрипции. Подобная транскрипция, инициируемая без участия каких-либо промоторных последовательностей, типична для митохондрий растений, детектируется в системах *in vitro* [28, 29] и, возможно, происходит *in vivo* [10]. Исходя из полученных данных, в последующих экспери-

ментах мы использовали наиболее длительный период инкубации (20 ч), так как в этом случае нарабатывается наибольшее количество транскрипта, уровень которого существенно превышает уровень транскрипта, синтезируемого с контрольной конструкции (рис. 3).

В следующей серии экспериментов для импорта в митохондрии использовали две генетические конструкции, содержащие промоторную область гена *RRN26*, но различающиеся входящей в их состав 3'-НТО (рис. 4а). НТО играют важную роль в контроле экспрессии генов в митохондриях растений, отвечая за стабильность транскрипта, поэтому тестирование вариантов генетической конструкции, содержащих последовательности различных 3'-НТО, было необходимым этапом исследования. Используемые конструкции содержали два варианта последовательности 3'-НТО: (1) участок митохондриальной плазмидоподобной ДНК кукурузы, кодирующий тРНК^{Trp} или

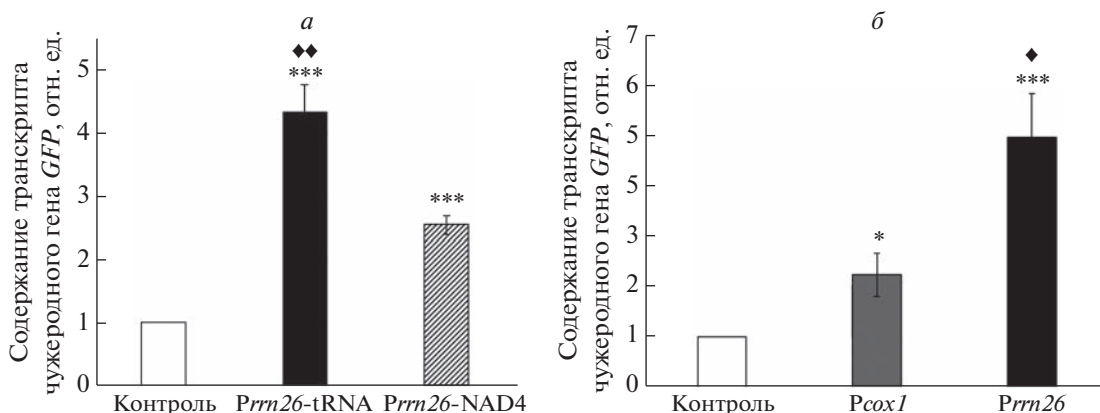


Рис. 4. Уровень транскрипта *GFP*, синтезирующегося *in organello* в митохондриях арабидопсиса, при использовании разных генетических конструкций: включающих в свой состав *Prnr26* и один из двух вариантов 3'-НТО (последовательности тРНК^{Trp} или 3'-НТО гена *NAD4*) (а); содержащих *Pcox1* или *Prnr26* и 3'-НТО из митохондриального генома *Z. mays* (последовательности тРНК) (б). Уровень неспецифической транскрипции с ДНК-матрицы без регуляторных элементов (контроль) принят за условную единицу. Допуски на диаграммах обозначают ошибки среднего; * (♦), ** (♦♦) и *** обозначают статистически значимые различия при $P \leq 0.05$, $P \leq 0.01$ и $P \leq 0.001$ соответственно. * обозначает отличия от контроля, ♦ обозначает отличия от *Prnr26-NAD4* (а) / *Pcox1* (б).

(2) 30-нуклеотидную последовательность на 3'-конце транскрипта гена *NAD4* арабидопсиса, служащую сайтом связывания белка MTSF1, играющего ключевую роль в стабилизации этого транскрипта [12].

Установлено, что содержание транскриптов, образующихся с ДНК-субстрата, несущего тРНК^{Trp} в 3'-НТО, выше, чем при использовании ДНК-субстрата, содержащего последовательность 3'-НТО гена *NAD4* (рис. 4а). Эти результаты позволяют судить о более эффективной стабилизации транскрипта, синтезируемого в системе *in organello*, в случае присутствия на его 3'-конце вторичных структур, содержащихся в молекуле тРНК.

Следующий этап работы — тестирование уровня транскрипции с различными промоторными последовательностями в митохондриях арабидопсиса — проведен с использованием генетических конструкций, содержащих последовательность тРНК^{Trp} в 3'-НТО (рис. 4б). Уровень транскрипции при использовании субстрата импорта, в котором *GFP* находился под контролем промоторной области *Prnr26*, был более высоким, чем в случае субстрата, в котором экспрессия гена контролировалась промоторной областью *Pcox1* (рис. 4б). Эти результаты согласуются с ранее полученными данными о дифференциальной экспрессии митохондриальных генов арабидопсиса, находящихся под контролем различных промоторов [8].

ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

На первом этапе работы нами показано, что ДНК-субстраты, содержащие либо разные промоторные области, либо разные 3'-НТО, импортировались в митохондриальный матрикс с та-

кой же эффективностью, как и фрагмент ДНК без регуляторных элементов. Эти результаты согласуются с данными о том, что процесс импорта ДНК не зависит от нуклеотидной последовательности [20].

Анализ содержания транскрипта чужеродного гена, нарабатанного в митохондриальном матриксе, после нескольких часов инкубации митохондрий выявил достаточно высокий уровень неспецифической транскрипции в условиях *in organello*. Об этом свидетельствовало обнаружение транскрипции гена *GFP* с субстрата импорта ДНК, не содержавшего регуляторных элементов (рис. 3). Неспецифические сайты инициации транскрипции — общая ситуация в случае транскрипции с митохондриальными промоторами в условиях *in vitro*, инициируемой обеими митохондриальными РНК-полимеразами, особенно RPO_{Trp} [28]. Этой РНК-полимеразе свойственна высокая транскрипционная активность *in vitro* даже при отсутствии в ДНК-субстрате промоторных последовательностей [28]. Исходя из этого, было необходимо определить продолжительность инкубации митохондрий после импорта ДНК, при которой содержание транскрипта, синтезированного с митохондриальных промоторов, будет существенно превышать уровень неспецифической транскрипции. Оптимальный результат получен после 20 ч инкубации, по истечении которой уровень целевого транскрипта в существенной степени превышал фоновый (рис. 3).

На следующем этапе работы мы проанализировали эффективность двух вариантов 3'-концевых последовательностей, обладающих потенциальной способностью стабилизировать синтезируемый транскрипт — шпилькоподобной структуры,

образуемой тРНК^{Трп}, и сайта связывания белка MTSF1, специфичного для гена *NAD4* арабидопсиса. Нами проведен анализ уровня нарабатываемого транскрипта *GFP* под контролем промоторной области гена *RRN26* и каждого из двух вариантов 3'-НТО. В результате оказалось, что последовательность тРНК^{Трп} на 3'-конце образующегося транскрипта более эффективно обеспечивает его стабильность, чем 3'-НТО гена *NAD4* арабидопсиса (рис. 4а). Такой результат выглядит достаточно неожиданным, особенно учитывая видовую специфичность последовательности 3'-НТО гена *NAD4*. MTSF1 — РНК-связывающий белок, относящийся к семейству PPR-белков, которое включает свыше 450 представителей, участвующих во всех этапах метаболизма РНК в органеллах арабидопсиса [30]. Показано, что 20-нуклеотидной последовательности, локализованной непосредственно на 3'-конце зрелого транскрипта *NAD4*, достаточно для связывания белка MTSF1 *in vitro*. Предполагается, что сразу после синтеза транскрипта-предшественника в митохондриях MTSF1 связывается с этой последовательностью и препятствует продвижению полинуклеотидфосфорилазы и других экзонуклеаз, расщепляющих РНК в направлении 3' → 5', в результате чего происходит стабилизация транскрипта и формирование его 3'-конца [12]. Тем не менее, нельзя исключить, что дополнительные последовательности могут играть роль в распознавании транскрипта *NAD4* белковым аппаратом экспрессии *in vivo*. Известно, что метаболизм РНК в митохондриях растений — сложный и не до конца изученный процесс, включающий в себя ряд взаимосвязанных этапов [3, 19]. Можно предположить, что образующийся в нашей системе чужеродный транскрипт не распознается в полной мере белковым аппаратом, ответственным за созревание и стабилизацию специфических митохондриальных РНК. Альтернативным объяснением полученного результата может быть отсутствие свободного белка MTSF1 в митохондриях, использованных в наших экспериментах.

С другой стороны, шпилькоподобные структуры, образуемые последовательностью тРНК^{Трп} митохондрий кукурузы, оказались достаточно эффективными для защиты транскрипта от деградации экзонуклеазами. Можно предположить, что этот механизм менее специфичен и зависит от меньшего числа факторов, что сделало его более эффективным в отношении стабилизации чужеродного транскрипта *in organello*. Ранее в изолированных митохондриях крысы была выявлена различная стабильность транскриптов, экспрессируемых с импортированных в митохондрии генетических конструкций, несущих репортерные гены с различными 3'-концевыми последовательностями [5]. В том случае, когда вслед за последовательностью репортерного гена в конструкции

располагалась последовательность, кодирующая тРНК, целевой транскрипт оказывался существенно более стабильным, чем когда в 3'-области находилась поли(А)-последовательность. Таким образом, несмотря на существенные различия в организации и экспрессии митохондриального генома растений и животных, в обоих случаях наличие последовательности тРНК в 3'-области оказывается эффективным механизмом стабилизации транскрипта в системе *in organello*.

Далее мы провели сравнительный анализ содержания транскриптов *GFP*, находящихся под контролем промоторных областей двух митохондриальных генов — *RRN26* и *COX1*. Эффективность функционирования промоторов митохондриальных генов арабидопсиса и других растений ранее изучали в условиях *in vitro* с использованием митохондриального экстракта [29, 31] либо рекомбинантных РНК-полимераз [28], однако в системе изолированных митохондрий подобная работа проведена впервые. Нами показано, что уровень транскрипта *GFP* под контролем *Prm26* выше, чем под контролем *Pcox1* (см. рис. 4б), из чего следует, что интенсивность транскрипции чужеродного гена в митохондриях растений *in organello* зависит от эффективности промотора митохондриального гена, входящего в состав импортируемой генетической конструкции.

Выявленные *in organello* различия в уровне экспрессии гена *GFP* под контролем разных промоторных последовательностей согласуются с данными о закономерностях экспрессии генов *RRN26* и *COX1 in vivo* [8]. Как показано с помощью метода *gus-on* транскрипции, интенсивность транскрипции гена *RRN26* в листьях арабидопсиса примерно в 2 раза превышает интенсивность транскрипции гена *COX1*, что отлично коррелирует с нашими данными [8]. *RRN26* представляет собой ген, наиболее активно транскрибируемый в митохондриях арабидопсиса [8]. В то же время, в экспериментах *in vitro* с использованием рекомбинантной РНК-полимеразы RPO_{Tm} показано, что уровень транскрипции, инициируемой с единственного промотора гена *RRN26*, многократно ниже интенсивности транскрипции с промотора *Patp6-1-200* гена *atp6*, транскрибирующегося *in vivo* на относительно невысоком уровне [28]. Из этого следует, что данные, полученные нами в системе *in organello*, корректно отражают эффективность митохондриальных промоторных областей, наблюдаемую в растениях арабидопсиса *in vivo*, в отличие от результатов, полученных в системе *in vitro*.

Ранее различия в уровне экспрессии репортерных генов, контролируемых двумя основными промоторными областями митохондриального генома млекопитающих — HSP и LSP — наблюдали в изолированных митохондриях крысы [5]. Уровень транскрипта, инициируемого с промотора

HSP, был в несколько раз выше, чем транскрипта, находившегося под контролем промотора LSP. Интересно отметить, что и в этом случае относительная интенсивность транскрипции была близка к показанной ранее *in vivo* [32], но не в системе *in vitro*, что служит указанием на отсутствие в последнем случае дополнительных факторов, необходимых для корректного протекания транскрипции [5, 33].

Вопрос о существовании дополнительных транскрипционных факторов в митохондриях растений остается открытым до настоящего времени. В то время как в митохондриях животных обнаружено несколько транскрипционных факторов [34], только косвенные данные свидетельствуют в пользу существования таких белков в митохондриях растений [35–37]. Наши данные, указывающие на существенные различия в относительной интенсивности транскрипции, инициируемой с промоторной области гена *RRN26 in organello* и *in vitro*, свидетельствуют в пользу существования дополнительных факторов транскрипции. Ранее было показано, что 9-нуклеотидный мотив CATAAGAGA – центральный элемент промоторной области гена *RRN26* и ряда других митохондриальных промоторов [7], слабо распознается рекомбинантными RPOTm и RPO^{Tnp}, находясь в составе промотора *Prn26*, и не распознается вообще в составе остальных проверенных промоторов (*Patp6*-1-156, *Patp9*-239 и *Patp8*-157) [28]. По нашему мнению, высокая эффективность этого промотора *in organello*, показанная нами, говорит в пользу участия в распознавании этой последовательности пока не идентифицированных белковых факторов митохондриальной локализации.

Следует отметить, что для инициации специфической транскрипции с митохондриальных промоторов арабидопсиса рекомбинантными РНК-полимеразами *in vitro* строго необходим ДНК-субстрат с суперспирализованной структурой [28]. Импорт суперспирализованных молекул ДНК невозможно обеспечить в условиях *in organello* в силу специфичности процесса переноса ДНК в митохондрии в отношении линейной ДНК [20], однако использование нами линейных субстратов позволило, тем не менее, выявить специфическую транскрипцию чужеродного гена, находящегося под контролем митохондриальных промоторов. Сходным образом линейная ДНК оказывается подходящим субстратом для инициации транскрипции с митохондриальных промоторов при использовании в системе *in vitro* грубого экстракта митохондрий, содержащего все митохондриальные белки [29, 38]. Эти данные в совокупности указывают на то, что в митохондриях растений присутствуют белковые факторы, модулирующие изменение конформации ДНК таким образом, чтобы сделать ее доступной для синтеза РНК.

Таким образом, нами впервые показана возможность детекции транскриптов чужеродного гена *GFP*, находящегося под контролем промоторных областей митохондриальных генов *COX1* или *RRN26*, после импорта рекомбинантных ДНК-матриц в изолированные митохондрии арабидопсиса. Обнаруженные ранее различия в уровне транскрипции митохондриальных генов *in vivo*, инициируемой с промоторных областей *Pcox1* или *Prn26*, подтверждаются в разработанной нами системе изучения экспрессии чужеродных генов *in organello*. При этом установлено, что уровень транскрипта гена *GFP*, синтезированного *in organello*, выше при использовании в качестве 3'-НТО последовательности тРНК^{Trp}, чем в присутствии 3'-НТО гена *NAD4* арабидопсиса. Полученные результаты создают возможность использования разработанной нами системы импорта ДНК *in organello* для изучения особенностей экспрессии митохондриальных генов растений, а также для подбора митохондриальных генетических регуляторных элементов, наиболее эффективных для обеспечения экспрессии чужеродного гена при трансформации митохондрий *in vivo*.

В работе использовано оборудование ЦКП “Биоаналитика” СИФИБР СО РАН (Иркутск).

Исследование выполнено за счет гранта Российского научного фонда № 22-74-00114, <https://rscf.ru/project/22-74-00114/>.

Настоящая статья не содержит описания выполненных авторами исследований с участием людей или использованием животных в качестве объектов.

Авторы заявляют об отсутствии конфликта интересов.

СПИСОК ЛИТЕРАТУРЫ

1. Larosa V., Remacle C. (2013) Transformation of the mitochondrial genome. *Int. J. Dev. Biol.* **57**, 659–665. <https://doi.org/10.1387/ijdb.130230cr>
2. Remacle C., Larosa V., Salinas T., Hamel P., Subrahmanian N., Bonnefoy N., Kempken F. (2012) Transformation and nucleic acid delivery to mitochondria. In: *Genomics of Chloroplasts and Mitochondria. Advances in Photosynthesis and Respiration*. **35**. Eds Bock R., Knoop V. Dordrecht: Springer, https://doi.org/10.1007/978-94-007-2920-9_19
3. Hammani K., Giegé P. (2014) RNA metabolism in plant mitochondria. *Trends Plant Sci.* **19**, 380–389. <https://doi.org/10.1016/j.tplants.2013.12.008>
4. Константинов Ю.М., Дитриш А., Вебер-Лотфи Ф., Ибрагим Н., Клименко Е.С., Тарасенко В.И., Болотова Т.А., Кулинченко М.В. (2016) Импорт ДНК в митохондрии. *Биохимия*. **81**, 1307–1321.
5. Koulintchenko M., Temperley R.J., Mason P.A., Dietrich A., Lightowers R.N. (2006) Natural competence of mammalian mitochondria allows the molecular investigation of mitochondrial gene expression. *Hum.*

- Mol. Genet.* **15**, 143–154.
<https://doi.org/10.1093/hmg/ddi435>
6. Tarasenko T.A., Klimenko E.S., Tarasenko V.I., Koulintchenko M.V., Dietrich A., Weber-Lotfi F., Konstantinov Y.M. (2021) Plant mitochondria import DNA via alternative membrane complexes involving various VDAC isoforms. *Mitochondrion*. **60**, 43–58.
<https://doi.org/10.1016/j.mito.2021.07.006>
 7. Kühn K., Weihe A., Börner T. (2005) Multiple promoters are a common feature of mitochondrial genes in *Arabidopsis*. *Nucl. Acids Res.* **33**, 337–346.
<https://doi.org/10.1093/nar/gki179>
 8. Kühn K., Richter U., Meyer E., Delannoy E., de Longevialle A.F., O'Toole N., Börner T., Millar A., Small I., Whelan J. (2009) Phage-type RNA polymerase RPOTmp performs gene-specific transcription in mitochondria of *Arabidopsis thaliana*. *Plant Cell*. **21**, 2762–2779.
<https://doi.org/10.1105/tpc.109.068536>
 9. Moller I.M., Rasmusson A.G., Van Aken O. (2021) Plant mitochondria – past, present and future. *Plant J.* **108**, 912–959.
<https://doi.org/10.1111/tpj.15495>
 10. Holec S., Lange H., Kuhn K., Alioua M., Borner T., Gagliardi D. (2006) Relaxed transcription in *Arabidopsis* mitochondria is counterbalanced by RNA stability control mediated by polyadenylation and polynucleotide phosphorylase. *Mol. Cell. Biol.* **26**, 2869–2876.
<https://doi.org/10.1128/MCB.26.7.2869-2876.2006>
 11. Perrin R., Meyer E.H., Zaepfel M., Kim Y.J., Mache R., Grienemberger J.M., Gualberto J.M., Gagliardi D. (2004) Two exoribonucleases act sequentially to process mature 3'-ends of atp9 mRNAs in *Arabidopsis* mitochondria. *J. Biol. Chem.* **279**, 25440–25446.
<https://doi.org/10.1074/jbc.M401182200>
 12. Haili N., Arnal N., Quadrado M., Amiar S., Tcherkez G., Dahan J., Briozzo P., Colas des Francs-Small C., Vrielynck N., Mireau H. (2013) The pentatricopeptide repeat MTSF1 protein stabilizes the nad4 mRNA in *Arabidopsis* mitochondria. *Nucl. Acids Res.* **41**, 6650–6663.
<https://doi.org/10.1093/nar/gkt337>
 13. Ruwe H., Wang G., Gusewski S., Schmitz-Linneweber C. (2016) Systematic analysis of plant mitochondrial and chloroplast small RNAs suggests organelle-specific mRNA stabilization mechanisms. *Nucl. Acids Res.* **44**, 7406–7417.
<https://doi.org/10.1093/nar/gkw466>
 14. Forner J., Weber B., Thuss S., Wildum S., Binder S. (2007) Mapping of mitochondrial mRNA termini in *Arabidopsis thaliana*: T-elements contribute to 5' and 3'-end formation. *Nucl. Acids Res.* **35**, 3676–3692.
<https://doi.org/10.1093/nar/gkm270>
 15. MacIntosh G.C., Castandet B. (2020) Organellar and secretory ribonucleases: major players in plant RNA homeostasis. *Plant Physiol.* **183**, 1438–1452.
<https://doi.org/10.1104/pp.20.00076>
 16. Dombrowski S., Brennicke A., Binder S. (1997) 3'-Inverted repeats in plant mitochondrial mRNAs are processing signals rather than transcription terminators. *EMBO J.* **16**, 5069–5076.
<https://doi.org/10.1093/emboj/16.16.5069>
 17. Kuhn J., Tengler U., Binder S. (2001) Transcript lifetime is balanced between stabilizing stem-loop structures and degradation promoting polyadenylation in plant mitochondria. *Mol. Cell. Biol.* **21**, 731–742.
<https://doi.org/10.1128/MCB.21.3.731-742.2001>
 18. Wang C., Aubé F., Planchard N., Quadrado M., Dargel-Graffin C., Nogué F., Mireau H. (2017) The pentatricopeptide repeat protein MTSF2 stabilizes a nad1 precursor transcript and defines the 3' end of its 5'-half intron. *Nucl. Acids Res.* **45**, 6119–6134.
<https://doi.org/10.1093/nar/gkx162>
 19. Wang C., Blondel L., Quadrado M., Dargel-Graffin C., Mireau H. (2022) Pentatricopeptide repeat protein MITOCHONDRIAL STABILITY FACTOR 3 ensures mitochondrial RNA stability and embryogenesis. *Plant Physiol.* **190**, 669–681.
<https://doi.org/10.1093/plphys/kiac309>
 20. Koulintchenko M., Konstantinov Y., Dietrich A. (2003) Plant mitochondria actively import DNA via the permeability transition pore complex. *EMBO J.* **22**, 1245–1254.
<https://doi.org/10.1093/emboj/cdg128>
 21. Sweetlove L.J., Taylor N.L., Leaver C.J. (2007) Isolation of intact, functional mitochondria from the model plant *Arabidopsis thaliana*. *Meth. Mol. Biol.* **372**, 125–136.
https://doi.org/10.1007/978-1-59745-365-3_9
 22. Bradford M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248–254.
[https://doi.org/10.1016/0003-2697\(76\)90527-3](https://doi.org/10.1016/0003-2697(76)90527-3)
 23. Douce R., Neuburger M. (1989) The uniqueness of plant mitochondria. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **40**, 371–414.
<https://doi.org/10.1146/annurev.pp.40.060189.002103>
 24. Tarasenko T.A., Subota I.Yu., Tarasenko V. I., Konstantinov Y.M., Koulintchenko M.V. (2020) Plant mitochondrial subfractions have different ability to import DNA. *Theor. Exp. Plant Physiol.* **32**, 5–18.
<https://doi.org/10.1007/s40626-020-00167-w>
 25. Тарасенко Т.А., Тарасенко В.И., Кулинченко М.В., Клименко Е.С., Константинов Ю.М. (2019) Импорт ДНК в митохондрии растений: комплексный подход для изучения *in organello* и *in vivo*. *Биохимия*. **84**, 1036–1048.
<https://doi.org/10.1134/S032097251907011X>
 26. Farré J.C., Araya A. (2001) Gene expression in isolated plant mitochondria: high fidelity of transcription, splicing and editing of a transgene product in electroporated organelles. *Nucl. Acids Res.* **29**, 2484–2491.
<https://doi.org/10.1093/nar/29.12.2484>
 27. Tarasenko V.I., Katyshev A.I., Yakovleva T.V., Garnik E.Y., Chernikova V.V., Konstantinov Y.M., Koulintchenko M.V. (2016) RPOTmp, an *Arabidopsis* RNA polymerase with dual targeting, plays an important role in mitochondria, but not in chloroplasts. *J. Exp. Botany*. **67**, 5657–5669.
<https://doi.org/10.1093/jxb/erw327>
 28. Kühn K., Bohne A.V., Liere K., Weihe A., Börner T. (2007) *Arabidopsis* phage-type RNA polymerases: accurate *in vitro* transcription of organellar genes. *Plant Cell*. **19**, 959–971.
<https://doi.org/10.1105/tpc.106.046839>

29. Binder S., Hatzack F., Brennicke A. (1995) A novel pea mitochondrial *in vitro* transcription system recognizes homologous and heterologous mRNA and tRNA promoters. *J. Biol. Chem.* **270**, 22182–2218. <https://doi.org/10.1074/jbc.270.38.22182>
30. Rovira A.G., Smith A.G. (2019) PPR proteins — orchestrators of organelle RNA metabolism. *Physiol. Plant.* **166**, 451–459. <https://doi.org/10.1111/ppl.12950>
31. Hanic-Joyce P.J., Gray M.W. (1991) Accurate transcription of a plant mitochondrial gene *in vitro*. *Mol. Cell Biol.* **11**, 2035–2039. <https://doi.org/10.1128/mcb.11.4.2035-2039.1991>
32. Attardi G., Chomyn A., King M.P., Kruse B., Polosa P.L., Murdter N.N. (1990) Regulation of mitochondrial gene expression in mammalian cells. *Biochem. Soc. Trans.* **18**, 509–513. <https://doi.org/10.1042/bst0180509>
33. Micol V., Fernández-Silva P., Attardi G. (1997) Functional analysis of *in vivo* and *in organello* footprinting of HeLa cell mitochondrial DNA in relationship to ATP and ethidium bromide effects on transcription. *J. Biol. Chem.* **272**, 18896–18904. <https://doi.org/10.1074/jbc.272.30.18896>
34. Kotrys A.V., Szczesny R.J. (2019) Mitochondrial gene expression and beyond—novel aspects of cellular physiology. *Cells*. **9**, 17. <https://doi.org/10.3390/cells9010017>
35. Newton K.J., Winberg B., Yamato K., Lupold S., Stern D.B. (1995) Evidence for a novel mitochondrial promoter preceding the *cox2* gene of perennial teosintes. *EMBO J.* **14**, 585–593. <https://doi.org/10.1002/j.1460-2075.1995.tb07034.x>
36. Xiao S., Zang J., Pei Y., Liu J., Liu J., Song W., Shi Z., Su A., Zhao J., Chen H. (2020) Activation of mitochondrial *orf355* gene expression by a nuclear-encoded DREB transcription factor causes cytoplasmic male sterility in maize. *Mol. Plant.* **13**, 1270–1283. <https://doi.org/10.1016/j.molp.2020.07.002>
37. Тарасенко В.И., Субота И.Ю., Кобзев В.Ф., Константинов Ю.М. (2005) Выделение митохондриальных белков, специфично связывающихся с промоторной областью гена *cox1* кукурузы. *Мол. биология*. **39**(3), 394–402.
38. Rapp W.D., Stern D.B. (1992) A conserved 11 nucleotide sequence contains an essential promoter element of the maize mitochondrial *atp1* gene. *EMBO J.* **11**, 1065–1073. <https://doi.org/10.1002/j.1460-2075.1992.tb05145.x>

Differential Expression of a Foreign Gene in Arabidopsis Mitochondria *in organello*

V. I. Tarasenko¹, T. A. Tarasenko¹*, I. V. Gorbenko¹,
Yu. M. Konstantinov¹, and M. V. Koulintchenko^{1,2}

¹ Siberian Institute of Plant Physiology and Biochemistry Siberian Branch, Russian Academy of Sciences, Irkutsk, 664033 Russia

² Kazan Institute of Biochemistry and Biophysics, Kazan Scientific Center, Russian Academy of Sciences, Kazan, 420111 Russia

*e-mail: bolotova_t.a@mail.ru

Genetic transformation of higher eukaryotes mitochondria *in vivo* is one of the unresolved and important problems. For efficient expression of foreign genetic material in mitochondria, it is necessary to select regulatory elements that ensure a high level of transcription and transcript stability. This work is aimed at studying the effectiveness of regulatory elements of mitochondrial genes flanking exogenous DNA using the phenomenon of natural competence of plant mitochondria. For this purpose, genetic constructs carrying the *GFP* gene under the control of the promoter regions of the *RRN26* or *COX1* genes and one of the two 3'-untranslated regions (3'-UTR) of mitochondrial genes were imported into isolated *Arabidopsis* mitochondria, followed by transcription *in organello*. It was shown that the level of *GFP* expression under the control of promoters of the *RRN26* or *COX1* genes *in organello* correlates with the level of transcription of these genes observed *in vivo*. At the same time, the presence of the tRNA^{Trp} sequence in the 3'-UTR leads to a higher level of the *GFP* transcript than the presence in this region of the 3'-UTR of the *NAD4* gene containing the binding site of the MTSF1 protein. The results obtained open up prospects for creating a system for efficient transformation of the mitochondrial genome.

Keywords: regulation of gene expression, DNA import, *GFP*, promoter region, untranslated region, RPOTmp, *Arabidopsis thaliana* mitochondria, *in organello*

ГЕНОМИКА.
ТРАНСКРИПТОМИКА

УДК 577.21+616.13.002.2-004.6

ИДЕНТИФИКАЦИЯ ФУНКЦИОНАЛЬНО ЗНАЧИМЫХ
ПОЛИМОРФНЫХ ВАРИАНТОВ В ГЕНАХ микроРНК
ПРИ АТЕРОСКЛЕРОЗЕ СОННЫХ АРТЕРИЙ

© 2023 г. А. А. Зарубин^а, К. В. Маннанова^б, Ю. А. Королёва^а, А. А. Слепцов^а, М. С. Кузнецов^с,
Б. Н. Козлов^с, М. С. Назаренко^{а, б, *}

^аНаучно-исследовательский институт медицинской генетики, Томский национальный исследовательский
медицинский центр Российской академии наук, Томск, 634050 Россия

^бСибирский государственный медицинский университет Министерства здравоохранения Российской Федерации,
Томск, 634050 Россия

^сНаучно-исследовательский институт кардиологии, Томский национальный исследовательский медицинский центр
Российской академии наук, Томск, 634034 Россия

*e-mail: maria.nazarenko@medgenetics.ru

Поступила в редакцию 19.07.2022 г.

После доработки 05.10.2022 г.

Принята к публикации 05.10.2022 г.

МикроРНК, участвующие в регуляции экспрессии генов, вовлечены в патогенез широкого спектра многофакторных заболеваний, в том числе атеросклероза, его факторов риска и осложнений, что делает актуальным изучение функционально значимых полиморфных вариантов в генах микроРНК у пациентов с клинически выраженным атеросклерозом сонных артерий. Нами проведено секвенирование экзотов и микроРНК из одних и тех же образцов атеросклеротических бляшек сонных артерий мужчин ($n = 8$, возраст 66–71 лет) с клинически выраженным атеросклерозом. В подтверждающем исследовании использовали расширенную выборку из 112 пациентов и 72 индивидов контрольной группы (этнических славян, жителей Западной Сибири). В образцах сонных артерий, пораженных атеросклерозом, обнаружены 206 пре-микроРНК и 76 зрелых микроРНК, содержащих 321 и 97 однонуклеотидных вариантов (SNV) соответственно. Объединение данных секвенирования экзотов и микроРНК показало, что 18 генов микроРНК, которые экспрессируются в атеросклеротических бляшках до зрелого состояния, несут 24 SNV. Из них наибольшей потенциальной функциональной значимостью в отношении экспрессии микроРНК, предсказанной *in silico*, обладали варианты rs2910164:C>G (*MIR146A*), rs2682818:A>C (*MIR618*), rs3746444:A>G (*MIR499A*), rs776722712:C>T (*MIR186*) и rs199822597:G>A (*MIR363*). В образцах, полученных от пациентов с генотипом AC rs2682818 гена *MIR618*, выявлено снижение уровня экспрессии miR-618 по сравнению с генотипом CC ($\log_2FC = 4.8$; $p = 0.012$). Обнаружено, что rs2910164:C (*MIR146A*) ассоциирован с риском развития клинически выраженного атеросклероза сонных артерий ($OR = 2.35$; 95%CI: 1.43–3.85; $p = 0.001$). Интегративный подход позволил выявить новые наиболее информативные с точки зрения предсказания функциональной значимости полиморфные варианты в генах микроРНК. Вариант rs2682818:A>C (*MIR618*) может быть связан с изменением экспрессии miR-618 в атеросклеротических бляшках сонных артерий, а rs2910164:C (*MIR146A*) ассоциирован с риском развития клинически выраженного атеросклероза.

Ключевые слова: микроРНК, секвенирование, атеросклероз сонных артерий

DOI: 10.31857/S0026898423030205, **EDN:** CHZTLI

ВВЕДЕНИЕ

Атеросклеротическое поражение сонных артерий часто приводит к развитию таких осложнений, как хроническая ишемия головного мозга, сосудистая деменция, транзиторная ишемическая атака и ишемический инсульт [1]. На сегодняшний день значительно улучшилось понимание патофизиологических процессов, лежащих в основе формирования атеросклероза и механиз-

мов его клинических осложнений, открыты ключевые сигнальные пути и молекулы, участвующие в инициации и прогрессии атеросклеротических бляшек [2].

Среди таких молекул особый интерес представляют микроРНК — малые (miRNA; в среднем 22 нуклеотида) одноцепочечные некодирующие РНК, которые регулируют экспрессию генов в ядре на уровне транскрипции и на посттран-

скрипционном уровне в цитоплазме, связываясь с 3'-нетранслируемой областью мРНК. Связывание может быть как полностью комплементарным с деградацией мРНК гена-мишени, так и частично комплементарным (более характерно для микроРНК) с подавлением трансляции [3]. При частичной комплементарности одна микроРНК может иметь сразу несколько сотен мРНК-мишеней. Верно и обратное — одна мРНК может связываться с разными микроРНК.

Полиморфизм в генах микроРНК (в регуляторных участках, в при- и пре-микроРНК, в области нуклеотидной последовательности зрелой микроРНК), а также в генах их биогенеза и генах-мишенях может быть функционально значимым для формирования атеросклероза и его осложненного течения [4]. Варианты в генах микроРНК могут влиять на процессинг и уровень экспрессии соответствующих молекул, на их биогенез, активность, а также на взаимодействие зрелой микроРНК с ее мРНК-мишенью [5]. Однако нуклеотидная вариабельность в области генов (при/пре-) микроРНК при сердечно-сосудистых заболеваниях изучена недостаточно детально [5, 6].

Многочисленные исследования показывают, что микроРНК участвуют почти во всех этапах атерогенеза, способствуя или подавляя данный патологический процесс [7–9]. В базе данных HMDD v3.0: the Human microRNA Disease Database (обновление от 27 марта 2019 г.) содержится информация о 116 экспериментально подтвержденных микроРНК, связанных с атеросклерозом, из них 26 микроРНК ассоциированы с атеросклерозом сонных артерий и 67 — с ишемическим инсультом [10].

В большинстве работ, направленных на поиск ассоциаций полиморфных вариантов генов микроРНК (miR-let-7, miR-27a, -146a, -149, -196a2, -200b, -423, -499, -618, -4513) с клиническими осложнениями атеросклероза, анализировали преимущественно отдельные однонуклеотидные варианты (SNV) в популяциях Азии и получали противоречивые результаты [11–16].

В других работах изменение экспрессии микроРНК в клетках сосудов, пораженных атеросклерозом, анализировали с использованием различных групп сравнения и методов, включая ПЦР в режиме реального времени и технологии микрочипов [17, 18]. В результате идентифицирована 31 микроРНК (miR-1, let-7f, -9, -10a, -10b, -16, -19a, -19b, -21, -22, -24, -25, -29b, -29c, -34a, -92a, -100, -106b, -125a, -127, -133a, -133b, -143, -145, -146a, -150, -155, -221, -223, -486 и -497), экспрессия которых в атеросклеротических бляшках коронарных и сонных артерий статистически значимо отличается от экспрессии в интактных тканях сосудов, либо в нестабильных (симптоматических) и стабильных (асимптоматических) атеросклеро-

тических бляшках в двух или более исследованиях [19]. Однако вплоть до настоящего времени нет работ, в которых анализируется вариабельность генов микроРНК и экспрессия микроРНК в атеросклеротических бляшках у одних и тех же пациентов.

В связи с этим актуальным представляется исследование функционально значимых вариантов в генах микроРНК и анализ их ассоциации с заболеваниями, обусловленными атеросклерозом, в европейских популяциях. Изучение полиморфных вариантов в генах микроРНК, участвующих в патогенезе атеросклероза, поможет выявить новые молекулярные механизмы развития заболевания, а также возможные биомаркеры и мишени для таргетной терапии. Полученные данные могут быть использованы для идентификации и стратификации индивидов, прогнозирования риска развития клинических осложнений атеросклероза.

В нашей работе с помощью секвенирования экзозома и микроРНК охарактеризован спектр функционально значимых однонуклеотидных полиморфных вариантов в генах микроРНК в атеросклеротических бляшках сонных артерий одних и тех же пациентов.

ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

Материалом для исследования послужили образцы пораженных атеросклерозом сонных артерий, полученные в ходе каротидной эндартерэктомии от восьми пациентов мужского пола (66–71 лет, славяне, жители Западной Сибири). Степень стеноза сонных артерий при ультразвуковом обследовании варьировала от 67 до 90%. Все пациенты курили, в анамнезе у них отмечена артериальная гипертензия и ишемическая болезнь сердца; у двух разных пациентов было острое нарушение мозгового кровообращения (ОНМК) и инфаркт миокарда (ИМ), а еще у одного — сахарный диабет типа 2. Все пациенты принимали лекарственные препараты из групп статинов, антигипертензивных препаратов и антикоагулянтов/дезагрегантов. Биоптаты тканей сонных артерий замораживали и хранили в жидком азоте до процедуры выделения ДНК и РНК.

ДНК выделяли стандартным методом фенол–хлороформной экстракции. ДНК-библиотеки для секвенирования экзозома получали с использованием SureSelect Clinical Research Exome Enrichment Kit (“Agilent”, США). Качество ДНК-библиотек оценивали с помощью прибора Bioanalyzer 2100 (“Agilent”). Секвенирование проводили на приборе HiSeq 1500 (“Illumina”, США) в режиме 2 × 150 п.н. Первичные данные хранятся в репозитории NCBI BioProject, PRJNA758796 [20].

Анализ генетических вариантов проводили с использованием алгоритма Genome Analysis Toolkit (GATK) для поиска унаследованных SNV, коротких инсерций и делеций [21]. Качество прочтений оценивали с помощью программы FastQC v0.11.5 [22]. Выравнивание на сборку генома GRCh38 проводили с помощью GATK BwaSpark, поиск SNV – с использованием GATK Haplotype-Caller [21]. Из дальнейшего анализа исключали инсерции и делеции, а также однонуклеотидные варианты, аннотированные в нескольких локусах генома.

РНК выделяли с использованием реагента TRIzol (“ThermoFisher”, США) и набора PureLink RNA Micro Scale Kit (“ThermoFisher”). Из всех образцов тканей сонных артерий выделена РНК хорошего качества (RINs >7.0), кроме одного образца, который исключили из дальнейшего исследования. Для приготовления библиотеки кДНК использовали набор NEBNext Multiplex Small RNA Library Prep Set for Illumina (“New England Biolabs”, Великобритания). Качество библиотек оценивали с помощью прибора Bioanalyzer 2100 (“Agilent”). Секвенирование осуществляли на приборе HiSeq 1500 (“Illumina”) в режиме 1 × 50 п.н. Данные секвенирования микроРНК анализировали с использованием пакета программ miARma-Seq [23]. Качество прочтений оценивали с помощью программы FastQC v0.11.5. Для удаления адапторов и обрезки последовательностей использовали пакет программ Trim Galore [24]. Выравнивание осуществляли на сборку генома GRCh38 с помощью программного обеспечения BWA-MEM [25]. Количество прочтений микроРНК подсчитывали с использованием алгоритма featureCounts [26]. Уровни экспрессии микроРНК оценивали с использованием пакета “edgeR” в программной среде R. Идентификацию SNV в данных секвенирования микроРНК проводили с помощью пакета программ GATK [21]. Информация о предшественниках (miRNA_primary_transcript) и зрелых микроРНК (miRNA) получена из базы miRBase v22 [27]. Функциональную значимость SNV для экспрессии микроРНК предсказывали с помощью программы miRVaS [28]. Связь микроРНК с атеросклерозом и его осложнениями оценивали с использованием информации, полученной из базы данных HMDD v3.2 [29].

Для подтверждающего исследования и анализа ассоциаций полиморфизма rs2910164 гена *MIR146A* с клинически выраженным атеросклерозом сонных артерий у славян, жителей Западной Сибири, использовали расширенную выборку пациентов ($n = 112$, возраст 65 (59; 69) лет, 78 мужчин и 34 женщины). Степень стеноза сонных артерий при ультразвуковом исследовании (УЗИ) составила $\geq 70\%$, что стало показанием к каротидной эндартерэктомии. У всех пациентов в

анамнезе была артериальная гипертензия. Курили 68 (60.7%) пациентов, а сахарный диабет типа 2 выявлен у 2 (1.8%) пациентов. ОНМК и ИМ зарегистрированы у 38 (33.9%) пациентов, одновременно оба эти события выявлены у 10 (9%) индивидов.

Контрольная группа сформирована из относительно здоровых индивидов без сердечно-сосудистой патологии в анамнезе и признаков ишемии головного мозга при клиническом обследовании ($n = 72$, возраст 69 (67;73) лет, 45 мужчин и 26 женщин). У всех индивидов проведено УЗИ, выявившее начальные стадии атеросклероза сонных артерий, но без гемодинамически значимых изменений (стеноз не превышал 24%).

Генотипирование образцов лейкоцитов периферической крови индивидов проведено с помощью ПЦР с последующим анализом полиморфизма длин рестрикционных фрагментов (ПЦР-ПДРФ), согласно Yamashita J. и соавт. [30]. Распределение генотипов по полиморфному варианту rs2910164 гена *MIR146A* в контрольной группе не выявило отклонений от равновесия Харди–Вайнберга.

Полученные данные анализировали с помощью пакета программ R (The R Foundation). Сравнение количественных данных проводили с использованием критерия Вилкоксона, а качественных – критерия χ^2 Пирсона с поправкой Йетса на непрерывность и/или точного критерия Фишера. Численные значения приведены далее в формате медиана [1-квартиль, 3-квартиль]. Относительный риск у носителей ассоциированного с болезнью аллеля по полиморфному генетическому варианту оценивали по величине отношения шансов (OR), для которой рассчитывали 95%-ный доверительный интервал (95% ДИ). Пороговый уровень значимости принят равным $p = 0.05$.

Проведение работы одобрено Комитетом по биомедицинской этике НИИ медицинской генетики Томского НИМЦ. Все доноры подписали добровольное информированное согласие на участие в исследовании.

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Данные экзомного секвенирования ДНК, выделенной из атеросклеротических бляшек сонных артерий восьми пациентов, характеризовались 55.9 [55.0; 57.5] миллионами парноконцевых прочтений длиной 150 п.н. на образец. В нуклеотидной последовательности предшественников микроРНК найдено 380 SNV, из которых у 321 покрытие ≥ 10 выявлено хотя бы в одном образце (приходящихся на 206 генов микроРНК), из них 234 (72.9%) составили транзиции и 87 (27.1%) – трансверсии. В области нуклеотидной последовательности зрелой микроРНК выявлено 113 SNV, в

том числе с покрытием $\geq \times 10$ — 97 (76 генов), из них 75 (77.3%) транзиций и 22 (22.7%) трансверсий. Эти результаты согласуются с большей частотой транзиций в области генов микроРНК, чем замен других типов [31].

В результате высокопроизводительного секвенирования микроРНК, выделенных из атеросклеротических бляшек сонных артерий семи пациентов, получено в среднем 7.8 [7.2; 12.3] миллиона одноконцевых прочтений длиной 50 п.н., из них 467 [253; 852] тысяч приходилось на микроРНК. Экспрессия на уровне больше 10 CPM (количество прочтений данной микроРНК на миллион прочтений в образце) хотя бы в одном образце зарегистрирована у 364 (19%) из 1917 известных микроРНК, согласно базе miRbase v.22. Причем большую часть микроРНК (169 (61%) из 278), обнаруженных в атеросклеротических бляшках без детализации -5p/-3p, ранее не связывали с атеросклерозом, в том числе сонных артерий и ишемическим инсультом, согласно базе данных HMDD v3.2. Более того, лишь 28 (8%) из 364 микроРНК с детализацией -5p/-3p или 76 (27%) из 278 микроРНК без данного уточнения ранее анализировали в качестве биомаркеров ИБС, острого коронарного синдрома, ИМ, инсульта [32].

Наиболее высокий уровень экспрессии выявлен у miR-143-3p — 18.7 [17.3; 18.9], miR-21-5p — 17.0 [16.4; 17.5] и miR-100-5p — 15.9 [15.4; 16.3]. Экспрессия этих микроРНК в атеросклеротических бляшках сонных артерий суммарно составляет 36% от экспрессии всех микроРНК, причем 23% приходилось на miR-143-3p. Эти результаты согласуются с тем, что экспрессия лишь нескольких микроРНК обуславливает существенную долю суммарного уровня микроРНК в анализируемом образце [33].

С использованием GATK проведен поиск генетических вариантов микроРНК в данных секвенирования. В результате идентифицировано 427 SNV, из которых покрытие $\geq \times 10$ хотя бы в одном образце установлено у 207 (48%), но только три SNV (1.5%) (rs4534339, rs771605638 и rs775920236) располагались в области нуклеотидной последовательности зрелой микроРНК (miR-1843 и miR-100 соответственно). Причем rs4534339 подтвержден и данными экзомного секвенирования — все пациенты имели генотип TT (частота минорного аллеля (MAF) согласно gnomAD у европейцев — 0.99).

При сопоставлении результатов экзомного секвенирования и экспрессии микроРНК выявлено 24 SNV, расположенных в 18 генах микроРНК, которые представляют интерес с точки зрения их влияния на регуляцию экспрессии соответствующих микроРНК в атеросклеротических бляшках сонных артерий (табл. 1). Из этих 24 SNV 14 (58%) были частыми, согласно базе данных gnomAD (MAF gnomAD $\geq 5\%$). Частые

варианты rs71428439 и rs2292832 расположены в области гена *MIR149*, а редкие rs796224492, rs796361602, rs796973621, rs796153424, rs878994369 и rs796405877 — в гене *MIR10527*. Потенциально наибольшей функциональной значимостью для экспрессии микроРНК, предсказанной *in silico*, обладали rs776722712:C>T (*MIR186*), rs2910164:C>G (*MIR146A*), rs2682818:A>C (*MIR618*), rs3746444:A>G (*MIR499A*) и rs199822597:G>A (*MIR363*).

Выявлены пять микроРНК с наиболее высоким уровнем экспрессии в атеросклеротических бляшках сонных артерий: miR-27a-3p — 12.9 [12.0; 13.3], miR-146a-5p — 15.9 [15.4; 16.2], miR-423-5p — 12.1 [10.4; 12.3], miR-423-3p — 11.6 [10.8; 11.7] и miR-186-5p — 10.3 [9.4; 10.4] (табл. 1).

С риском развития ишемического инсульта — осложнением атеросклероза сонных артерий, ранее связывали аллель С/генотип CC rs2292832 (*MIR149*) в азиатских популяциях [34, 35], аллель G/генотип GG rs2910164 (*MIR146A*) у жителей Южной Кореи [12, 13, 36], а в популяции Китая генотипы GT+TT rs2682818 (*MIR618*) [37] и генотип GG rs3746444 (*MIR499A*; табл. 1) [13].

С увеличением риска ИБС связаны аллель T rs2292832 (*MIR149*) в Южной Корее [38], аллель A rs6505162 (*MIR423*) в Индии [39], а в популяциях Азии — генотипы CC rs2910164 (*MIR146A*) [12, 15] и GG rs3746444 (*MIR499A*; табл. 1) [15]. Риск ИМ увеличен у носителей генотипов CC rs895819 (*MIR27A*) [40] и GG rs3746444 (*MIR499A*; табл. 1) [41]. Однако проведенный недавно метаанализ (восемь исследований, 2507 пациентов с ИМ и 3796 здоровых жителей Азии, 11 полиморфизмов генов микроРНК) показал, что, наоборот, риск ИМ повышен у носителей генотипа GG rs2910164 (*MIR146A*), а также AA+AG rs3746444 (*MIR499A*) (табл. 1) [42].

Исследования, проведенные на культурах клеток *in vitro*, показали вовлеченность miR-186-5p, miR-149-5p, miR-146a-5p, miR-182-5p, miR-27a-3p, miR-363-3p в развитие атеросклеротического поражения артерий (табл. 1) [43–50].

На рис. 1 представлен уровень экспрессии 19 микроРНК в атеросклеротических бляшках сонных артерий в зависимости от генотипа пациентов по SNV, расположенных в области генов микроРНК. Однако экспрессия только одной из них — miR-618 — в атеросклеротических бляшках сонных артерий пациентов с генотипом AC rs2682818 гена *MIR618* была в 27 раз ниже, чем у носителей генотипа CC ($\log_2 FC = 4.8$; $p = 0.012$; рис. 1). Полиморфизм rs2682818 может влиять на экспрессию miR-618, препятствуя формированию вторичной структуры шпильки и процессингу предшественника miR-618 до ее зрелой формы. Показано, что аллель A rs2682818 ассоциирован со снижением уровня зрелой miR-618 в клетках HeLa по сравнению с аллелем C [51]. Кроме того,

Таблица 1. Спектр и характеристика SNV в генах микроРНК, экспрессируемых в атеросклеротических бляшках сонных артерий

№	dbSNP:ref/alt	Ген микроРНК	Генотип, кол-во пациентов	Уровень экспрессии микроРНК в атеросклеротических бляшках сонных артерий ($n = 7$)	MAF alt (gnomAD)	Локализация SNV относительно микроРНК	MFE	MFE deltaG ref/alt	Ассоциация с атеросклерозом сонных и коронарных артерий (HMD, NSBI) и его осложнениями
1	rs776722712:C>T	MIR186	CC, 7; CT, 1	miR-186-5p = 10.3 [9.4; 10.4]	0.0002	mature3p/miR-186	mature	-40.6/-36.9	АТ [43, 44]
2	rs4534339:C>T	MIR1843	TT, 8	miR-1843 = 5 [4.8; 5.7]	0.9397	NA	NA	NA	–
3	rs71428439:A>G	MIR149	AA, 7; AG, 1	miR-149-5p = 4.5 [4.1; 6]	0.1333	arm3p/miR-149	no change	-56.1/-56.1	ИМ [42]
4	rs2292832:T>C	MIR149	TC, 5; CC, 3	miR-149-5p = 4.5 [4.1; 6]	0.6889	arm3p/miR-149	no change	-56.1/-58.3	ИИ [34, 35], ИМ [42], АТ [45]
5	rs28645567:G>A	MIR378DI	GG, 7; GA, 1	miR-378d = 3.8 [3.6; 4.4]	0.0766	arm5p/miR-378d-1	no change	-19.6/-17.3	–
6	rs2910164:C>G	MIR146A	CC, 1; CG, 2; GG, 5	miR-146a-5p = 12.1 [11.8; 12.7]	0.7018	mature3p, seed/miR-146a	seed	-38.8/-41.8	ИИ [12, 13, 36], ИБС [12, 15], ИМ [42], АТ [46, 47]
7	rs76481776:C>T	MIR182	CC, 7; CT, 1	miR-182-5p = 5.9 [5.1; 7.8]	0.05786	arm3p/miR-182	arm	-47.3/-47.9	АТ [48]
8	rs7911488:A>G	MIR1307	AA, 4; AG, 4	miR-1307-3p = 8.1 [7.5; 8.4]; miR-1307-5p = 6.3 [6.1; 6.7]	0.2697	loop/miR-1307	no change	-55.4/-55.4	–
9	rs17091403:C>T	MIR2110	CC, 7; CT, 1	miR-2110 = 5.9 [5.4; 6.1]	0.0751	arm5p/miR-2110	mature	-34.6/-35.4	–
10	rs796224492:T>C	MIR10527	TT, 3; TC, 5	miR-10527-5p = 3.2 [3.0; 3.6]	0.0467	NA	NA	NA	–
11	rs796361602:C>T	MIR10527	CC, 3; CT, 5	miR-10527-5p = 3.2 [3.0; 3.6]	0.0547	NA	NA	NA	–
12	rs796973621:C>T	MIR10527	CC, 3; CT, 5	miR-10527-5p = 3.2 [3.0; 3.6]	0.0428	NA	NA	NA	–
13	rs796153424:A>T	MIR10527	AA, 3; AT, 5	miR-10527-5p = 3.2 [3.0; 3.6]	0.0212	NA	NA	NA	–
14	rs878994369:T>C	MIR10527	TT, 4; TC, 4	miR-10527-5p = 3.2 [3.0; 3.6]	0.0139	NA	NA	NA	–

Таблица 1. Окончание

№	dbSNP:ref/alt	Ген микроРНК	Генотип, кол-во пациентов	Уровень экспрессии микроРНК в атеросклеротических бляшках сонных артерий (<i>n</i> = 7)	MAF alt (gnomAD)	Локализация SNV относительно микроРНК	MFE	MFE deltaG ref/alt	Ассоциация с атеросклерозом сонных и коронарных артерий (HMD, NCBI) и его осложнениями
15	rs79640587:G>A	MIR10527	GG, 6; GA, 2	miR-10527-5p = 3.2 [3.0; 3.6]	0.0071	NA	NA	NA	—
16	rs2682818:A>C	MIR618	AA, 1; AC, 2; CC, 5	miR-618 = 2.7 [1.6; 3.1]	0.8077	arm3p/miR-618	seed	−34.5/−38.3	ИИ [37]
17	rs75330474:C>T	MIR323B	CC, 7; CT, 1	miR-323b-3p = 3.4 [2.1; 3.7]	0.0249	mature5p/miR-323b	no change	−40.2/−40.2	—
18	rs41280052:G>T	MIR184	GG, 8	miR-184 = 2.3 [1.6; 4.0]	0.0110	arm5p/miR-184	arm	−37.2/−37	—
19	rs6505162:A>C	MIR423	AA, 3; AC, 3; CC, 2	miR-423-5p = 12.1 [10.4; 12.3]; miR-423-3p = 11.6 [10.8; 11.7];	0.4244	arm3p/miR-423	no change	−48.6/−48.6	ИБС [39]
20	rs745666:G>C	MIR3615	GG, 4; GC, 3; CC, 1	miR-3615 = 6.1 [6.0; 7.1]	0.4086	arm3p/miR-3615	no change	−45.4/−45.4	—
21	rs95819:T>C	MIR27A	TT, 1; TC, 6; CC, 1	miR-27a-3p = 12.9 [12.0; 13.3]; miR-27a-5p = 7.9 [6.4; 8.0]	0.3732	loop/miR-27a	no change	−37.8/−37.8	ИМ [40], АТ [49]
22	rs3746444:A>G	MIR499A	AA, 5; AG, 2; GG, 1	miR-499a-5p = 4.5 [3.9; 4.9]	0.1862	mature5p/miR-499b	seed	−25.4/−21	ИИ [13]; ИБС [15]; ИМ [41, 42]
23	rs199822597:G>A	MIR363	GG, 6; AA, 2	miR-363-3p = 6.5 [5.7; 8.2]	0.0022	mature5p/miR-363	mature	−26.4/−23.7	АТ [50]
24	rs5907732:GT	MIR320D2	GG, 2; TT, 6	miR-320d = 3.4 [2.6; 5.3]	0.7043	upstream/miR-320d-2	NA	NA	—

Примечание: dbSNP – согласно базе данных NCBI; ref/alt – референсный/альтернативный аллель; MAF – частота минорного аллеля, gnomAD – база данных Genome Aggregation Database; MFE – минимальная свободная энергия; ИИ – ишемический инсульт; ИМ – инфаркт миокарда, АТ – атеросклероз; mature – зрелая микроРНК; arm – плечо микроРНК; seed – область “седа” микроРНК; loop – петля микроРНК, upstream – расположены выше по отношению к микроРНК; no change – нет изменений; NA – нет данных. Полуширинным шрифтом в столбце 5 выделены микроРНК, уровень которых выше 10 log₂(CPM); столбец 6 – частые генетические варианты в популяции (частота MAF gnomAD ≥ 5%); столбец 7, 8 – показатели, важные для предсказания *in silico* SNV, обладающие наибольшей потенциальной функциональной значимостью в отношении экспрессии микроРНК.

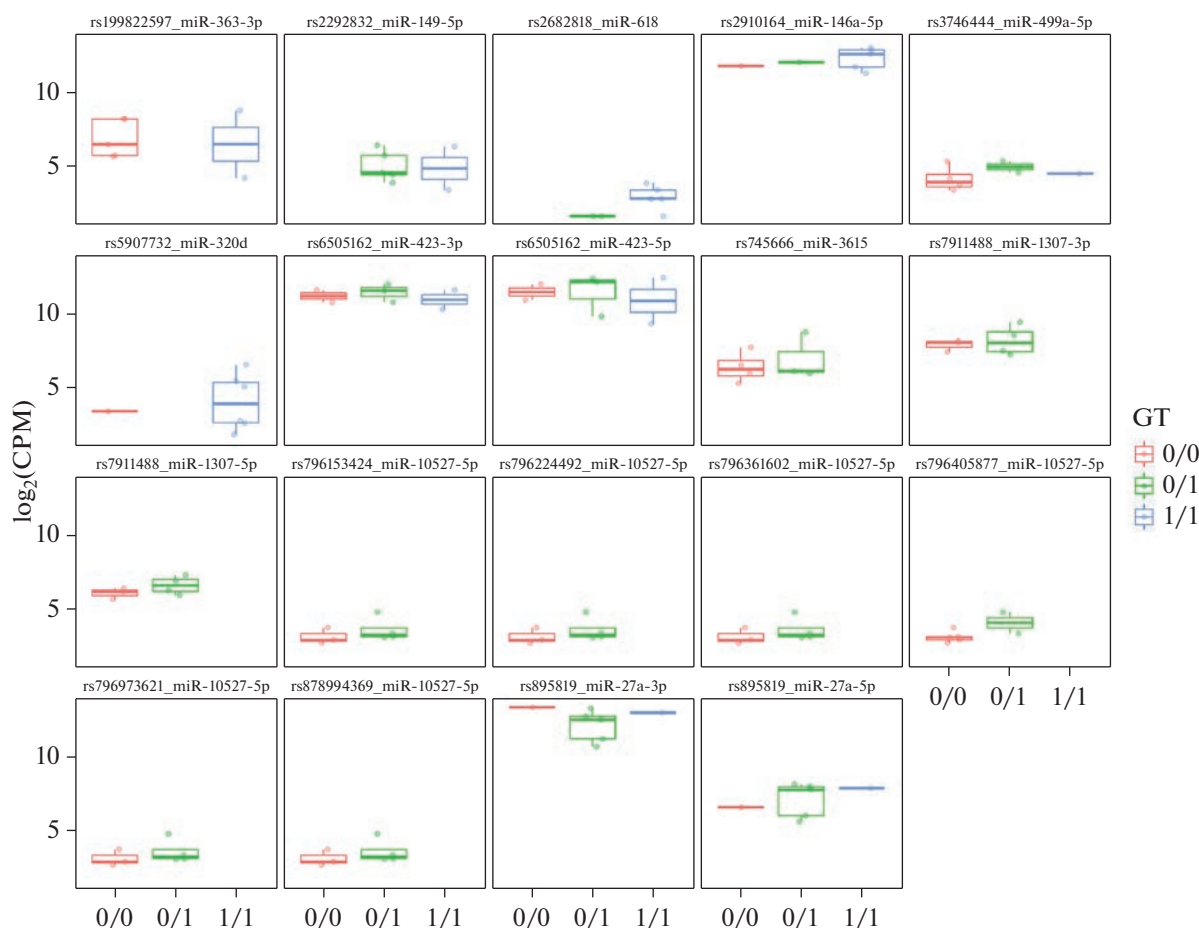


Рис. 1. Уровень экспрессии 19 микроРНК в атеросклеротических бляшках сонных артерий в зависимости от генотипа пациентов по SNV, расположенных в области генов микроРНК. GT – генотипы; 0/0 – гомозиготы по референсному аллелю; 1/1 – гомозиготы по альтернативному аллелю; 0/1 – гетерозиготы. Над каждой ячейкой приведено название SNV и соответствующей микроРНК; по оси X приведены генотипы (0/0, 0/1 и 1/1), а по оси Y – уровень экспрессии микроРНК.

генотипы AC/AA rs2682818 выступают в роли негативного прогностического фактора рецидива ишемического инсульта [37].

Из всех микроРНК, экспрессирующихся в атеросклеротических бляшках сонных артерий, наиболее хорошо изучена miR-146a (табл. 1). Экспрессия miR-146a повышается в атеросклеротических бляшках человека и мыши, при этом она ингибирует передачу сигналов через NF- κ B в эндотелиальных клетках и макрофагах путем интерференции с TRAF6 и IRAK1/2, тем самым снижая экспрессию провоспалительных цитокинов [7, 46, 52–54]. Таким образом, miR-146a представляет собой противовоспалительную микроРНК, которая придает атеропротективные свойства стенке сосуда.

Генетические варианты в пре-микроРНК затрагивают процессинг отдельных микроРНК и, соответственно, снижают уровень зрелой микроРНК. В частности, наиболее активно при ате-

росклерозе и его осложнениях изучали вариант rs2910164:C>G гена *MIR146A*. Однако ассоциации этого полиморфизма с фенотипами атеросклероза сонных и коронарных артерий анализировали, главным образом, в популяциях Азии. Лишь в одной работе показана связь генотипа CC rs2910164 гена *MIR146A* с риском рестеноза коронарных артерий у европейцев в популяции Германии [55]. Согласно данным gnomAD, частоты минорного аллеля C rs2910164 различаются в популяциях Восточной Азии и Европы – 63 и 23% соответственно.

В связи с этим, в рамках настоящего исследования проведено генотипирование лейкоцитов крови пациентов с осложненным течением атеросклероза сонных артерий и у относительно здоровых жителей Западной Сибири, этнических славян, по полиморфному варианту rs2910164 гена *MIR146A* (табл. 2). Сравнение частот аллелей и генотипов данного полиморфизма в этих двух группах выявило ассоциацию аллеля C и генотипов

Таблица 2. Распределение аллелей и генотипов rs2910164 (G>C) гена *MIR146A* у пациентов с осложненным течением атеросклероза сонных артерий и у относительно здоровых жителей Западной Сибири славянского происхождения

Генотип/аллель	Здоровые индивиды, абс. кол-во (%)	Пациенты с клинически выраженным атеросклерозом, абс. кол-во (%)	χ^2 ; OR (95%CI); p
GG	46 (63.9)	42 (37.5)	$\chi^2 = 11.20$; OR = 2.95 (95% CI: 1.60–5.45); $p = 0.001$
GC + CC	26 (36.1)	70 (62.5)	
G	116 (80.5)	143 (63.8)	$\chi^2 = 10.96$ OR = 2.35 (95% CI: 1.43–3.85) $p = 0.001$
C	28 (19.5)	81 (36.2)	

(CC+GC) rs2910164 гена *MIR146A* с риском клинически выраженного атеросклероза сонных артерий: OR = 2.35 (95%CI: 1.43–3.85), $p = 0.001$, OR = 2.95 (95% CI: 1.60–5.45) $p = 0.001$.

Следует отметить, что в популяциях Азии генотип CC rs2910164 (*MIR146A*) связан с риском ИБС, но аллель G/генотип GG ассоциирован с риском осложненного течения атеросклероза сонных и коронарных артерий – ОНМК и ИМ [12, 13, 36]. В нашей работе установлена ассоциация аллеля C и генотипов (CC+GC) rs2910164 гена *MIR146A* с риском развития клинически выраженного атеросклероза сонных артерий жителей Западной Сибири славянского происхождения, однако связь с ОНМК остается неясной, поскольку острые сосудистые события зарегистрированы у 38 человек из общей выборки, что недостаточно для анализа.

Полиморфизм rs2910164 приводит к замене нуклеотидов G>C в последовательности премикроРНК гена *MIR146A* и далее в “пассажирской цепи” зрелой miR-146a-3p. В результате нарушается процессинг pre-miR-146a, конформация ее вторичной структуры и стабильность, уменьшается продукция зрелой miR-146a в клеточной линии U2OS [56]. В этой работе показано, что генотип CC и аллель C rs2910164 связаны со снижением экспрессии miR-146a в клетках, что увеличивает содержание мРНК ее мишеней (IRAK1, TRAF6), внося вклад в формирование провоспалительного профиля и, следовательно, риска заболевания.

Однако в мононуклеарных клетках периферической крови пациентов с ИБС и генотипом CC rs2910164 гена *MIR146A* выявлено повышение уровня miR-146a и снижение уровня IRAK-1, TRAF-6, NF- κ B и C-реактивного белка по сравнению с носителями генотипа GG [52]. Кроме того, в работе Xiong и соавт. показана ассоциация генотипов GC/CC rs2910164 с риском развития ИБС и увеличением экспрессии miR-146a в мо-

нонуклеарных клетках крови пациентов [57]. В то же время при сахарном диабете типа 2, который часто регистрируется у пациентов с атеросклеротическим поражением артерий, снижение уровня miR-146a в мононуклеарных клетках крови связано с генотипом CC rs2910164 [58]. Однако нами не выявлено связи между экспрессией miR-146a-5p в атеросклеротических бляшках сонных артерий и генотипом rs2910164 гена *MIR146A*. Возможно, существуют клеточно- и тканеспецифичные различия в экспрессии miR-146a у носителей разных генотипов rs2910164 гена *MIR146A* в зависимости от стадии и русла атеросклеротического поражения артерий, коморбидного статуса пациентов.

С другой стороны, относительно малый размер выборок, используемых для анализа связи между полиморфизмом гена и экспрессией микроРНК в атеросклеротических бляшках сонных артерий, может быть причиной различий в результатах, полученных в нашей и других работах. Кроме того, результаты изучения ассоциации rs2910164 гена *MIR146A* с патологией необходимо реплицировать на больших по размеру и разных по этническому составу выборках, в том числе, включающих пациентов разного пола, с атеросклеротическим поражением артерий разной локализации и их осложнениями. Не исключено, что кумулятивный эффект не одного, а многих генетических вариантов, в том числе локализованных в области генов микроРНК, и факторов среды играет существенную роль в формировании предрасположенности к атеросклеротическому поражению артерий.

Результаты анализа ассоциаций полиморфизма генов микроРНК с атеросклерозом необходимо интерпретировать в контексте оценки экспрессии пре- и зрелой микроРНК и их мишеней в клетках/тканях органов-мишеней патологии. С целью установления причинно-следственных отношений и детальных молекулярно-генетических механизмов эти работы следует дополнять функциональными исследованиями с использо-

ванием культур клеток и модельных животных. Так, показано, что miR-146a вносит вклад в формирование и стабильность атеросклеротических бляшек, регулируя пролиферацию гладкомышечных клеток через сигнальный путь Notch и процесс воспаления в макрофагах IRAK1/TRAF6/NF- κ B [59, 60]. Однако функциональная значимость miR-618 при атеросклерозе остается неясной.

ЗАКЛЮЧЕНИЕ

В настоящем исследовании впервые охарактеризован спектр SNV в генах микроРНК у проживающих в Западной Сибири этнических славян, с клинически выраженным атеросклерозом сонных артерий. С помощью секвенирования экзонов и микроРНК оценена экспрессия зрелых микроРНК в атеросклеротических бляшках сонных артерий. Выявлены 24 SNV, расположенных в 18 генах микроРНК, которые экспрессируются в атеросклеротических бляшках сонных артерий. Показано снижение экспрессии miR-618 в атеросклеротических бляшках сонных артерий у пациентов с генотипом AC rs2682818 гена *MIR618* по сравнению с генотипом CC ($\log_2FC = 4.8$; $p = 0.012$). Установлена ассоциация аллеля C и генотипов (GC+CC) полиморфного варианта rs2910164 гена *MIR146A* с риском развития клинически выраженного атеросклероза сонных артерий у славян, проживающих в Западной Сибири (OR = 2.35; 95%CI: 1.43–3.85; $p = 0.001$, OR = 2.95; 95% CI: 1.60–5.45; $p = 0.001$ соответственно).

Работа выполнена в рамках Государственного задания Министерства науки и высшего образования Российской Федерации.

Проведение данной работы одобрено Комитетом по биомедицинской этике НИИ медицинской генетики Томского НИМЦ. Все доноры подписали добровольное информированное согласие на участие в исследовании.

Авторы заявляют об отсутствии конфликта интересов.

СПИСОК ЛИТЕРАТУРЫ

1. Song P., Fang Z., Wang H., Cai Y., Rahimi K., Zhu Y., Fowkes F., Fowkes F., Rudan I. (2020) Global and regional prevalence, burden, and risk factors for carotid atherosclerosis: a systematic review, meta-analysis, and modelling study. *Lancet Glob. Hlth.* **8**(5), e721–e729.
2. Libby P., Bornfeldt K.E., Tall A.R. (2016) Atherosclerosis: successes, surprises, and future challenges. *Circ. Res.* **118**(4), 531–534.
3. Gebert L.F.R., MacRae I.J. (2019) Regulation of microRNA function in animals. *Nat. Rev. Mol. Cell. Biol.* **20**(1), 21–37.
4. Borghini A., Andreassi M.G. (2018) Genetic polymorphisms offer insight into the causal role of microRNA in coronary artery disease. *Atherosclerosis.* **269**, 63–70.
5. Króliczewski J., Sobolewska A., Lejnowski D., Collawn J.F., Bartoszewski R. (2018) microRNA single polynucleotide polymorphism influences on microRNA biogenesis and mRNA target specificity. *Gene.* **640**, 66–72.
6. Li Y., Huo C., Pan T., Li L., Jin X., Lin X., Chen J., Zhang J., Guo Z., Xu J., Li X. (2019) Systematic review regulatory principles of non-coding RNAs in cardiovascular diseases. *Brief. Bioinform.* **20**(1), 66–76.
7. Feinberg M.W., Moore K.J. (2016) MicroRNA regulation of atherosclerosis. *Circ. Res.* **118**(4), 703–720.
8. Ghanbari M., Franco O.H., de Looper H.W., Hofman A., Erkeland S.J., Dehghan A. (2015) Genetic variations in microRNA-binding sites affect microRNA-mediated regulation of several genes associated with cardio-metabolic phenotypes. *Circ. Cardiovasc. Genet.* **8**(3), 473–486.
9. Fasolo F., Di Gregoli K., Maegdefessel L., Johnson J.L. (2019) Non-coding RNAs in cardiovascular cell biology and atherosclerosis. *Cardiovasc. Res.* **115**(12), 1732–1756.
10. Huang Z., Shi J., Gao Y., Cui C., Zhang S., Li J., Zhou Y., Cui Q. (2019) HMDD v3.0: a database for experimentally supported human microRNA-disease associations. *Nucl. Acids Res.* **47**(D1), D1013–D1017.
11. He Y., Yang J., Kong D., Lin, J., Xu C., Ren H., Ouyang P., Ding Y., Wang K. (2015) Association of miR-146a rs2910164 polymorphism with cardio-cerebrovascular diseases: a systematic review and meta-analysis. *Gene.* **565**(2), 171–179.
12. Bao M.H., Xiao Y., Zhang Q.S., Luo H.Q., Luo J., Zhao J., Li G.Y., Zeng J., Li J.M. (2015) Meta-analysis of miR-146a polymorphisms association with coronary artery diseases and ischemic stroke. *Int. J. Mol. Sci.* **16**(7), 14305–14317.
13. Zhu J., Yue H., Qiao C., Li Y. (2015) Association between single-nucleotide polymorphism (SNP) in miR-146a, miR-196a2, and miR-499 and risk of ischemic stroke: a meta-analysis. *Med. Sci. Monit.* **21**, 3658–3663.
14. Zhao D., Li Y., Yu X., Zhu Y., Ma B. (2019) Associations between miR-146a rs2910164 polymorphisms and risk of ischemic cardio-cerebrovascular diseases. *Medicine.* **98**(42), e17106.
15. Bastami M., Choupani J., Saadatani Z., Zununi Vahed S., Mansoori Y., Daraei A., Samadi Kafil H., Masotti A., Nariman-Saleh-Fam Z. (2019) miRNA polymorphisms and risk of cardio-cerebrovascular diseases: a systematic review and meta-analysis. *Int. J. Mol. Sci.* **20**(2), 293.
16. Elfaki I., Mir R., Mir M.M., AbuDuhier F.M., Babakr A.T., Barnawi J. (2019) Potential impact of microRNA gene polymorphisms in the pathogenesis of diabetes and atherosclerotic cardiovascular disease. *J. Personalized Med.* **9**(4), 51.
17. Raitoharju E., Oksala N., Lehtimäki T. (2013) MicroRNAs in the atherosclerotic plaque. *Clin. Chem.* **59**(12), 1708–1721.
18. Maitrias P., Metzinger-Le Meuth V., Nader J., Reix T., Caus T., Metzinger L. (2017) The involvement of miRNA in carotid-related stroke. *Arterioscler. Thromb. Vasc. Biol.* **37**(9), 1608–1617.

19. Назаренко М.С., Королёва Ю.А., Зарубин А.А., Слепцов А.А. (2022) Регулом микроРНК при различных фенотипах атеросклероза. *Молекуляр. биология*. **56**(2), 227–243.
20. Sleptcov A.A., Zarubin A.A., Bogaychuk P.M., Kuznetsov M.S., Kozlov B.N., Nazarenko M.S. (2021) Human exome sequence data in support of somatic mosaicism in carotid atherosclerosis. *Data Brief*. **39**, 107656.
21. Van der Auwera G.A., O'Connor B.D. (2020) *Genomics in the Cloud: Using Docker, GATK, and WDL in Terra*. O'Reilly Media, Inc. p. 496.
22. Andrews S. (2010) FastQC: a quality control tool for high throughput sequence data. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.
23. Andrés-León E., Núñez-Torres R., Rojas A.M. (2016) miARma-Seq: a comprehensive tool for miRNA, mRNA and circRNA analysis. *Sci. Rep.* **6**, 25749.
24. Krueger F., James F., Ewels P., Afyounian E., Schuster-Boeckler B. (2021) FelixKrueger/TrimGalore: v0.6.7 – DOI via Zenodo. Zenodo.
25. Li H. (2013) Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv Preprint arXiv*. 1303.3997.
26. Liao Y., Smyth G.K., Shi W. (2014) featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* (Oxford). **30**(7), 923–930.
27. Kozomara A., Birgaoanu M., Griffiths-Jones S. (2019) miRBase: from microRNA sequences to function. *Nucl. Acids Res.* **47**(D1), D155–D162.
28. Cammaerts S., Strazisar M., Dierckx J., Del Favero J., De Rijk P. (2016) miRVaS: a tool to predict the impact of genetic variants on miRNAs. *Nucl. Acids Res.* **44**(3), e23.
29. Huang Z., Shi J., Gao Y., Cui C., Zhang S., Li J., Zhou Z., Cui Q. (2019) HMDD v3. 0: a database for experimentally supported human microRNA–disease associations. *Nucl. Acids Res.* **47**(D1), D1013–D1017.
30. Yamashita J., Iwakiri T., Fukushima S., Jinnin M., Miyashita A., Hamasaki T., Makino T., Aoi J., Masuguchi S., Inoue Y., Ihn H. (2013) The rs2910164 G > C polymorphism in microRNA-146a is associated with the incidence of malignant melanoma. *Melanoma Res.* **23**(1), 13–20.
31. Omariba G., Xu F., Wang M., Li K., Zhou Y., Xiao J. (2020) Genome-wide analysis of microRNA-related single nucleotide polymorphisms (SNPs) in mouse genome. *Sci. Rep.* **10**(1), 5789.
32. Sharma A.R., Shashikiran U., Uk A.R., Shetty R., Sathyamoorthy K., Rai P.S. (2020) Aberrant DNA methylation and miRNAs in coronary artery diseases and stroke: a systematic review. *Brief. Funct. Genomics*. **19**(4), 259–285.
33. De Rie D., Abugessaisa I., Alam T., Arner E., Arner P., Ashoor H., Åström G., Babina M., Bertin N., Burroughs A.M., Carlisle A.J., Daub C.O., Detmar M., Deviatiiarov R., Fort A., Gebhard C., Goldowitz D., Guhl S., Ha T.J., Harshbarger J., Hasegawa A., Hashimoto K., Herlyn M., Heutink P., Hitchens K.J., Hon C.C., Huang E., Ishizu Y., Kai C., Kasukawa T., Klinken P., Lassmann T., Lecellier C.H., Lee W., Lizio M., Makeev V., Mathelier A., Medvedeva Y.A., Mejhert N., Mungall C.J., Noma S., Ohshima M., Okada-Hatakeyama M., Persson H., Rizzu P., Roudnický F., Sætrom P., Sato H., Severin J., Shin J.W., Swoboda R.K., Tarui H., Toyoda H., Vitting-Seerup K., Winteringham L., Yamaguchi Y., Yasuzawa K., Yoneda M., Yumoto N., Zabierowski S., Zhang P.G., Wells C.A., Summers K.M., Kawaji H., Sandelin A., Rehli M.; FANTOM Consortium, Hayashizaki Y., Carninci P., Forrest A.R.R., de Hoon M.J.L. (2017) An integrated expression atlas of miRNAs and their promoters in human and mouse. *Nat. Biotechnol.* **35**(9), 872–878.
34. Zou D., Liu C., Zhang Q., Li X., Qin G., Huang Q., Meng Y., Chen L., Wei J. (2018) Association between polymorphisms in microRNAs and ischemic stroke in an Asian population: evidence based on 6,083 cases and 7,248 controls. *Clin. Interv. Aging*. **13**, 1709–1726.
35. Zhu R., Wang Q.W., Zhao J., Liu X., He Z. (2020) miR-149 and miR-499 gene polymorphism and the incident of ischemic stroke in the Asian population: from a case-control study to meta-analysis. *Clin. Neurol. Neurosurg.* **193**, 105789.
36. Wang M., Wang S., Wang X., Wu J., Wu Y., Wang Z., Wang J., Wu T., Hu Y. (2020) Carotid intima-media thickness, genetic risk, and ischemic stroke: a family-based study in Rural China. *Int. J. Environ. Res. Publ. Hlth.* **18**(1), 119.
37. Zhang Z., Xu G., Cai B., Zhang H., Zhu W., Liu X. (2017) Genetic variants in microRNAs predict recurrence of ischemic stroke. *Mol. Neurobiol.* **54**(4), 2776–2780.
38. Sung J.H., Kim S.H., Yang W.I., Kim W.J., Moon J.Y., Kim I.J., Cha D.H., Cho S.Y., Kim J.O., Kim K.A., Kim O.J., Lim S.W., Kim N.K. (2016) miRNA polymorphisms (miR-146a, miR-149, miR-196a2 and miR-499) are associated with the risk of coronary artery disease. *Mol. Med. Rep.* **14**(3), 2328–2342.
39. Jha C.K., Mir R., Elfaki I., Khullar N., Rehman S., Javid J., Banu S., Chahal S. (2019) Potential impact of microRNA-423 gene variability in coronary artery disease. *Endocr. Metab. Immune Disord. Drug Targets.* **19**(1), 67–74.
40. Cai M.Y., Cheng J., Zhou M.Y., Liang L.L., Lian S.M., Xie X.S., Xu S., Liu X., Xiong X.D. (2018) The association between pre-miR-27a rs895819 polymorphism and myocardial infarction risk in a Chinese Han population. *Lipids Hlth Dis.* **17**(1), 7.
41. Chen C., Hong H., Chen L., Shi X., Chen Y., Weng Q. (2014) Association of microRNA polymorphisms with the risk of myocardial infarction in a Chinese population. *Tohoku J. Exp. Med.* **233**(2), 89–94.
42. Yang Y., Shi X., Du Z., Zhou G., Zhang X. (2021) Associations between genetic variations in microRNA and myocardial infarction susceptibility: a meta-analysis and systematic review. *Herz.* **47**(6), 524–535. <https://doi.org/10.1007/s00059-021-05086-3>
43. Sun B., Cao Q., Meng M., Wang X. (2020) MicroRNA-186-5p serves as a diagnostic biomarker in atherosclerosis and regulates vascular smooth muscle cell proliferation and migration. *Cell Mol. Biol. Lett.* **25**, 27.
44. Zhang S., Zhu X., Li G. (2020) E2F1/SNHG7/miR-186-5p/MMP2 axis modulates the proliferation and

- migration of vascular endothelial cell in atherosclerosis. *Life Sci.* **257**, 118013.
45. Ye Z.M., Yang S., Xia Y.P., Hu R.T., Chen S., Li B.W., Chen S.L., Luo X.Y., Mao L., Li Y., Jin H., Qin C., Hu B. (2019) LncRNA MIAT sponges miR-149-5p to inhibit efferocytosis in advanced atherosclerosis through CD47 upregulation. *Cell Death Dis.* **10**(2), 138.
 46. Hou J., Deng Q., Deng X., Zhong W., Liu S., Zhong Z. (2021) MicroRNA-146a-5p alleviates lipopolysaccharide-induced NLRP3 inflammasome injury and pro-inflammatory cytokine production via the regulation of TRAF6 and IRAK1 in human umbilical vein endothelial cells (HUVECs). *Ann. Transl. Med.* **9**(18), 1433.
 47. Huang S.F., Zhao G., Peng X.F., Ye W.C. (2021) The pathogenic role of long non-coding RNA H19 in atherosclerosis via the miR-146a-5p/ANGPTL4 pathway. *Front. Cardiovasc. Med.* **8**, 770163.
 48. Qin S.B., Peng D.Y., Lu J.M., Ke Z.P. (2018) MiR-182-5p inhibited oxidative stress and apoptosis triggered by oxidized low-density lipoprotein via targeting toll-like receptor 4. *J. Cell Physiol.* **233**(10), 6630–6637.
 49. Choe N., Kwon D.H., Ryu J., Shin S., Cho H.J., Joung H., Eom G.H., Ahn Y., Park W.J., Nam K.I., Kim Y.K., Kook H. (2020) miR-27a-3p targets ATF3 to reduce calcium deposition in vascular smooth muscle cells. *Mol. Ther. Nucl. Acids.* **22**, 627–639.
 50. Zhou T., Li S., Yang L., Xiang D. (2021) microRNA-363-3p reduces endothelial cell inflammatory responses in coronary heart disease via inactivation of the NOX4-dependent p38 MAPK axis. *Aging.* **13**(8), 11061–11082.
 51. Fu A., Hoffman A.E., Liu R., Jacobs D.I., Zheng T., Zhu Y. (2014) Targetome profiling and functional genetics implicate miR-618 in lymphomagenesis. *Epi-genetics.* **9**(5), 730–737.
 52. Ramkaran P., Khan S., Phulukdaree A., Moodley D., Chuturgoon A.A. (2014) miR-146a polymorphism influences levels of miR-146a, IRAK-1, and TRAF-6 in young patients with coronary artery disease. *Cell. Biochem. Biophys.* **68**(2), 259–266.
 53. Raitoharju E., Lyytikäinen L.P., Levula M., Oksala N., Mennander A., Tarkka M., Klopp N., Illig T., Kähönen M., Karhunen P.J., Laaksonen R., Lehtimäki T. (2011) miR-21, miR-210, miR-34a, and miR-146a/b are up-regulated in human atherosclerotic plaques in the Tampere Vascular Study. *Atherosclerosis.* **219**(1), 211–217.
 54. Takahashi Y., Satoh M., Minami Y., Tabuchi T., Itoh T., Nakamura M. (2010) Expression of miR-146a is associated with the Toll-like receptor 4 signal in coronary artery disease: effect of renin-angiotensin system blockade and statins on miRNA-146a/b and Toll-like receptor 4 levels. *Clin. Sci. (Lond).* **119**(9), 395–405.
 55. Hamann L., Glaeser C., Schulz S., Gross M., Franke A., Nöthlings U., Schumann R.R. (2014) A micro RNA-146a polymorphism is associated with coronary restenosis. *Int. J. Immunogenet.* **41**(5), 393–396.
 56. Jazdzewski K., Murray E.L., Franssila K., Jarzab B., Schoenberg D.R., de la Chapelle A. (2008) Common SNP in pre-miR-146a decreases mature miR expression and predisposes to papillary thyroid carcinoma. *Proc. Natl. Acad. Sci. USA.* **105**(20), 7269–7274.
 57. Xiong X.D., Cho M., Cai X.P., Cheng J., Jing X., Cen J.M., Liu X., Yang X.L., Suh Y. (2014) A common variant in pre-miR-146 is associated with coronary artery disease risk and its mature miRNA expression. *Mutat. Res.* **761**, 15–20.
 58. Alipoor B., Ghaedi H., Meshkani R., Omrani M.D., Sharifi Z., Golmohammadi T. (2018) The rs2910164 variant is associated with reduced miR-146a expression but not cytokine levels in patients with type 2 diabetes. *J. Endocrinol. Invest.* **41**(5), 557–566.
 59. Cao J., Zhang K., Zheng J., Dong R. (2015) MicroRNA-146a and -21 cooperate to regulate vascular smooth muscle cell proliferation via modulation of the Notch signaling pathway. *Mol. Med. Rep.* **11**(4), 2889–2895.
 60. Chu T., Xu X., Ruan Z., Wu L., Zhou M., Zhu G. (2022) miR-146a contributes to atherosclerotic plaque stability by regulating the expression of TRAF6 and IRAK-1. *Mol. Biol. Rep.* **49**(6), 4205–4216. <https://doi.org/10.1007/s11033-022-07253-z>

Identification of Functionally Significant Polymorphic Variants in miRNA Genes in Carotid Atherosclerosis

A. A. Zarubin¹, K. V. Mannanova², I. A. Koroleva¹, A. A. Sleptcov¹,
M. S. Kuznetsov³, B. N. Kozlov³, and M. S. Nazarenko^{1,2,*}

¹Research Institute of Medical Genetics, Tomsk National Research Medical Center, Russian Academy of Sciences, Tomsk, 634050 Russia

²Siberian State Medical University, Tomsk, 634050 Russia

³Cardiology Research Institute, Tomsk National Research Medical Center, Russian Academy of Sciences, Tomsk, 634050 Russia

*e-mail: maria.nazarenko@medgenetics.ru

miRNAs are vital molecules of gene expression. They are involved in the pathogenesis of various common diseases, including atherosclerosis, its risk factors and complications. A detailed characterization of the spectrum of functionally significant polymorphisms of miRNA genes of patients with advanced carotid atherosclerosis is an actual research task. We analyzed miRNA expression and exome sequencing data of carotid atherosclerotic plaques of the same male patients ($n = 8$, 66–71 years of age, 67–90% degree of carotid artery stenosis). For further study and analysis of the association between rs2910164 polymorphism of the *MIR146A*

gene and advanced carotid atherosclerosis, we recruited 112 patients and 72 relatively healthy Slavic residents of Western Siberia. 321 and 97 single nucleotide variants (SNVs) were detected in the nucleotide sequences of pre- and mature miRNAs in carotid atherosclerotic plaques. These variants were located in 206 and 76 miRNA genes, respectively. Integration the data of exome sequencing and miRNA expression revealed 24 SNVs of 18 miRNA genes which were processed to mature form in carotid atherosclerotic plaques. SNVs with the greatest potential functional significance for miRNA expression predicted *in silico* were rs2910164:C>G (*MIR146A*), rs2682818:A>C (*MIR618*), rs3746444:A>G (*MIR499A*), rs776722712:C>T (*MIR186*), rs199822597:G>A (*MIR363*). The expression of miR-618 was lower in carotid atherosclerotic plaques of patients with the AC rs2682818 genotype of the *MIR618* gene compared with the CC genotype ($\log_2FC = 4.8$; $p = 0.012$). We also found the association of rs2910164:C (*MIR146A*) with the risk of advanced carotid atherosclerosis (OR = 2.35; 95% CI: 1.43–3.85; $p = 0.001$). Integrative analysis of polymorphism in miRNA genes and miRNA expression is informative for identifying functionally significant polymorphisms in miRNA genes. The rs2682818:A>C (*MIR618*) is a candidate for regulating miRNA expression in carotid atherosclerotic plaques. The rs2910164:C (*MIR146A*) is associated with the risk of advanced carotid atherosclerosis.

Keywords: miRNA, sequencing, carotid atherosclerosis

ГЕНОМИКА.
ТРАНСКРИПТОМИКА

УДК 575.85

ВОВЛЕЧЕННОСТЬ ГЕНА *Gagr*, ДОМЕСТИЦИРОВАННОГО ГЕНА *gag*
РЕТРОВИРУСОВ, В ПУТИ СТРЕССОВОГО ОТВЕТА
У РАЗНЫХ ВИДОВ ДРОЗОФИЛЫ

© 2023 г. А. Н. Гигин^а, Л. Н. Нефедова^{а, *}

^аБиологический факультет Московского государственного университета им. М.В. Ломоносова,
Москва, 119234 Россия

*e-mail: nefedova@mail.bio.msu.ru

Поступила в редакцию 21.11.2022 г.

После доработки 02.12.2022 г.

Принята к публикации 05.12.2022 г.

Ген *Gagr* — доместицированный ген *gag* ретроэлементов у *Drosophila melanogaster*, функция которого связана со стрессовым ответом. Ген *Gagr* и его гомологи у разных видов *Drosophila* имеют высоко консервативную структуру, однако демонстрируют определенную изменчивость промоторной области, связанную, по-видимому, с постепенным приобретением доместицированным геном новой функции и вовлечением в новые сигнальные пути. В настоящей работе изучено влияние окислительного стресса, вызванного персульфатом аммония, на жизнеспособность различных видов рода *Drosophila* (*D. melanogaster*, *D. mauritiana*, *D. simulans*, *D. yakuba*, *D. teissieri*, *D. pseudoobscura*), проведен анализ корреляции структуры промоторных областей и стресс-индуцированных изменений в экспрессии гена *Gagr* и его гомологов у разных видов дрозофил, а также сравнение стресс-индуцированных изменений в экспрессии генов-маркеров окислительного стресса: активатора сигнального пути Jak-STAT *upd3*, эффектора пути Jak-STAT *vir-1* и мишени сигнального пути IMD Rel. Обнаружено, что у *D. simulans* и *D. mauritiana* существенно повышена чувствительность к персульфату аммония, что коррелирует со сниженным уровнем транскрипции ортологов гена *vir-1*. Последнее обусловлено уменьшением числа сайтов связывания транскрипционного фактора STAT92E, компонента сигнального пути Jak-STAT, в промоторной области ортологов *vir-1* у *D. simulans* и *D. mauritiana*. Согласованные изменения экспрессии генов *Gagr*, *upd3*, *vir-1* наблюдаются у всех видов подгруппы *melanogaster*, кроме *D. pseudoobscura*, что свидетельствует о возрастании роли *Gagr* в регуляции путей стрессового ответа в ходе филогенеза рода *Drosophila*.

Ключевые слова: доместикация, ретроэлементы, *gag*, *Drosophila*, стрессовый ответ

DOI: 10.31857/S0026898423030047, EDN: CHBVPQ

ВВЕДЕНИЕ

Молекулярная доместикация последовательностей мобильных элементов может играть огромную роль в образовании новых семейств генов, несущих в себе значимые эволюционные новшества и во многом определяющие облик крупных таксонов эукариот. Роль доместицированных генов ретроэлементов у млекопитающих изучена достаточно подробно: обнаружено несколько семейств генов ретротранспозонного происхождения (*PNMA*, *MART*, синцитины и *Vex/Tcl*), члены которых играют определяющую роль в формировании плаценты и в эмбриогенезе [1–5], а также в функционировании центральной нервной системы [6, 7]. В защите от ретровирусов, попадающих в организм извне, могут участвовать и другие доместицированные гены. Так, у мышей есть четыре таких гена — *Fv1*, *Fv4*, *Rmcf1* и *Rmcf2* [8–10]. Доместицированные последовательности ретроэлементов найдены и у других по-

звоночных: ген *Arc*, необходимый для формирования и поддержания долговременной памяти в центральной нервной системе, консервативен у тетрапод [11], а последовательности, кодирующие SCAN-домен, характерны для многих транскрипционных факторов (ТФ) у амниот — самой разнообразной и успешной группы тетрапод [12]. Функции многих доместицированных последовательностей ретроэлементов изучены слабо; некоторые из них охарактеризованы недавно: например, роль последовательности выше гена *Pparg*, произошедшей от ретроэлементов *LINE*, как энхансера этого гена, необходимого для адипогенеза [13], или роль *PRLH1* — транскрипта, происходящего от эндогенного ретровируса ERV-9, в репарации двухцепочечных разрывов [14].

Случаи молекулярной доместикации ретроэлементов и их последовательностей у беспозвоночных изучены значительно слабее, чем у позвоночных. Большая часть исследований ретроэле-

ментов проведена на плодовых мухах (*Drosophila melanogaster*), для которых характерно высокое разнообразие ретроэлементов (включая самую представительную группу — *Gypsy*, очень близкую по свойствам к ретровирусам) [15]. На текущий момент у *D. melanogaster* известно лишь четыре доместицированных гена, произошедших от ретроэлементов, — *Gagr*, *dARC1*, *dARC2* (гомологи гена *gag* ретроэлементов группы *Gypsy*) и *Iris* (гомолог *env* ретроэлементов группы *Bel*) [11, 16, 17]. Ген *Gagr* и его белковый продукт проявляют определенную изменчивость среди видов рода *Drosophila*, связанную с постепенным приобретением новой функции и вовлечением в новые сигнальные пути [18]. Установление связи между изменчивостью последовательности гена *Gagr* в ряду видов дрозофил с видоизменением его функций позволит лучше понять ход процесса молекулярной доместикации в контексте эволюции.

Ранее было показано, что экспрессия гена *Gagr* у *D. melanogaster* индуцируется индуктором окислительного стресса — персульфатом аммония; при этом также индуцируется экспрессия ряда других стресс-активируемых генов, вероятно, имеющих общие пути регуляции с доместицированным геном: *vir-1*, *upd3* и *Rel* [18]. Ген *vir-1* — ген с неизвестными функциями, экспрессия которого находится под контролем каскада Jak-STAT [19]. У *D. melanogaster* экспрессия *vir-1* повышается в ответ на окислительный стресс [20], а также на вирусную инфекцию [19]. Ген *upd3* — один из трех генов семейства —цитокинов Unpaired, запускающих сигнальный путь Jak-STAT через рецептор Domeless [21], является мишенью ТФ Sd/Yki (из пути Hippo) и AP-1 [22, 23]. Экспрессия *upd3* у *D. melanogaster* индуцируется в ответ на окислительный стресс и через каскад JNK, активирующий как Yki, так и AP-1 [23, 24]. Ген *Rel* кодирует мишень сигнального пути IMD, ТФ, который активируется в ответ на присутствие патогенных микроорганизмов и некоторых окислителей [25, 26].

В задачи нашей работы входило изучение влияния окислительного стресса, вызванного персульфатом аммония, на жизнеспособность различных видов рода *Drosophila* (*D. melanogaster*, *D. mauritiana*, *D. simulans*, *D. yakuba*, *D. teissieri*, *D. pseudoobscura*), анализ корреляции структуры промоторных областей и стресс-индуцированных изменений экспрессии гена *Gagr* и его гомологов у разных видов дрозофил, а также сравнение стресс-индуцированных изменений экспрессии генов-маркеров окислительного стресса (*vir-1*, *Rel* и *upd3*) и их ортологов у разных видов дрозофил.

ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

Виды рода *Drosophila* и условия культивирования.

В работе использовали лабораторную линию *D. melanogaster* дикого типа (Canton-S) и виды

подрода *Sophophora* (*D. simulans*, *D. mauritiana*, *D. yakuba*, *D. teissieri* и *D. pseudoobscura*) из коллекции кафедры генетики МГУ. Все использованные в работе виды культивировали на стандартной агаризованной среде при 21°C. В экспериментах использовали только самок на стадии имаго, сбор которых осуществляли спустя 6–8 дней после вылета. Для индукции окислительного стресса имаго помещали на агаризованную питательную среду, содержащую персульфат аммония (“Fermentas”, США). В экспериментах по изучению экспрессии использовали персульфат аммония в конечной концентрации 0.1 М. В качестве контроля использовали мух, помещенных на агаризованную питательную среду без добавления персульфата аммония. Для снижения дисперсии значений уровня экспрессии изучаемых генов в дальнейшем, учитывая индивидуальную реакцию на воздействие, и для повышения достоверности результатов эксперимент проводили параллельно на 10 группах мух (шесть тестируемых и четыре контрольные) по 10 самок в каждой, помещенных в отдельные пробирки, при температуре 21°C. Жизнеспособность при окислительном стрессе изучали, используя персульфат аммония в конечных концентрациях 0.01 и 0.05 М. Использование более низкой концентрации позволило несколько увеличить время выживания особей на среде с персульфатом и увидеть значимые различия между более и менее чувствительными к персульфату видами. В опытах использовали семь групп мух каждого вида (пять тестируемых и две контрольные), помещенных в отдельные пробирки и в сумме дававших не менее 50 самок на опыт и 20 на контроль. Количество выживших мух фиксировали спустя 18–30, 42–52 и 66–74 ч после посадки с промежутками 2 ч в каждом интервале.

Выделение суммарной РНК и ОТ-ПЦР. РНК выделяли с помощью реактива ExtractRNA (“Евроген”, Россия) по протоколу, рекомендованному фирмой-производителем. РНК выделяли не из отдельных мух, а из пулов (по пять самок), отобранных из каждой группы (экспериментальной и контрольной) спустя 24 ч после посадки. Таким образом, в каждом эксперименте получено шесть независимых проб РНК из мух, находившихся в стрессовых условиях, и четыре независимые контрольные пробы. Концентрацию РНК измеряли спектрофотометрически на приборе NanoDrop (“Peglab”). Перед постановкой реакции обратной транскрипции (ОТ) образцы РНК обрабатывали ДНКазой I (“Fermentas”) согласно протоколу фирмы. Для приготовления реакционной смеси использовали набор реагентов “пятикратная реакционная смесь с горячим стартом на основе Taq-полимеразы с интеркалирующим красителем SYBR Green I” (“Евроген”) в соответствии с протоколом фирмы-производителя. Реакцию ставили в амплификаторе Mini Opticon Real-Time PCR

Таблица 1. Праймеры, использованные в работе

Ген	Прямой праймер (5' → 3')	Обратный праймер (5' → 3')
<i>Gagr amst</i>	CCTCGATGGCAGTGATCCTG	ATTTGTCGCGTGAAGACCG
<i>Gagr y</i>	CCTCGATGGCAGTGATCCTG	CGCTCGTTCTGTGCACTTAT
<i>Gagr p</i>	AAC TTCGATGGCAGCGATC	CCGATCAGTGCACTTGTCT
<i>vir-1 ams</i>	GGCCGAATTCGATCAGGATC	GCCAAATCGTTGACGGCTAT
<i>vir-1 ty</i>	GGCCGAATTCGATCAGGATC	GCCACATCGTTGACGGCTAT
<i>vir-1 p</i>	AACACTGGACCAGGATCAGG	GCAAAATCGTTGACGGCGAT
<i>Rel amsty</i>	TCATACACACCGCCAAGAAG	AGTTCATGTCCTTGGCCTCA
<i>Rel p</i>	TCATACACACGGCCAAGAAG	TTTCCAGCTTCTGTACCTCC
<i>upd3 ams</i>	GACCAGGAACCCAGTGGA	GGCAAAC TGGGACATGGC
<i>upd3 ty</i>	CAATGTGACCAACCTGCTCT	CGCGGCAGGATCTTGTAG
<i>upd3 p</i>	GCCAACACGCGTGAAAGA	ACTGCTGCAACTGGTACATG
<i>RpL32 amsty</i>	GCACCAAGCACTTCATCCG	GACGCACTCTGTTGTCGATCA
<i>RpL32 p</i>	AGCACTTCATCCGTCACCA	GTAACCGATGTTGGGCATCA
<i>aTub84B</i>	TACGCCAAGCGTGCCTT	GTCCATGCCGACCTCCTC

Сокращения: *a* – *D. mauritiana*, *m* – *D. melanogaster*, *p* – *D. pseudoobscura*, *s* – *D. simulans*, *t* – *D. teissieri*, *y* – *D. yakuba*.

System (“Bio-Rad Laboratories”, США). Определяли относительную экспрессию генов *Gagr*, *vir-1*, *upd3* и *Rel*, нормированную на экспрессию двух референсных генов – α *Tub84B* и *RpL32*. Фрагменты кДНК амплифицировали с использованием праймеров, подобранных к последовательностям мРНК тестируемых генов, с помощью модулей Primer3 и MUSCLE пакета программ UGENE v.41 (табл. 1).

ПЦР проводили, используя следующую программу: предварительный этап плавления ДНК – 10 мин, 95°C, затем 40 циклов: плавление – 15 с, 95°C; отжиг праймеров – 30 с, 55°C; синтез – 60 с, 72°C. После синтеза в каждом цикле измеряли уровень флуоресценции в пробах. Одновременно ставили отрицательные контроли для каждого образца: пробы, обработанные ДНКазой I и не прошедшие ОТ, а также пробы, содержащие реактивы для ПЦР (включая праймеры), инкубированные без внесения амплифицируемой ДНК. Для проверки специфичности реакции после завершения ПЦР определяли температуру плавления продуктов. Продукт ПЦР анализировали с помощью электрофореза в агарозном геле. Относительный уровень экспрессии рассчитывали методом Δ Сt с нормированием на два референсных гена. Статистическую значимость различий нормированной экспрессии генов оценивали с помощью теста Манна–Уитни ($p < 0.05$).

Множественное выравнивание и биоинформатический анализ биологических последовательностей. В работе использовали нуклеотидные и аминокислотные последовательности, загруженные из баз данных FlyBase (<https://flybase.org>) и NCBI (<https://www.ncbi.nlm.nih.gov/>). Выравнивание по-

следовательностей ДНК и белков проводили на платформе Unipro Ugene v.41 с помощью инструмента MUSCLE (<https://www.ebi.ac.uk/msa>). Поиск и визуализацию консервативных мотивов осуществляли с помощью Интернет-ресурса WebLogo 3 (weblogo.threeplusone.com).

РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ

Влияние окислительного стресса, вызванного персульфатом аммония, на жизнеспособность видов рода Drosophila

В качестве индуктора окислительного стресса использовали персульфат аммония [27]. Окисляя биологические макромолекулы, персульфат-анион и продукты его распада вызывают обширные повреждения клеточных структур, способствуя быстрой активации стрессовых каскадов. Показана высокая эффективность персульфата аммония как индуктора экспрессии *Gagr* и маркеров окислительного стресса у *D. melanogaster* [18].

Для выявления общих закономерностей функционирования сигнальных путей ответа на окислительный стресс у разных видов дрозофил оценили жизнеспособность мух каждого вида на среде, содержащей персульфат аммония. Концентрация окислителя 0.1 М, использованная для определения экспрессии генов-маркеров окислительного стресса в работе [18], оказалась неподходящей для длительной экспозиции: гибель мух нарастала слишком быстро, не позволяя зафиксировать подробную динамику (у некоторых видов 100% мух погибли менее чем за сутки). Понижение концентрации в предварительных опытах до 0.05 М расширило возможный диапазон наблюдений до 3 сут.

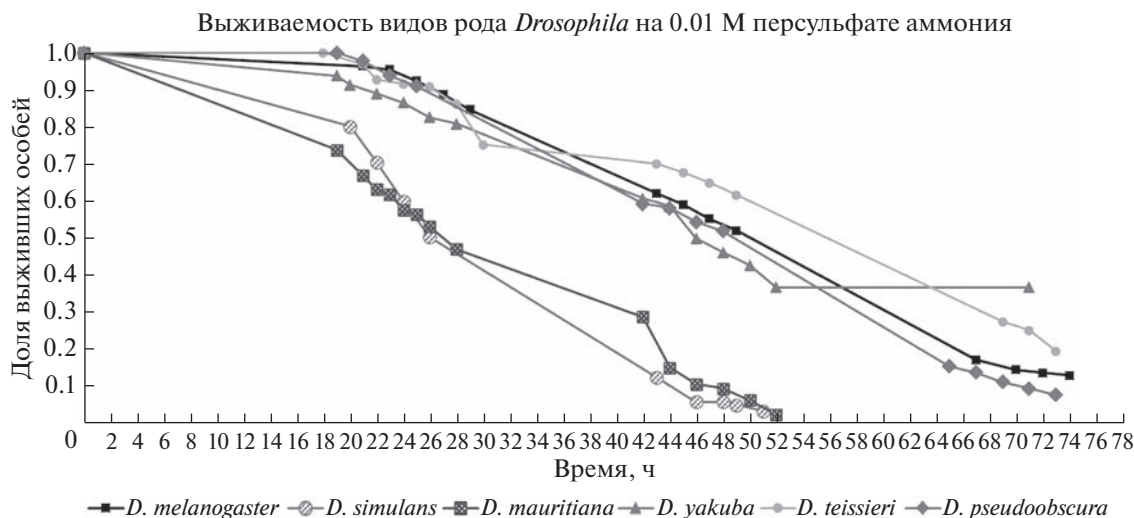


Рис. 1. Доля выживших мух в зависимости от времени экспозиции на 0.05 М персульфате аммония.

Результаты эксперимента представлены на рис. 1. Время, за которое происходит гибель половины мух большинства изученных видов, составляет порядка 48 ч, но два близкородственных вида (*D. simulans* и *D. mauritiana*) из подгруппы *melanogaster* проявляют повышенную чувствительность к 0.05 М персульфату. У этих видов гибель половины особей наступает уже через 24 ч. Повышенная чувствительность к персульфату аммония может быть связана с изменениями экспрессии стресс-активируемых генов, специфичными для *D. simulans* и *D. mauritiana*, поскольку этот эффект отсутствует у *D. melanogaster*, а динамика жизнеспособности при окислительном стрессе принципиально схожа с динамикой у менее близкородственных видов (включая *D. pseudoobscura*).

Влияние окислительного стресса на экспрессию гена *Gagr* и стрессовых маркеров у видов рода *Drosophila*

Влияние окислительного стресса на экспрессию гена *Gagr* и генов-маркеров окислительного стресса изучали с использованием 0.1 М персульфата аммония. Оценивали экспрессию как самого гена *Gagr*, так и ряда других стресс-активируемых генов, вероятно, имеющих общие пути регуляции с доместицированным геном *Gagr*: *vir-1*, *upd3* и *Rel*.

Согласно полученным данным, у разных видов экспрессия гена *Gagr* в стрессовых условиях может существенно различаться (рис. 2). Эффективную индукцию демонстрируют виды *D. melanogaster*, *D. mauritiana* и *D. yakuba*. У *D. simulans* и *D. teissieri* индукция не зафиксирована, но уровень экспрессии *Gagr* в контроле у этих видов сопоставим с уровнем у *D. yakuba* в стрессовых условиях. У *D. pseudoobscura* экспрессия *Gagr* также

повышена и не реагирует на присутствие окислителя.

Эффективность индукции *Gagr*, зарегистрированная у разных видов, согласуется с различиями в структуре промотора доместицированного гена, исследованными в работе [18]. Промотор *Gagr* у *D. melanogaster* содержит сайт связывания ТФ пути JNK *kaу* и два сайта связывания ТФ пути Jak-STAT STAT92E. У *D. mauritiana* и *D. simulans* промоторы гомологов *Gagr* содержат сайт связывания *kaу* и два сайта связывания STAT92E. У *D. yakuba* и *D. teissieri* ортологичные гены содержат сайт связывания *kaу* и только один сайт связывания STAT92E. У *D. pseudoobscura* есть только сайт связывания *kaу*, но нет сайтов связывания STAT92E.

Повышение экспрессии *vir-1* в условиях окислительного стресса обнаружено у большинства исследованных видов. У *D. teissieri* и *D. pseudoobscura* эффективность индукции снижена (в 2.5 раза по сравнению с более чем десятикратным увеличением экспрессии у *D. melanogaster* и *D. yakuba*). У двух близкородственных видов (*D. simulans* и *D. mauritiana*) *vir-1* не индуцируется и экспрессируется на низком уровне. Эти же виды показали наименьшую жизнеспособность в условиях окислительного стресса. Таким образом, жизнеспособность этих видов в условиях окислительного стресса может коррелировать с уровнем экспрессии гена *vir-1*.

Мы провели поиск и сравнительный анализ сайтов связывания ТФ STAT92E в промоторных областях *vir-1* у разных видов, поскольку ген *vir-1*, как и *Gagr*, является мишенью ТФ STAT92E. В поиске мы опирались на данные, полученные в работе [19], в которой экспериментально обнаруже-

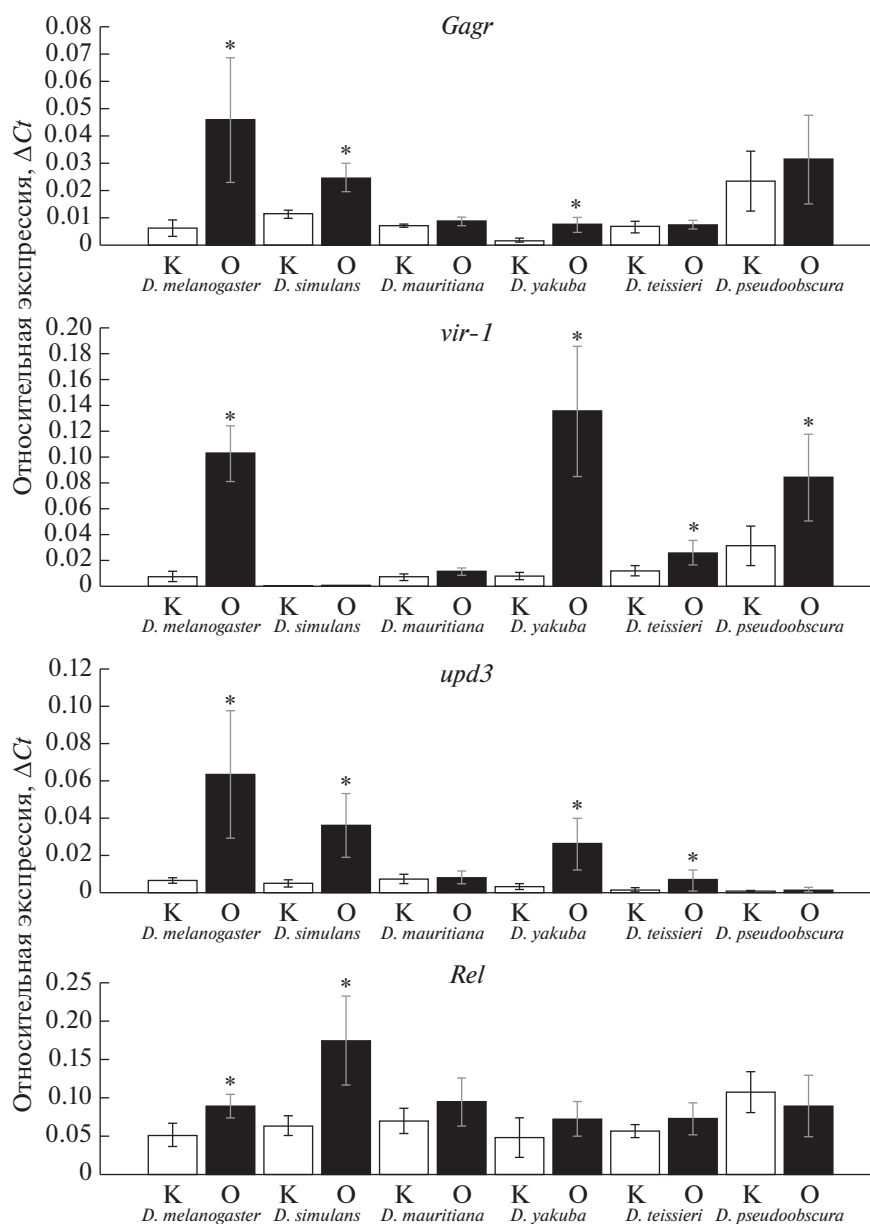


Рис. 2. Относительный уровень экспрессии генов *Gagr*, *vir-1*, *upd3* и *Rel* у самок *D. melanogaster* и их гомологов у *D. mauritiana*, *D. simulans*, *D. yakuba*, *D. teissieri*, *D. pseudoobscura*. К — уровень экспрессии у контрольных особей (не подвергавшихся окислительному стрессу); О — уровень экспрессии у особей, подвергавшихся воздействию персульфата аммония. Уровень экспрессии нормирован на экспрессию двух референсных генов, α *Tub84B* и *RpL32*, для *D. melanogaster* и их гомологов для других видов. Данные представлены в виде среднего со стандартным отклонением. *Значимые различия (тест Манна–Уитни, $p < 0.05$, $N = 6$).

ны два сайта связывания ТФ STAT92E в промоторной области *vir-1* *D. melanogaster*.

Согласно полученным данным (рис. 3), промоторные области всех исследуемых видов, кроме *D. simulans* и *D. mauritiana*, содержат два сайта связывания ТФ STAT92E. Таким образом, слабый уровень экспрессии *vir-1* и индукции в условиях стресса у этих видов может объясняться более низким уровнем индукции экспрессии ТФ STAT92E.

От экспрессии гена *upd3* в стрессовых условиях во многом зависит эффективность активации пути Jak-STAT, а значит, индукция экспрессии генов *vir-1* и *Gagr* (при наличии хотя бы одного сайта связывания STAT92E в промоторе). У *D. melanogaster*, *D. mauritiana*, *D. yakuba* и *D. teissieri* экспрессия *upd3* значительно повышается в присутствии персульфата аммония, обеспечивая эффективную передачу сигнала по пути Jak-STAT.

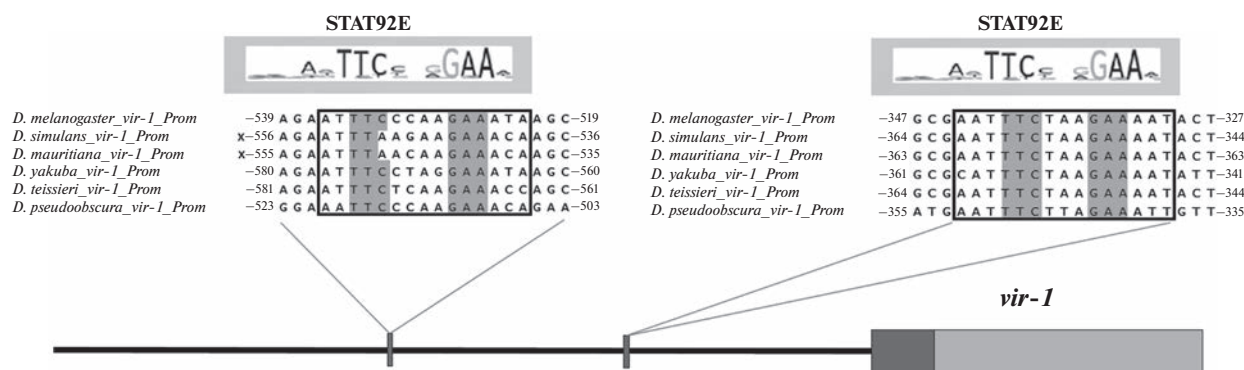


Рис. 3. Промоторные области гена *vir-1* и его ортологов у видов *Drosophila*. В рамках показано множественное выравнивание ортологических последовательностей промоторных областей *vir-1* у разных видов, соответствующих консенсусной последовательности сайта связывания ТФ STAT92E.

У *D. simulans* и *D. pseudoobscura* наблюдается иная ситуация. *D. simulans* не демонстрирует индукции *upd3*, что может существенно ограничивать STAT-зависимую активацию *vir-1* и *Gagr* и приводить к наблюдаемой низкой жизнеспособности в условиях стресса. У *D. pseudoobscura* экспрессия *upd3* в опыте остается на низком уровне, что может свидетельствовать о пониженной роли пути Jak-STAT в реакциях ответа на индуцированный персульфатом стресс у этого вида и об альтернативном пути активации гена *vir-1*.

Экспрессия гена *Rel* у большинства изученных видов значимо не повышалась в ответ на стресс, вызванный персульфатом аммония. Это может быть связано с особенностями функционирования сигнального пути IMD, необходимого для передачи сигнала на каскады JNK и MAPK, но в то же время ограничивающего биологические эффекты этих путей (как напрямую, так и вследствие развития иммунологических реакций). Экспрессия *Rel* в опыте значительно (в 3 раза) повышена только у *D. mauritiana* и вместе со слабой индукцией *vir-1* может обуславливать низкую жизнеспособность этого вида.

ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Важнейшую роль в обеспечении ответа дрозофилы на окислительный стресс играют пути IMD [26], JNK [24], MAPK (p38) [28, 29] и Jak-STAT [23]. Ген *Gagr* приобретен предком дрозофил в результате молекулярной доместикации ретроэлемента, в ходе дальнейшего филогенеза он был вовлечен в сложную систему ответа на окислительный стресс, став частью JNK-, а затем и STAT-регулируемой группы стресс-индуцируемых генов [17, 18]. Ген *Gagr* и его белковый продукт предположительно участвуют в обеспечении регенеративной пролиферации и выхода клеток из стрессового состояния, способствуя, возможно, выживанию организма в присутствии окислителей. Структура

гена *Gagr* и свойства его продукта различаются у разных видов, что может означать различную степень вовлеченности доместического гена в клеточные процессы [17, 18].

Полученные результаты свидетельствуют о различном характере экспрессии *Gagr* и функционирования пути Jak-STAT в условиях окислительного стресса у разных видов дрозофил. Эффективность стресс-индукции *Gagr* зависит от количества сайтов связывания ТФ STAT92E на промоторе доместического гена [18]. Виды, предположительно обладающие двумя сайтами связывания STAT92E (*D. melanogaster*, *D. mauritiana*), имеют более высокий уровень экспрессии *Gagr* в опыте, чем *D. yakuba*, содержащий только один сайт связывания. У всех этих видов существует также сходство стресс-индукции *Gagr* и *upd3* (от которого зависит STAT-регуляция доместического гена) [21].

У вида *D. pseudoobscura*, не входящего в подгруппу *melanogaster*, экспрессия ортолога *Gagr* находится на высоком уровне на фоне низкой экспрессии *upd3* и не зависит от присутствия стрессового фактора. Это согласуется с другой структурой промотора (без сайтов связывания STAT92E) и, предположительно, иными свойствами белкового продукта ортолога *Gagr* у этого вида, предполагающими меньшую вовлеченность доместического гена в стрессовые каскады [18].

Наличие хотя бы одного сайта связывания STAT92E в промоторе *Gagr* может играть существенную роль в обеспечении совместной регуляции этого гена с *vir-1* — маркером окислительного стресса [19, 20]. Согласно экспериментальным данным, наибольшая стресс-индукция *vir-1* наблюдается у тех видов, у которых эффективно индуцируется *Gagr* (кроме *D. mauritiana*) (рис. 4). Молекулярные функции обоих этих генов не вполне ясны, но их совместная экспрессия может быть важна для выживания мух в стрессовых

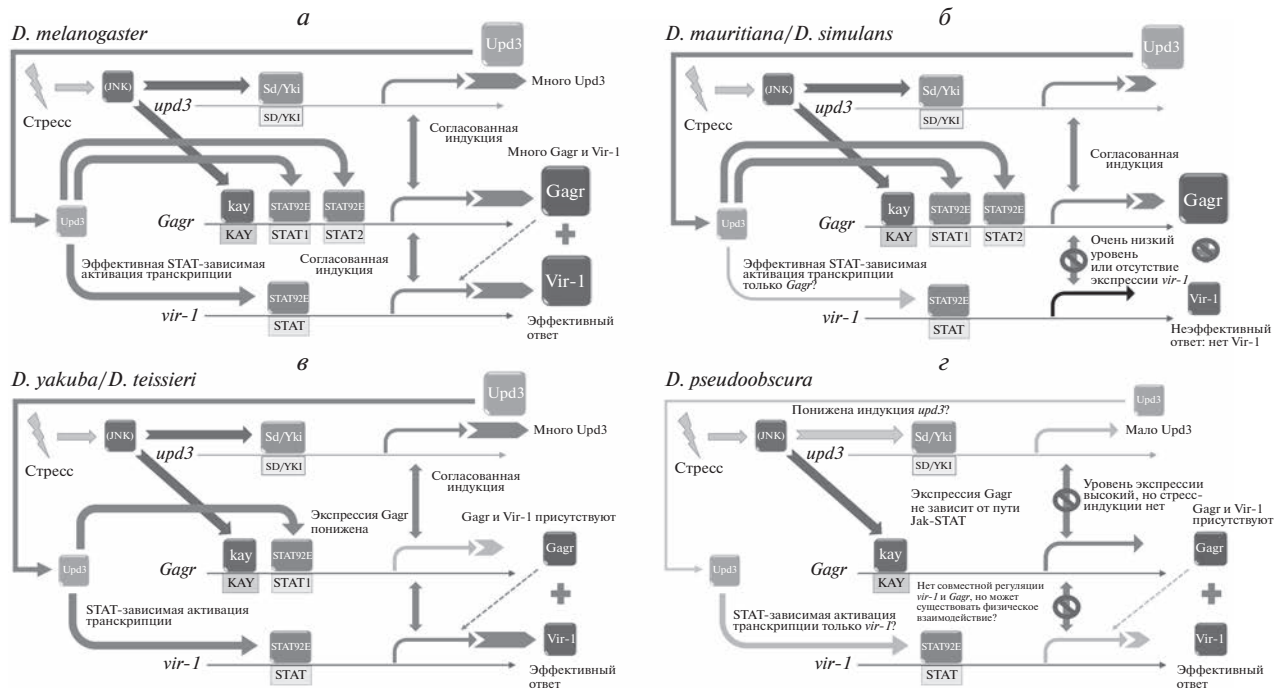


Рис. 4. Предполагаемая схема стресс-зависимой регуляции *Gagr* у видов рода *Drosophila*. а — Схема регуляции у *D. melanogaster*, б — *D. mauritiana* и *D. simulans*, имеющие два сайта связывания STAT92E в промоторе *Gagr*, но различный характер индукции *vir-1*. в — Схема регуляции у *D. yakuba* и *D. teissieri*, имеющих один (действующий) сайт связывания STAT. г — Схема регуляции у *D. pseudoobscura*, не имеющей сайтов связывания STAT в промоторе доместицированного гена. Пояснения в тексте.

условиях. Согласно полученным результатам, самую низкую жизнеспособность на персульфате аммония показали виды, у которых наблюдалась низкая экспрессия *vir-1* в стрессовых условиях (*D. simulans* и *D. mauritiana*). Для *D. simulans* также характерно отсутствие индукции *upd3*, что может означать сниженную роль Jak-STAT в ответе на стресс у этого вида (рис. 4б). С другой стороны, более слабая стресс-индукция *vir-1* (и *upd3*) у *D. teissieri* по сравнению с *D. yakuba* на фоне одинаковой экспрессии *Gagr* не повлияла на жизнеспособность. Это может быть связано с непосредственным участием *Gagr* в регенеративной пролиферации и, возможно, во взаимодействии с *vir-1*, компенсирующим его пониженную экспрессию (рис. 4в). У *D. pseudoobscura*, не входящего в подгруппу *melanogaster*, характер экспрессии *vir-1* в условиях стресса различается и, по-видимому, никак не согласован с экспрессией *Gagr* (не реагирующей на стресс в обоих случаях). Это согласуется с предполагаемым отсутствием STAT-зависимой регуляции *Gagr* у этого вида (рис. 4г).

В отличие от стресс-индуцируемой экспрессии *Gagr*, *vir-1* и *upd3*, демонстрирующей межвидовую изменчивость, ген *Rel* сильно индуцируется только у *D. mauritiana* и очень слабо — у *D. melanogaster*. Практически полное отсутствие межвидовых различий в экспрессии *Rel* по сравнению с другими

тестируемыми генами свидетельствует об отсутствии прямого влияния сигнального пути IMD на стресс-индукцию *Gagr* у исследованных видов дрозофил.

Полученные результаты свидетельствуют о роле вовлеченности доместицированного гена в клеточные сигнальные каскады стрессового ответа в ходе филогенеза рода *Drosophila*, но конкретные молекулярные функции гена *Gagr* и его белкового продукта, а также его партнеров до сих пор недостаточно изучены. Слабо охарактеризованы и функции ряда генов, имеющих общую с *Gagr* JNK/STAT-регуляцию, и по характеру экспрессии являющихся маркерами стрессовых воздействий (в том числе *vir-1*). Функциональный анализ генетического и белкового окружения *Gagr* позволит лучше понять природу эволюционных преобразований, происходящих в процессе молекулярной доместики.

Работа поддержана грантом Российского научного фонда (№ 22-24-00305).

Статья не содержит каких-либо исследований с участием людей или животных в качестве объектов исследований.

Авторы заявляют об отсутствии конфликта интересов.

СПИСОК ЛИТЕРАТУРЫ

- Dupressoir A., Marceau G., Vernochet C., Benit L., Kanellopoulos C., Sapin V., Heidmann T. (2005) Syncytin-A and syncytin-B, two fusogenic placenta-specific murine envelope genes of retroviral origin conserved in Muridae. *Proc. Natl. Acad. Sci. USA*. **102**(3), 725–730.
- Emera D., Wagner G.P. (2012) Transposable element recruitments in the mammalian placenta: impacts and mechanisms. *Brief. Funct. Genomics*. **11**(4), 267–276.
- Henke C., Strissel P.L., Schubert M.-T., Mitchell M., Stolt C.C., Faschingbauer F., Beckmann M.W., Strick R. (2015) Selective expression of sense and antisense transcripts of the Sushi-ichirelated retrotransposon-derived family during mouse placentogenesis. *Retrovirology*. **12**, 9.
- Kämmerer U., Germeyer A., Stengel S., Kapp M., Denner J. (2011) Human endogenous retrovirus K (HERV-K) is expressed in villous and extravillous cytotrophoblast cells of the human placenta. *J. Reprod. Immunol.* **91**(1–2), 1–8.
- Mallet F., Bouton O., Prudhomme S., Cheynet V., Oriol G., Bonnaud B., Lucotte G., Duret L., Mandrand B. (2004) The endogenous retroviral locus ERVWE1 is a *bona fide* gene involved in hominoid placental physiology. *Proc. Natl. Acad. Sci. USA*. **101**(6), 1731–1736.
- Cho G., Lim Y., Golden J.A. (2011) XLMR candidate mouse gene, *Zcchc12* (*Sign1*) is a novel marker of Cajal–Retzius cells. *Gene Exp. Patterns*. **11**, 216–220.
- Navas-Pérez E., Vicente-García C., Mirra S., Burguera D., Fernández-Castillo N., Ferrán J.L., López-Mayorga M., Alaiz-Noya M., Suárez-Pereira I., Antón-Galindo E., Ulloa F., Herrera-Úbeda C., Cuscó P., Falcón-Moya R., Rodríguez-Moreno A., D’Aniello S., Cormand B., Marfany G., Soriano E., Carrión Á.M., Carvajal J.J., García-Fernández J. (2020) Characterization of an eutherian gene cluster generated after transposon domestication identifies *Bex3* as relevant for advanced neurological functions. *Genome Biol.* **21**, 267.
- Benit L., Parseval N. De, Casella J.F., Callebaut I., Cordonnier A., Heidmann T. (1997) Cloning of a new murine endogenous retrovirus, MuERV-L, with strong similarity to the human HERV-L element and with a *gag* coding sequence closely related to the *Fv1* restriction gene. *J. Virol.* **71**(7), 5652–5657.
- Kaneko-Ishino T., Ishino F. (2012) The role of genes domesticated from LTR retrotransposons and retroviruses in mammals. *Front. Microbiol.* **3**, 262.
- Yap M.W., Colbeck E., Ellis S.A., Stoye J.P. (2014) Evolution of the retroviral restriction gene *Fv1*: inhibition of non-MLV retroviruses. *PLoS Pathogens*. **10**(3), e1003968.
- Hantak M.P., Einstein J., Kearns R.B., Shepherd J.D. (2021) Intercellular communication in the nervous system goes viral. *Trends Neurosci.* **44**(4), 248–259.
- Schrader L., Schmitz J. (2019) The impact of transposable elements in adaptive evolution. *Mol. Ecology*. **28**, 1537–1549.
- Chen W., Schwalie P.C., Pankevich E.V., Gubelmann C., Raghav S.K., Dainese R., Cassano M., Imbeault M., Jang S.M., Russeil J., Delessa T., Duc J., Trono D., Wolfrum C., Deplancke B. (2019) ZFP30 promotes adipogenesis through the KAP1-mediated activation of a retrotransposon-derived *Pparg2* enhancer. *Nat. Commun.* **10**(1), 1809.
- Deng B., Xu W., Wang Z., Liu C., Lin P., Li B., Huang Q., Yang J., Zhou H., Qu L. (2019) An LTR retrotransposon-derived lncRNA interacts with RNF169 to promote homologous recombination. *EMBO Rep.* **20**(11), e47650.
- Nefedova L., Kim A. (2017) Mechanisms of LTR-retroelement transposition: lessons from *Drosophila melanogaster*. *Viruses*. **9**(4), 81.
- Malik H.S., Henikoff S. (2005) Positive selection of Iris, a retroviral envelope-derived host gene in *Drosophila melanogaster*. *PLoS Genet.* **1**(4), e44.
- Nefedova L.N., Kuzmin I. V., Makhnovskii P.A., Kim A.I. (2014) Domesticated retroviral *GAG* gene in *Drosophila*: new functions for an old gene. *Virology*. **450–451**, 196.
- Makhnovskii P., Balakireva Y., Nefedova L., Lavrenov A., Kuzmin I., Kim A. (2020) Domesticated *gag* gene of *Drosophila* LTR retrotransposons is involved in response to oxidative stress. *Genes*. **11**(4), 396.
- Dostert C., Jouanguy E., Irving P., Troxler L., Galiana-Arnoux D., Hetru C., Hoffmann J.A., Imler J.-L. (2005) The Jak-STAT signaling pathway is required but not sufficient for the antiviral response of *Drosophila*. *Nat. Immunol.* **6**(9), 946–953.
- Gruenewald C., Botella J.A., Bayersdorfer F., Navarro J.A., Schneuwly S. (2009) Hyperoxia-induced neurodegeneration as a tool to identify neuroprotective genes in *Drosophila melanogaster*. *Free Rad. Biol. Med.* **46**(12), 1668–1676.
- Wright V.M., Vogt K.L., Smythe E., Zeidler M.P. (2011) Differential activities of the *Drosophila* JAK/STAT pathway ligands Upd, Upd2 and Upd3. *Cell. Signal.* **23**(5), 920–927.
- Santabarbara-Ruiz P., Lopez-Santillan M., Martinez-Rodriguez I., Binagui-Casas A., Perez L., Milan M., Corominas M., Serras F. (2015) ROS-Induced JNK and p38 signaling is required for unpaired cytokine activation during *Drosophila* regeneration. *PLoS Genet.* **11**(10), e1005595.
- Staley B.K., Irvine K.D. (2010) Warts and Yorkie mediate intestinal regeneration by influencing stem cell proliferation. *Curr. Biol.* **20**(17), 1580–1587.
- Biteau B., Karpac J., Hwangbo D., Jasper H. (2011) Regulation of *Drosophila* lifespan by JNK signaling. *Exp. Gerontol.* **46**(5), 349–354.
- Ermolaeva M.A., Schumacher B. (2014) Systemic DNA damage responses: organismal adaptations to genome instability. *Trends Genet.* **30**(3), 95–102.
- Myllymäki H., Valanne S., Rämetsä M. (2014) The *Drosophila* Imd signaling pathway. *J. Immunol.* **192**, 3455–3462.
- Jakob H., Leininger S., Lehmann T., Jacobi S., Gutewort S. (2007) Peroxo compounds, inorganic. In: *Ullmann's Encyclopedia of Industrial Chemistry*. Wiley-VCH Verlag GmbH. https://doi.org/10.1002/14356007.a19_177.pub2
- Belozarov V.E., Lin Z.-Y., Gingras A.-C., McDermott J.C., Michael Siu K.W. (2012) High resolution protein interaction map of the *Drosophila melanogaster* p38 mitogen-activated protein kinases reveals limited functional redundancy. *Mol. Cell. Biol.* **32**(18), 3695–3706.
- Brun S., Vidal S., Spellman P., Takahashi K., Tricoire H., Lemaitre B. (2006) The MAPKKK Mekk1 regulates the expression of *Turandot* stress genes in response to septic injury in *Drosophila*. *Genes Cells*. **11**, 397–407.

Involvement of the *Gagr* Gene, a Domesticated *gag* Gene of Retrovirus, in the Stress Response Pathway in Different *Drosophila* Species

A. N. Gigin¹ and L. N. Nefedova^{1, *}

¹*Department of Biology, Lomonosov Moscow State University, Moscow, 119234 Russia*

**e-mail: nefedova@mail.bio.msu.ru*

The *Gagr* gene is a domesticated *gag* retroelement gene in *Drosophila melanogaster*, whose function is associated with a stress response. The protein products of the *Gagr* gene and its homologues in different *Drosophila* species have a highly conserved structure; however, they demonstrate a certain variability in the promoter region of the gene, apparently associated with the gradual acquisition of a new function and involvement in new signaling pathways. In this work we studied the effect of oxidative stress caused by ammonium persulfate on the survival of various species of the genus *Drosophila* (*D. melanogaster*, *D. mauritiana*, *D. simulans*, *D. yakuba*, *D. teissieri*, *D. pseudoobscura*), analyzed the correlation between the structure of promoter regions and stress-induced changes in the expression of the *Gagr* gene and its homologues in different *Drosophila* species and comparison of stress-induced changes in the expression of oxidative stress markers: Jak-STAT signaling pathway activator gene *upd3*, Jak-STAT pathway effector *vir-1*, and signaling pathway target IMD *Rel*. It was found that in *D. simulans* and *D. mauritiana* sensitivity to ammonium persulfate is significantly increased, which correlates with a reduced level of transcription of *vir-1* gene orthologues. The latter is due to a decrease in the number of binding sites for the transcription factor STAT92E, a component of the Jak-STAT signaling pathway, in the *vir-1* promoter region. Consistent changes in the expression of the *Gagr*, *upd3*, *vir-1* genes are observed in all species of the *melanogaster* subgroup, except *D. pseudoobscura*, which indicates an increase in the role of *Gagr* in the regulation of stress response pathways during the phylogenesis of the genus *Drosophila*.

Keywords: domestication, retroelements, *gag*, *Drosophila*, stress response

ГЕНОМИКА.
ТРАНСКРИПТОМИКА

УДК 577.218

ПРОФИЛИРОВАНИЕ ЭКСПРЕССИИ микроРНК
В В-КРУПНОКЛЕТОЧНОЙ ЛИМФОМЕ

© 2023 г. Ю. А. Веряскина^{a, b, *}, С. Е. Титов^{a, c}, И. Б. Ковынев^d, С. С. Фёдорова^d,
Я. Ю. Шебуняева^d, О. В. Антоненко^a, Т. И. Поспелова^d, И. Ф. Жимулёв^a

^aИнститут молекулярной и клеточной биологии Сибирского отделения Российской академии наук,
Новосибирск, 630090 Россия

^bФедеральный исследовательский центр Институт цитологии и генетики
Сибирского отделения Российской академии наук, Новосибирск, 630090 Россия

^cАО “Вектор-Бест”, Новосибирск, 630117 Россия

^dНовосибирский государственный медицинский университет Министерства здравоохранения Российской Федерации,
Новосибирск, 630091 Россия

*e-mail: microrna@inbox.ru

Поступила в редакцию 10.11.2022 г.

После доработки 05.12.2022 г.

Принята к публикации 13.12.2022 г.

Неходжкинские лимфомы (НХЛ) представляют собой гетерогенную группу онкологических заболеваний, различающихся по патогенезу и прогнозу. К основным методам лечения НХЛ относятся химиотерапия, иммунохимиотерапия и лучевая терапия. Однако значительная часть этих опухолей отличается химиорезистентностью или быстрым рецидивированием после непродолжительной ремиссии, индуцированной химиотерапией. В связи с этим актуален поиск альтернативных терапевтических циторедуктивных методов. Аберрантная экспрессия микроРНК (miРНК) — один из механизмов возникновения и опухолевой прогрессии злокачественных лимфоидных неоплазий. Мы проанализировали профиль экспрессии miРНК в биопсийном материале лимфоузлов, пораженных диффузной В-крупноклеточной лимфомой (DLBCL). Ключевым материалом исследования стали гистологические препараты лимфатических узлов, полученные в результате эксцизионной диагностической биопсии и обработанные с использованием классических гистоморфологических методов фиксации в формалине. Группу исследования составили пациенты с DLBCL ($n = 52$), а контрольную — пациенты с реактивной лимфоаденопатией (RL) ($n = 40$). Показано, что в DLBCL уровень экспрессии miR-150 снижен более чем в 12 раз ($p = 3.6 \times 10^{-15}$) в сравнении с RL. Биоинформационный анализ выявил участие miR-150 в регуляции гемопоэза и лимфопоэза. Полученные данные позволяют рассматривать miR-150 в качестве перспективной терапевтической мишени, обладающей большим потенциалом в клинической практике.

Ключевые слова: микроРНК, гемопоэз, лимфома, неходжкинская лимфома, диффузная В-крупноклеточная лимфома

DOI: 10.31857/S0026898423030175, **EDN:** CHWPDХ

Согласно классификации гематолимфоидных опухолей WHO-НАЕМ5, опубликованной Всемирной организации здравоохранения в 2022 году, лимфомы представляют собой гетерогенную группу онкологических заболеваний, различающихся по прогнозу и патогенезу. Неходжкинские лимфомы (НХЛ) делят на В- и Т-клеточные. К основным подтипам НХЛ относятся следующие: диффузная В-крупноклеточная лимфома (DLBCL), фолликулярная лимфома, лимфома Беркитта, хро-

нический лимфолейкоз, ассоциированные со слизистыми оболочками опухоли лимфоидной ткани, лимфома из клеток мантийной зоны и лимфома маргинальной зоны. Классификация лимфом основана на комбинации данных морфологии, иммунофенотипа, специфических генетических аномалий и клинических признаков заболевания. У пациентов с диагностированным заболеванием клиническое течение крайне вариабельно: от индолентного до агрессивного. Так что стратегии их лечения и ответа на терапию тоже сильно различаются, как и клинические результаты. Основными методами лечения НХЛ считаются химиотерапия, иммунохимиотерапия и лучевая терапия,

Сокращения: НХЛ — неходжкинская лимфома; DLBCL (diffuse large B-cell lymphoma) — диффузная В-крупноклеточная лимфома; miРНК (microRNA) — микроРНК; RL (reactive lymphadenopathy) — реактивная лимфоаденопатия.

однако значительная часть этого типа опухолей химиорезистентна или рецидивирует после лечения [1]. В связи с этим поиск новых терапевтических подходов сохраняет актуальность.

Успехи в изучении молекулярных механизмов, лежащих в основе лимфомагенеза, привели к разработке таргетных препаратов [2]. В последние годы в качестве таких лекарственных средств рассматривают микроРНК (miРНК) [3]. Аберрантная экспрессия miРНК описана для большинства видов злокачественных опухолей, включая лимфомы [4]. В экспериментах *in vitro* показано, что использование длинной интерферирующей РНК для одновременного ингибирования 13 онко-miРНК значительно снижало клеточную пролиферацию, индуцировало остановку клеточного цикла и апоптоз в клеточных линиях DLBCL в основном за счет повышения экспрессии белков PTEN, p27^{Kip1}, TIMP3, RECK и подавления p38/MAPK, Survivin, CDK4, c-Мyc [5]. Н. Due и соавт. [6] сообщали, что повышение уровня miR-155 в клеточных линиях DLBCL индуцировало изменения в чувствительности этих клеток к винкристину. Кроме того, трансфекция миметиков miR-197 и miR-187 в клеточные линии DLBCL повышала их чувствительность к доксорубину, усиливая апоптоз [7, 8]. Повышение экспрессии miR-10a и miR-26a ингибировало пролиферацию и индуцировало апоптоз клеток DLBCL [9, 10]. Повышение экспрессии miR-223-3p снижало пролиферацию и ускоряло апоптоз клеток в мантийноклеточной лимфоме *in vitro* и *in vivo* через модулирование сигнального пути CHUK/NF- κ B2 [11]. miR-373 значительно замедляла рост клеток Т-клеточной лимфомы [12]. М. Morales-Martinez и соавт. [13] отмечали, что miR-7 регулирует химиочувствительность НХЛ посредством негативной регуляции генов *YY1* и *KLF4* [13]. Показано, что miR-150 можно рассматривать в качестве потенциального терапевтического сенситизатора, который регулирует путь PI3K/AKT/mTOR при лечении НК/Т-клеточной лимфомы [14]. К. Musilova и соавт. [15] показали, что трансфекция miR-150 в клетки фолликулярной лимфомы приводила к значительному снижению доли клеток в S-фазе, что коррелировало со снижением пролиферации опухолевых клеток.

Возможность доставки синтетических миметиков или ингибиторов miРНК открыла новые терапевтические перспективы. MRX34 (миметик miR-34), mesomiR-1 (миметик miR-16) и коб-марсен (анти-miR-155) проявляли противоопухолевую активность в клинических испытаниях фазы I. Эти исследования не были специально разработаны для НХЛ, но в них участвовали пациенты с В-клеточными НХЛ [3, 16, 17].

Эти данные подтверждают терапевтический потенциал miРНК. Персонализированная ком-

бинированная терапия может напрямую воздействовать на главную проблему при лечении опухолей различной локализации — лекарственную устойчивость.

Целью работы было определение профиля экспрессии miРНК в материале лимфоузлов DLBCL и поиск miРНК, которые в перспективе могут быть использованы в качестве таргетных препаратов, позволяющих максимально персонализировать проводимую терапию.

ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

Исследуемые группы. В экспериментальную группу вошли пациенты с DLBCL ($n = 52$). Контрольную группу составили пациенты с реактивной лимфоаденопатией (RL) ($n = 40$). От каждого пациента было получено письменное информированное согласие, все данные обезличены. Исследование было одобрено Этическим комитетом Новосибирского государственного медицинского университета.

Исследование проведено на гистологических препаратах биоптатов опухолевых лимфатических узлов, которые были обработаны с использованием классических методов фиксации в формалине, обезжизивания в изопропиловом спирте, обезжиривания ксилолом и пропитаны парафином.

Экстракция РНК. Для выделения нуклеиновых кислот из фиксированных в формалине парафинизированных образцов проводили их депарафинизацию. Для этого в пробирку, содержащую 3 парафиновых среза ткани лимфатического узла толщиной 15 мкм, добавляли 1 мл минерального масла и тщательно перемешивали в течение 10 с на вортексе (“BioSan”, Латвия), затем пробирку переносили в термошейкер и инкубировали при 65°C и частоте перемешивания 1300 об/мин в течение 2 мин. Полученную суспензию центрифугировали при 13000–15000 g в течение 4 мин и удаляли надосадочную жидкость. В осадок вносили 1 мл 96%-ного этанола, перемешивали в течение 10 с на вортексе и центрифугировали при 13000–15000 g в течение 4 мин. Надосадочную жидкость удаляли, а в осадок вносили 1 мл 70%-ного этанола, центрифугировали при той же скорости в течение 2 мин. Полученный осадок использовали для выделения нуклеиновых кислот.

Выделение нуклеиновых кислот из образцов проводили с использованием набора реагентов РеалБест экстракция 100 (АО “Вектор-Бест”, Россия).

Выбор miРНК. На основании анализа данных Е. Sebestyén и соавт. [18] для исследования были выбраны miРНК, число копий которых в образце превышало 100 единиц, и они были представлены, как минимум, в 80% исследованных авторами образ-

Таблица 1. Сравнительный анализ уровней miРНК, дифференциально экспрессируемых в образцах пациентов с DLBCL и RL

miРНК	Уровень различий ^a	<i>p</i> -value ^b	miРНК	Уровень различий ^a	<i>p</i> -value ^b
miR-26b	–3.64	1.5×10^{-13}	miR-29b	–2.12	1.1×10^{-7}
miR-30b	–1.96	1.2×10^{-4}	let-7c	–2.75	3×10^{-6}
miR-150	–12.47	3.6×10^{-15}	let-7f	–1.76	2.9×10^{-9}
miR-451a	–2.57	4.8×10^{-3}	miR-9	1.55	7×10^{-3}
miR-574	–1.56	3.3×10^{-3}	miR-23a	–1.35	1.7×10^{-2}
let-7a	–2.33	1.1×10^{-10}	miR-23b	–2.01	1×10^{-3}
miR-124	1.90	1.2×10^{-2}	miR-96	1.54	1.9×10^{-2}
miR-144	–1.78	5×10^{-3}	miR-148b	–1.83	6.2×10^{-12}
miR-15a	–1.47	1.1×10^{-4}	miR-128	–2.00	4.1×10^{-7}
miR-16	–1.75	9×10^{-6}	miR-221	–2.47	1.5×10^{-4}
miR-196b	–2.50	1.7×10^{-3}			

^aРазличия в уровнях экспрессии представлены как кратные величины.^bРазличия считали статистически значимыми при $p < 0.05$.

цов. Этим критериям удовлетворяло 29 miРНК: miR-124-3p, -144-5p, -15a-5p, -16-5p, -196b-5p, -221-3p, -29b-3p, -148b-3p, -150-5p, -18a-5p, -183-5p, -185-5p, -205-5p, -20a-5p, -23a-3p, -23b-3p, -26b-5p, -30b-5p, -34a-5p, -451a, -9-5p, -128-3p, -141-3p, -200b-3p, -574-3p, -96-5p, let-7a-5p, let-7c-5p и let-7f-5p. Для нормализации использовали среднее геометрическое значений C_T трех miРНК: miR-378-3p, -191-5p и -103a-3p, — которые тоже были выбраны на основе литературных данных [19–21]. Все олигонуклеотиды были синтезированы в АО “Вектор-Бест”. В зависимости от системы значение E (эффективность реакции) варьировало от 92.5 до 99.7%.

Обратная транскрипция (ОТ). Синтез кДНК проводили в реакционной смеси объемом 30 мкл, содержащей 3 мкл выделенной РНК, 16.2 мкл 40%-ного раствора трегалозы, 3 мкл 10× буфера для ОТ (500 мМ Трис-НСl, рН 8.3 (при 25°C), 500 мМ KCl, 40 мМ MgCl₂), 3 мкл 4 мМ раствора дезоксинуклеозидтрифосфатов, 3 мкл 10%-ного раствора BSA, 0.32 мкл обратной транскриптазы (АО “Вектор-Бест”), 1.5 мкл 10 мМ раствора соответствующего праймера для ОТ. Смесь инкубировали в течение 15 мин при 16°C и 15 мин при 42°C с последующей инактивацией в течение 2 мин при 95°C.

ПЦР в реальном времени. Уровни экспрессии miРНК оценивали методом ПЦР в реальном времени на амплификаторе CFX96 (“Bio-Rad Laboratories”, США). Реакционная смесь объемом 30 мкл содержала 3 мкл образца кДНК, ПЦР-буфер (АО “Вектор-Бест”), 0.5 мкМ каждого праймера и 0.25 мкМ зонд. Условия реакции: 2 мин при 50°C, 2 мин при 94°C и 50 циклов денатурации (10 с при 94°C), отжига и элонгации цепи (20 с при 60°C).

Статистический анализ. Статистический анализ выполнен с использованием программы Statistica v13.1 и *U*-критерия Манна–Уитни. Значения $p < 0.05$ считали статистически значимыми.

РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ

Анализ экспрессии исследуемых miРНК в клинических образцах

Методом ОТ-ПЦР в реальном времени проанализированы уровни экспрессии 29 miРНК: miR-124, -144, -15a, -16, -196b, -221, -29b, -148b, -150, -18a, -183, -185, -205, -20a, -23a, -23b, -26b, -30b, -34a, -451a, -9, -128, -141, -200b, -574, -96, let-7a, let-7c и let-7f — в 52 образцах DLBCL и 40 образцах RL. В образцах DLBCL по сравнению с RL выявлено статистически значимое изменение экспрессии для 21 miРНК: снижение для miR-26b, -30b, -150, -451a, -574, -144, -15a, -16, -196b, -221, -29b, -23a, -23b, -148, -128, let-7a, let-7c, let-7f и повышение для miR-124, -9 и -96 ($p < 0.05$) (табл. 1).

Наиболее значимые различия, более чем в 2 раза, в экспрессии miРНК зарегистрированы для 10 представителей: miR-26b, -150, -451a, -196b, -221, -29b, -23b, -128, let-7a и let-7c. Распределение относительных уровней экспрессии этих miРНК, включая медианное значение и межквартильный диапазон, представлено на рис. 1.

Таким образом, наибольший размах в изменении экспрессии среди исследованных miРНК выявлен для miR-150: ее уровень понижен более чем в 12 раз в клетках DLBCL в сравнении с неопухлой тканью. По-видимому, aberrантная экспрессия этой miРНК ассоциирована с развитием опухоли.

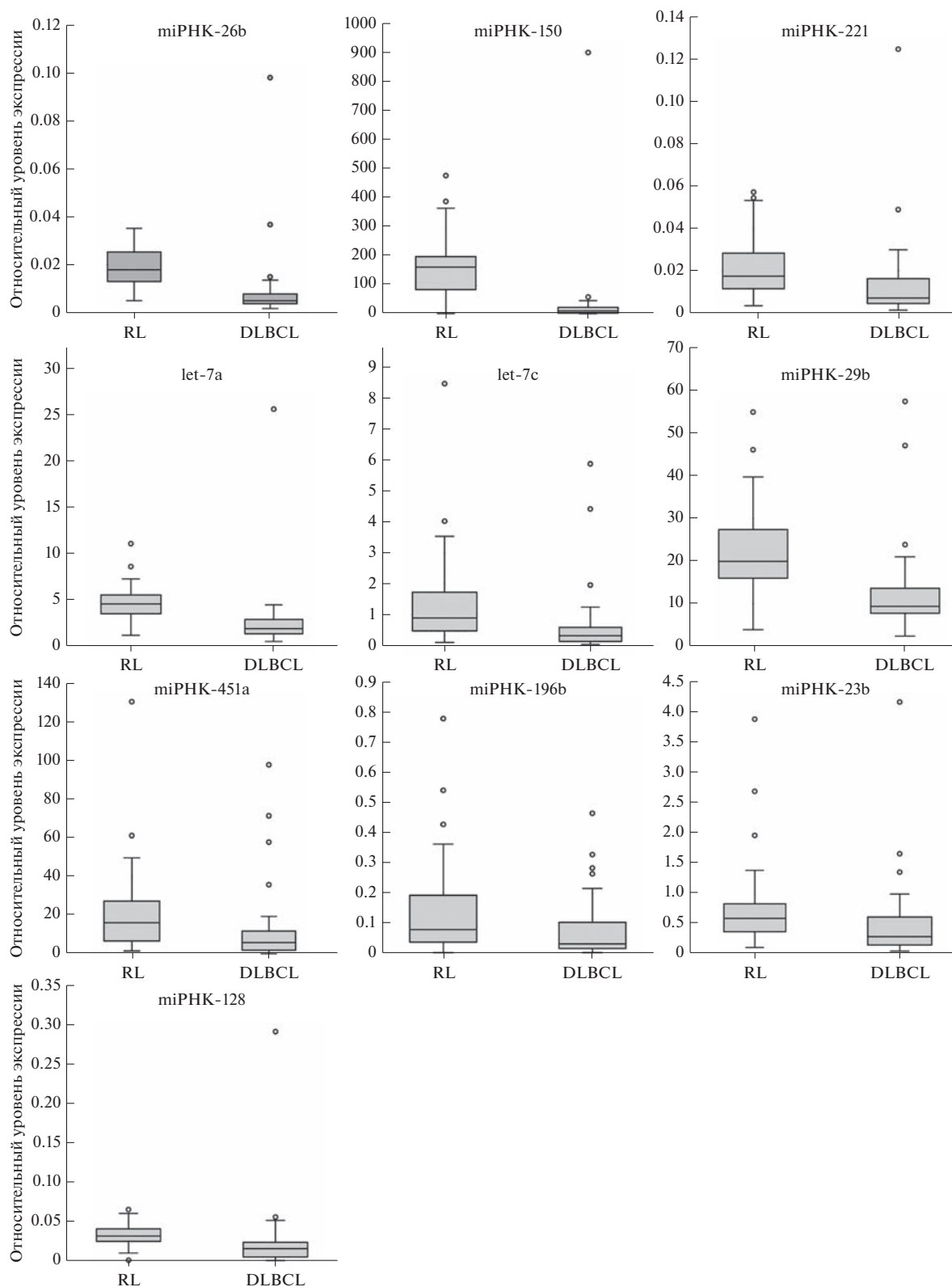


Рис. 1. Относительные уровни дифференциально экспрессируемых miPHK в образцах DLBCL и RL. Горизонтальная черная линия в пределах бокса – медиана, бокс – межквартильный диапазон, усы – диапазон без выбросов, выбросы обозначены кружками. Для всех приведенных miPHK различия в экспрессии были статистически значимыми ($p < 0.05$ по критерию Манна–Уитни).

Таблица 2. Процессы, в регуляции которых участвуют гены-мишени miR-150

Процесс	<i>p</i> -value	Гены-мишени
Активация В-клеток	0.030	<i>CCR6, EP300, FLT3, MMP14, STAT5B, TP53</i>
Дифференцировка В-клеток	0.019	<i>EP300, FLT3, MMP14, STAT5B, TP53</i>
Контроль клеточного цикла	0.049	<i>BIRC5, EP300, POLD3, PRKCA, TP53</i>
Гемопоез	0.018	<i>CCR6, CREB1, EP300, FLT3, MMP14, MYB, PRKCA, STAT1, STAT5B, TP53, VEGFA, ZEB1</i>

Биоинформационный анализ генов-мишеней miR-150

Понимание механизмов регуляции генов относится к одной из основных задач молекулярной биологии и биоинформатики. С помощью ресурса miRPathDB 2.0 (<https://mpd.bioinf.uni-sb.de/>) определены экспериментально подтвержденные пути, которые значительно обогащены мишенями miR-150 и нарушения в которых могут быть ассоциированы с DLBCL (табл. 2).

ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Лимфоциты относятся к важнейшим клеточным компонентам иммунной системы. Большинство лимфом происходит из В-лимфоцитов, подвергшихся неопластической трансформации. Подтипы В-клеточной лимфомы гетерогенны в отношении генетических и клинических характеристик [22]. Развитие В-клеток представляет собой сложный процесс, в котором участвуют как факторы транскрипции и цитокины, так и miРНК [23]. S. Koralov и соавт. [24] показали, что в развитии В-лимфоцитов участвует рибонуклеаза Dicer — ключевой участник процессинга miРНК, — из чего можно сделать вывод о том, что в регуляцию этого биологического процесса вовлечен сложный сигнальный каскад.

Мы провели анализ уровней экспрессии 29 miРНК в образцах DLBCL и RL и выявили наиболее значимые различия, более чем 2 раза, для 10 miРНК: miR-26b, -150, -451a, -196b, -221, -29b, -23b, -128, let-7a и let-7c. Согласно литературным данным, эти miРНК участвуют в регуляции нормального гемопоэза и, следовательно, их aberrантная экспрессия может быть вовлечена в развитие как миелоидных, так и лимфоидных опухолей [25–31]. Стоит отметить, что экспрессия miR-150 снижена более чем в 12 раз ($p = 3.6 \times 10^{-15}$) в образцах DLBCL в сравнении с тканями неопухолевых узловых образований. Таким образом, miR-150 может быть одним из ключевых участников злокачественной трансформации В-лимфоцитов.

На сегодняшний день опубликован ряд работ, отражающих роль miРНК в дифференцировке

В-клеток и развитии В-клеточных лимфом [32]. miR-150 контролирует дифференцировку В-клеток, воздействуя на фактор транскрипции с-Myb [33]. с-Myb, играет важную роль во время развития В-клеток, в поддержании их пролиферации, а также в контроле клеточного цикла гемопоэтических клеток [34]. Нарушения на каждом из этих биологических процессов способствуют прогрессии опухолей различного генеза, характеризуя с-Myb в качестве важного звена в развитии как солидных, так и гематологических опухолей. В исследованиях механизмов канцерогенеза лимфомы Беркитта выявлена ключевая роль генов *ZDHHC11* и *ZDHHC11B* в поддержании пути *MYC*—miR-150—*MYB*, обеспечивающего развитие опухоли [35]. Лимфома Беркитта, как и DLBCL, — это агрессивная В-клеточная лимфома. Показано, что гены *MYC*, *MYB* и *ZDHHC11* участвуют и в онкогенезе DLBCL [36]. Таким образом, не исключено, что в развитии DLBCL могут быть задействованы аналогичные регуляторные пути, в которые вовлечена miR-150. M. Wang и соавт. [37] идентифицировали miR-150 в качестве супрессора опухоли, снижающего пролиферацию клеток лимфомы Беркитта, определив в качестве мишеней этой miРНК с-Myb и Survivin. В клетках НК/Т-клеточной лимфомы повышенная экспрессия miR-150 коррелирует с усилением апоптоза и снижением пролиферации клеток, что подтверждает роль этой miРНК как супрессора; при этом гены *DKC1* и *AKT2* идентифицированы в качестве ее прямых мишеней [38]. Кроме того, miR-150 участвует в развитии Т-клеток, регулируя экспрессию *NOTCH3*, а также сигнальный путь *AKT3/Bim* [39, 40]. Снижение экспрессии miR-150 способствует мультиорганный инвазии и метастазированию Т-клеточной лимфомы за счет усиления экспрессии мишени — *CCR6* [41].

С фундаментальной точки зрения miR-150 можно считать регулятором развития лимфом. Кроме того, на сегодняшний день уже показан потенциал этой miРНК в качестве клинико-биологического маркера. Так, M. Mraz и соавт. [42] выявили, что у пациентов с хроническим лимфолейкозом повышенный уровень miR-150 в крови коррелирует с более длительной общей выживаемостью.

мостью. А Х. Wang и др. [43] показали, что пониженная экспрессия miR-150 коррелирует с более коротким временем общей выживаемости и периодом без прогрессирования у пациентов с первой желудочно-кишечной формой DLBCL.

Как отмечалось выше, miРНК — это важные регуляторы нормального гемопоэза. В частности, miR-150 регулирует терминальный эритропоэз у человека [44]. J. Lu и соавт. [45] отмечали, что miR-150 модулирует развитие мегакариоцитарно-эритроцитарных предшественников, при этом уровень ее экспрессии значительно снижается в условиях повышенной потребности в эритропоэзе. Очень часто у пациентов с DLBCL заболевание сопровождается анемией, что позволяет рассматривать miR-150 в качестве одного из потенциальных регуляторов и этого физиологического процесса. Так, С. Apple и соавт. [46] обнаружили дифференциальную экспрессию miR-150, miR-223, miR-15a и miR-24 в костном мозге у пациентов с травмой тазобедренного сустава, отмечая важную роль этих miРНК в эритропоэтической дисфункции, связанной с анемией.

Анализ уровней экспрессии miРНК может быть использован в качестве биомаркеров не только в тканях, но и в плазме крови. Н. Fayyad-Kazan и соавт. [47] считают, что уровни miR-150 и miR-342 в плазме можно рассматривать как перспективные биомаркеры в диагностике острого миелоидного лейкоза.

Таким образом, можно предположить, что miR-150 опосредует многие патофизиологические процессы, регулируя экспрессию своих генов-мишеней. На основании полученных данных miR-150 можно рассматривать в качестве перспективной терапевтической мишени с большим потенциалом в клинической практике.

Исследование выполнено при поддержке Российского фонда фундаментальных исследований (проект № 19-34-60024) и Российского научного фонда (проект № 20-14-00074). Исследование было финансировано в рамках государственного задания (проект № FWGZ-2021-0014).

Все процедуры, выполненные в данной работе, соответствуют этическим стандартам Институционального комитета по исследовательской этике и Хельсинкской декларации 1964 года и ее последующим изменениям или сопоставимым нормам этики. От всех пациентов получено информированное согласие.

Авторы заявляют об отсутствии конфликта интересов.

СПИСОК ЛИТЕРАТУРЫ

1. Alaggio R., Amador C., Anagnostopoulos I., Attygalle A.D., Araujo I.B.O., Berti E., Bhagat G., Borges A.M., Boyer D., Calaminici M., Chadburn A., Chan J.K.C.,

Cheuk W., Chng W.J., Choi J.K., Chuang S.S., Coup-land S.E., Czader M., Dave S.S., de Jong D., Du M.Q., Elenitoba-Johnson K.S., Ferry J., Geyer J., Gratzinger D., Guitart J., Gujral S., Harris M., Harrison C.J., Hartmann S., Hochhaus A., Jansen P.M., Karube K., Kempf W., Khoury J., Kimura H., Klapper W., Kovach A.E., Kumar S., Lazar A.J., Lazzi S., Leoncini L., Leung N., Leventaki V., Li X.Q., Lim M.S., Liu W.P., Louissaint A. Jr, Marcogliese A., Medeiros L.J., Michal M., Miranda R.N., Mitteldorf C., Montes-Moreno S., Morice W., Nardi V., Naresh K.N., Natkunam Y., Ng S.B., Oschlies I., Ott G., Parrens M., Pulitzer M., Rajkumar S.V., Rawstron A.C., Rech K., Rosenwald A., Said J., Sarkozy C., Sayed S., Saygin C., Schuh A., Sewell W., Siebert R., Sohani A.R., Lim Tooe R., Traverse-Glehen A., Vega F., Vergier B., Wechalekar A.D., Wood B., Xerri L., Xiao W. (2022) The 5th edition of the World Health Organization classification of haematolymphoid tumours: lymphoid neoplasms. *Leukemia*. 36, 1720–1748.

2. Wang L., Qin W., Huo Y.J., Li X., Shi Q., Rasko J.E.J., Janin A., Zhao W.L. (2020) Advances in targeted therapy for malignant lymphoma. *Signal Transduct. Target. Ther.* 5(1), 15.
3. Fuertes T., Ramiro A.R., de Yébenes V.G. (2020) miRNA-based therapies in B cell non-Hodgkin lymphoma. *Trends Immunol.* 41(10), 932–947.
4. Peng Y., Croce C.M. (2016) The role of microRNAs in human cancer. *Signal Transduct. Target. Ther.* 1, 15004.
5. Su Y., Sun B., Lin X., Zhao X., Ji W., He M., Qian H., Song X., Yang J., Wang J., Chen J. (2016) Therapeutic strategy with artificially-designed i-lncRNA targeting multiple oncogenic microRNAs exhibits effective anti-tumor activity in diffuse large B-cell lymphoma. *Oncotarget*. 7(31), 49143–49155.
6. Due H., Schönherz A.A., Ryø L., Primo M.N., Jespersen D.S., Thomsen E.A., Roug A.S., Xiao M., Tan X., Pang Y., Young K.H., Bøgsted M., Mikkelsen J.G., Dybkær K. (2019) MicroRNA-155 controls vincristine sensitivity and predicts superior clinical outcome in diffuse large B-cell lymphoma. *Blood Adv.* 3(7), 1185–1196.
7. Yang J.M., Jang J.Y., Jeon Y.K., Paik J.H. (2018) Clinicopathologic implication of microRNA-197 in diffuse large B cell lymphoma. *J. Transl. Med.* 16(1), 162.
8. Huang F., Jin Y., Wei Y. (2016) MicroRNA-187 induces diffuse large B-cell lymphoma cell apoptosis via targeting BCL6. *Oncol. Lett.* 11(4), 2845–2850.
9. Fan Q., Meng X., Liang H., Zhang H., Liu X., Li L., Li W., Sun W., Zhang H., Zen K., Zhang C.Y., Zhou Z., Chen X., Ba Y. (2016) miR-10a inhibits cell proliferation and promotes cell apoptosis by targeting BCL6 in diffuse large B-cell lymphoma. *Protein Cell*. 7(12), 899–912.
10. Farina F.M., Inguscio A., Kunderfranco P., Cortesi A., Elia L., Quintavalle M. (2017) MicroRNA-26a/cyclin-dependent kinase 5 axis controls proliferation, apoptosis and *in vivo* tumor growth of diffuse large B-cell lymphoma cell lines. *Cell Death Dis.* 8(6), e2890.
11. Yuan J., Zhang Q., Wu S., Yan S., Zhao R., Sun Y., Tian X., Zhou K. (2021) MiRNA-223-3p affects mant-

- le cell lymphoma development by regulating the CHUK/NF- κ B2 signaling pathway. *OncoTargets Ther.* **14**, 1553–1564.
12. Tian Y.Y., Jia C.M., Li Y., Wang Y., Jiang L., Liu A.C. (2016) Restoration of microRNA-373 suppresses growth of human T-cell lymphoma cells by repressing CCND1. *Eur. Rev. Med. Pharmacol. Sci.* **20**(21), 4435–4444.
 13. Morales-Martinez M., Vega G.G., Neri N., Nambo M.J., Alvarado I., Cuadra I., Duran-Padilla M.A., Huerta-Yepez S., Vega M.I. (2020) MicroRNA-7 regulates migration and chemoresistance in non-Hodgkin lymphoma cells through regulation of KLF4 and YY1. *Front. Oncol.* **10**, 588893.
 14. Wu S.J., Chen J., Wu B., Wang Y.J., Guo K.Y. (2018) MicroRNA-150 enhances radiosensitivity by inhibiting the AKT pathway in NK/T cell lymphoma. *J. Exp. Clin. Cancer Res.* **37**, 18.
 15. Musilova K., Devan J., Cerna K., Seda V., Pavlasova G., Sharma S., Oppelt J., Pytlik R., Prochazka V., Prouzova Z., Trbusek M., Zlamalikova L., Liskova K., Kruzova L., Jarosova M., Mareckova A., Kornauth C., Simonitsch-Klupp I., Schiefer A.I., Merkel O., Mocikova H., Burda P., Machova Polakova K., Kren L., Mayer J., Zent C.S., Trneny M., Evans A.G., Janikova A., Mraz M. (2018) *miR-150* downregulation contributes to the high-grade transformation of follicular lymphoma by upregulating FOXP1 levels. *Blood.* **132**(22), 2389–2400.
 16. Witten L., Slack F.J. (2020) *miR-155* as a novel clinical target for hematological malignancies. *Carcinogenesis.* **41**, 2–7.
 17. Chakraborty C., Sharma A.R., Sharma G., Lee S.S. (2020) Therapeutic advances of miRNAs: a preclinical and clinical update. *J. Adv. Res.* **28**, 127–138.
 18. Sebestyén E., Nagy Á., Marosvári D., Rajnai H., Kajtár B., Deák B., Matolcsy A., Brandner S., Storhoff J., Chen N., Bagó A.G., Bödör C., Reiniger L. (2022) Distinct miRNA expression signatures of primary and secondary central nervous system lymphomas. *J. Mol. Diagn.* **24**(3), 224–240.
 19. Costé É., Rouleux-Bonnin F. (2020) The crucial choice of reference genes: identification of *miR-191-5p* for normalization of miRNAs expression in bone marrow mesenchymal stromal cell and HS27a/HS5 cell lines. *Sci. Rep.* **10**(1), 17728.
 20. de Leeuw D.C., van den Ancker W., Denkers F., de Menezes R.X., Westers T.M., Ossenkoppele G.J., van de Loosdrecht A.A., Smit L. (2013) MicroRNA profiling can classify acute leukemias of ambiguous lineage as either acute myeloid leukemia or acute lymphoid leukemia. *Clin. Cancer Res.* **19**, 2187–2196.
 21. Ragni E., Colombini A., De Luca P., Libonati F., Viganò M., Perucca Orfei C., Zagra L., de Girolamo L. (2021) *miR-103a-3p* and *miR-22-5p* are reliable reference genes in extracellular vesicles from cartilage, adipose tissue, and bone marrow cells. *Front. Bioeng. Biotechnol.* **9**, 632440.
 22. Esmeray E., Küçük C. (2020) Genetic alterations in B cell lymphoma subtypes as potential biomarkers for noninvasive diagnosis, prognosis, therapy, and disease monitoring. *Turk. J. Biol.* **44**(1), 1–14.
 23. Souza O.F., Popi A.F. (2022) Role of microRNAs in B-cell compartment: development, proliferation and hematological diseases. *Biomedicines.* **10**(8), 2004.
 24. Koralov S.B., Muljo S.A., Galler G.R., Krek A., Chakraborty T., Kanellopoulou C., Jensen K., Cobb B.S., Merkenschlager M., Rajewsky N., Rajewsky K. (2008) Dicer ablation affects antibody diversity and cell survival in the B lymphocyte lineage. *Cell.* **132**(5), 860–874.
 25. Веряскина Ю.А., Титов С.Е., Ковынев И.Б., Федорова С.С., Поспелова Т.И., Жимулёв И.Ф. (2021) микроРНК при миелодиспластическом синдроме. *Acta Naturae.* **13**(2), 4–15.
 26. Emmrich S., Rasche M., Schöning J., Reimer C., Keihani S., Maroz A., Xie Y., Li Z., Schambach A., Reinhardt D., Klusmann J.H. (2014) *miR-99a/100~125b* tricistrons regulate hematopoietic stem and progenitor cell homeostasis by shifting the balance between TGF β and Wnt signaling. *Genes Dev.* **28**(8), 858–874.
 27. Velu C.S., Baktula A.M., Grimes H.L. (2009) Gfi1 regulates *miR-21* and *miR-196b* to control myelopoiesis. *Blood.* **113**(19), 4720–4728.
 28. Felli N., Fontana L., Pelosi E., Botta R., Bonci D., Facchiano F., Liuzzi F., Lulli V., Morsilli O., Santoro S., Valtieri M., Calin G.A., Liu C.G., Sorrentino A., Croce C.M., Peschle C. (2005) MicroRNAs 221 and 222 inhibit normal erythropoiesis and erythroleukemic cell growth via kit receptor down-modulation. *Proc. Natl. Acad. Sci. USA.* **102**(50), 18081–18086.
 29. Lee J.Y., Kim M., Heo H.R., Ha K.S., Han E.T., Park W.S., Yang S.R., Hong S.H. (2018) Inhibition of microRNA-221 and 222 enhances hematopoietic differentiation from human pluripotent stem cells via c-KIT upregulation. *Mol. Cells.* **41**(11), 971–978.
 30. Kurkewich J.L., Boucher A., Klopstein N., Baskar R., Kapur R., Dahl R. (2017) The *Mirn23a* and *Mirn23b* microRNA clusters are necessary for proper hematopoietic progenitor cell production and differentiation. *Exp. Hematol.* **59**, 14–29.
 31. Pelosi A., Careccia S., Lulli V., Romania P., Marziali G., Testa U., Lavorgna S., Lo-Coco F., Petti M.C., Calabretta B., Levrero M., Piaggio G., Rizzo M.G. (2013) *miRNA let-7c* promotes granulocytic differentiation in acute myeloid leukemia. *Oncogene.* **32**(31), 3648–3654.
 32. Zheng B., Xi Z., Liu R., Yin W., Sui Z., Ren B., Miller H., Gong Q., Liu C. (2018) The function of microRNAs in B-cell development, lymphoma, and their potential in clinical practice. *Front. Immunol.* **9**, 936.
 33. Xiao C., Calado D.P., Galler G., Thai T.H., Patterson H.C., Wang J., Rajewsky N., Bender T.P., Rajewsky K. (2007) *miR-150* controls B cell differentiation by targeting the transcription factor c-Myb. *Cell.* **131**(1), 146–159.
 34. Nakata Y., Shetzline S., Sakashita C., Kalota A., Rallapalli R., Rudnick S.I., Zhang Y., Emerson S.G., Gewirtz A.M. (2007) c-Myb contributes to G2/M cell cycle transition in human hematopoietic cells by direct

- regulation of cyclin B1 expression. *Mol. Cell Biol.* **27**(6), 2048–2058.
35. Dzikiewicz-Krawczyk A., Kok K., Slezak-Prochazka I., Robertus J.L., Bruining J., Tayari M.M., Rutgers B., de Jong D., Koerts J., Seitz A., Li J., Tillema B., Guikema J.E., Nolte I.M., Diepstra A., Visser L., Kluiver J., van den Berg A. (2017) ZDHHC11 and ZDHHC11B are critical novel components of the oncogenic MYC-miR-150-MYB network in Burkitt lymphoma. *Leukemia*. **31**(6), 1470–1473.
 36. Ziel-Swier L.J.Y.M., Liu Y., Seitz A., de Jong D., Koerts J., Rutgers B., Veenstra R., Razak F.R.A., Dzikiewicz-Krawczyk A., van den Berg A., Kluiver J. (2022) The role of the MYC/miR-150/MYB/ZDHHC11 network in Hodgkin lymphoma and diffuse large B-cell lymphoma. *Genes (Basel)*. **13**(2), 227.
 37. Wang M., Yang W., Li M., Li Y. (2014) Low expression of miR-150 in pediatric intestinal Burkitt lymphoma. *Exp. Mol. Pathol.* **96**(2), 261–266.
 38. Watanabe A., Tagawa H., Yamashita J., Teshima K., Nara M., Iwamoto K., Kume M., Kameoka Y., Takahashi N., Nakagawa T., Shimizu N., Sawada K. (2011) The role of microRNA-150 as a tumor suppressor in malignant lymphoma. *Leukemia*. **25**(8), 1324–1334.
 39. Ghisi M., Corradin A., Basso K., Frasson C., Serafin V., Mukherjee S., Mussolin L., Ruggero K., Bonanno L., Guffanti A., De Bellis G., Gerosa G., Stellin G., D'Agostino D.M., Basso G., Bronte V., Indraccolo S., Amadori A., Zanovello P. (2011) Modulation of microRNA expression in human T-cell development: targeting of NOTCH3 by miR-150. *Blood*. **117**(26), 7053–7062.
 40. Sang W., Sun C., Zhang C., Zhang D., Wang Y., Xu L., Zhang Z., Wei X., Pan B., Yan D., Zhu F., Yan Z., Cao J., Loughran T.P. Jr., Xu K. (2016) MicroRNA-150 negatively regulates the function of CD4⁺ T cells through AKT3/Bim signaling pathway. *Cell Immunol.* **306–307**, 35–40.
 41. Ito M., Teshima K., Ikeda S., Kitadate A., Watanabe A., Nara M., Yamashita J., Ohshima K., Sawada K., Tagawa H. (2014) MicroRNA-150 inhibits tumor invasion and metastasis by targeting the chemokine receptor CCR6, in advanced cutaneous T-cell lymphoma. *Blood*. **123**(10), 1499–1511.
 42. Mraz M., Chen L., Rassenti L.Z., Ghia E.M., Li H., Jepsen K., Smith E.N., Messer K., Frazer K.A., Kipps T.J. (2014) miR-150 influences B-cell receptor signaling in chronic lymphocytic leukemia by regulating expression of GAB1 and FOXP1. *Blood*. **124**(1), 84–95.
 43. Wang X., Kan Y., Chen L., Ge P., Ding T., Zhai Q., Yu Y., Wang X., Zhao Z., Yang H., Liu X., Li L., Qiu L., Qian Z., Zhang H., Wang Y., Zhao H. (2020) miR-150 is a negative independent prognostic biomarker for primary gastrointestinal diffuse large B-cell lymphoma. *Oncol. Lett.* **19**(5), 3487–3494.
 44. Sun Z., Wang Y., Han X., Zhao X., Peng Y., Li Y., Peng M., Song J., Wu K., Sun S., Zhou W., Qi B., Zhou C., Chen H., An X., Liu J. (2015) miR-150 inhibits terminal erythroid proliferation and differentiation. *Oncotarget*. **6**(40), 43033–43047.
 45. Lu J., Guo S., Ebert B.L., Zhang H., Peng X., Bosco J., Pretz J., Schlanger R., Wang J.Y., Mak R.H., Dombkowski D.M., Preffer F.I., Scadden D.T., Golub T.R. (2008) MicroRNA-mediated control of cell fate in megakaryocyte-erythrocyte progenitors. *Dev. Cell*. **14**(6), 843–853.
 46. Apple C.G., Miller E.S., Kannan K.B., Stortz J.A., Loftus T.J., Lopez M.C., Parvataneni H.K., Patrick M., Hagen J.E., Baker H.V., Efron P.A., Mohr A.M. (2021) The role of bone marrow microRNA (miR) in erythropoietic dysfunction after severe trauma. *Surgery*. **169**(5), 1206–1212.
 47. Fayyad-Kazan H., Bitar N., Najjar M., Lewalle P., Fayyad-Kazan M., Badran R., Hamade E., Daher A., Hussein N., ElDirani R., Berri F., Vanhamme L., Burny A., Martiat P., Rouas R., Badran B. (2013) Circulating miR-150 and miR-342 in plasma are novel potential biomarkers for acute myeloid leukemia. *J. Transl. Med.* **11**, 31.

The Profile of microRNA Expression in Diffuse Large B-Cell Lymphoma

Yu. A. Veryaskina^{1, 2, *}, S. E. Titov^{1, 3}, I. B. Kovynev⁴, S. S. Fyodorova⁴, Ya. Yu. Shebunyaeva⁴, O. V. Antonenko¹, T. I. Pospelova⁴, and I. F. Zhimulev¹

¹*Institute of Molecular and Cellular Biology, Siberian Branch, Russian Academy of Sciences, Novosibirsk, 630090 Russia*

²*Institute of Cytology and Genetics, Siberian Branch, Russian Academy of Sciences, Novosibirsk, 630090 Russia*

³*JSC Vector-Best, Novosibirsk, 630117 Russia*

⁴*Novosibirsk State Medical University, Ministry of Health of the Russian Federation, Novosibirsk, 630091 Russia*

*e-mail: microrna@inbox.ru

Non-Hodgkin's lymphoma (NHL) is a heterogeneous group of cancers characterized by different pathogenesis and prognosis. The main methods for treating NHL are chemotherapy, immunochemotherapy, and radiation therapy; however, most of these cancers are known to be chemoresistant or return rapidly after the short chemotherapy-induced remission. Therefore, searching for alternative cytoreductive therapy options is quite

relevant today. Aberrant microRNA (miRNA) expression is one of the mechanisms responsible for the emergence and progression of lymphoid malignancies. This study was aimed at identifying the miRNA expression profile in diagnostic biopsy specimens harvested from the lymph nodes affected by diffuse large B-cell lymphoma (DLBCL) and identifying miRNA markers, which can potentially be used to design a novel type of targeted anticancer drugs that would allow one to achieve maximum therapy personalization and increase its efficacy. The key study objects were histological specimens harvested from the lymph nodes by excisional diagnostic biopsy and treated using the conventional histomorphological formalin fixation methods. The study group consisted of patients with DLBCL ($n = 52$). The biopsy specimens harvested from patients with reactive lymphadenopathy (RL) ($n = 40$) constituted the control group. The miR-150 expression level was reduced over 12-fold ($p = 3.6 \times 10^{-15}$) compared to that in the tissues of non-cancerous nodular masses. Bioinformatic analysis revealed that miR-150 is involved in regulation of hematopoiesis and lymphopoiesis. The findings obtained in this study allow considering miR-150 a promising therapeutic target having a great potential for clinical applications.

Keywords: microRNA, hematopoiesis, lymphoma, non-Hodgkin lymphoma, B-cell non-Hodgkin lymphoma

УДК 576.52;577.352

ВЛИЯНИЕ РЕЦЕПТОРА- γ , АКТИВИРУЕМОГО ПРОЛИФЕРАТОРОМ ПЕРОКСИСОМ, НА МОДУЛЯЦИЮ СИНТЕЗА АНГИОПОЭТИН-ПОДОБНОГО БЕЛКА 4 В КЛЕТКАХ Caco-2, ОБРАБОТАННЫХ *Clostridium butyricum*¹

© 2023 г. X. Zhao^{a, *}, H. S. Huang^{b, *}, S. R. Shi^{c, **}

^aCollege of Agriculture and Forestry Science, Linyi University, Linyi, Shandong, 276000 China

^bShandong Longda Biotechnology Co., Ltd., Linyi, Shandong, 276400 China

^cPoultry Institute, Chinese Academy of Agricultural Sciences, Yangzhou, Jiangsu, 271018 China

*e-mail: kity850814@163.com

**e-mail: ssr236@163.com

Поступила в редакцию 01.04.2022 г.

После доработки 23.06.2022 г.

Принята к публикации 26.06.2022 г.

Ангиопоэтин-подобный белок 4 (ANGPTL4) считается одним из важных циркулирующих медиаторов, связывающих кишечные микроорганизмы с метаболизмом липидов хозяйского организма. В представленной работе изучено влияние рецептора- γ , активируемого пролифератором пероксисом (PPAR γ) на модуляцию синтеза ANGPTL4 в клетках Caco-2, обработанных *Clostridium butyricum*. Жизнеспособность клеток Caco-2 и экспрессию в них PPAR γ и ANGPTL4 определяли после сокультивирования с клетками *C. butyricum* в концентрации 1×10^6 , 1×10^7 и 1×10^8 КОЕ/мл. Показано повышение жизнеспособности клеток в присутствии *C. butyricum*. Кроме того, экспрессия и секреция PPAR γ и ANGPTL4 в клетках Caco-2 значительно возрастали в присутствии 1×10^7 и 1×10^8 КОЕ/мл *C. butyricum*. Более того, влияние PPAR γ на модуляцию синтеза ANGPTL4 в клетках Caco-2, обработанных 1×10^8 КОЕ/мл *C. butyricum*, проверено также на модели активации/ингибирования PPAR γ в клетках Caco-2 и с использованием метода иммунопреципитации хроматина (ChIP). Обнаружено, что *C. butyricum* стимулирует взаимодействие PPAR γ с сайтом связывания PPAR, локализованным перед сайтом старта транскрипции гена *angptl4* (chr19: 8362157–8362357) в клетках Caco-2. Однако PPAR γ это не единственный путь стимуляции продукции ANGPTL4 под действием *C. butyricum*. В целом, PPAR γ участвует в регуляции синтеза ANGPTL4 в клетках Caco-2, обработанных *C. butyricum*.

Ключевые слова: *Clostridium butyricum*, рецептор- γ , активируемый пролифератором пероксисом, ангиопоэтин-подобный белок 4, клетки Caco-2

DOI: 10.31857/S0026898423030217, EDN: CIBNGM

Effects of Peroxisome Proliferator-Activated Receptor γ on Modulating Angiopoietin-Like Protein 4 Synthesis in Caco-2 Cells Exposed to *Clostridium butyricum*

X. Zhao^{1, *}, H. S. Huang^{2, *}, and S. R. Shi^{3, **}

¹ College of Agriculture and Forestry Science, Linyi University, Linyi, Shandong, 276000 China

² Shandong Longda Biotechnology Co., Ltd., Linyi, Shandong, 276400 China

³ Poultry Institute, Chinese Academy of Agricultural Sciences, Yangzhou, Jiangsu, 271018 China

*e-mail: kity850814@163.com

**e-mail: ssr236@163.com

Angiopoietin-like protein 4 (ANGPTL4) is considered to be one of the important circulating mediators linking intestinal microorganisms and host lipid metabolism. The objective of this study was to assess the effects

¹ Статья представлена авторами на английском языке.

of peroxisome proliferator-activated receptor γ (PPAR γ) on modulating ANGPTL4 synthesis in Caco-2 cells exposed to *Clostridium butyricum*. The viability of Caco-2 cells and the expression of PPAR γ and ANGPTL4 in Caco-2 cells were detected after the Caco-2 cells were co-cultured with *C. butyricum* at the concentration of 1×10^6 , 1×10^7 and 1×10^8 CFU/mL. The results showed that cell viability was enhanced by *C. butyricum*. Besides, PPAR γ and ANGPTL4 expression and secretion in Caco-2 cells was significantly increased by 1×10^7 and 1×10^8 CFU/mL of *C. butyricum*. Furthermore, the effects of PPAR γ on modulating ANGPTL4 synthesis in Caco-2 cells regulated by 1×10^8 CFU/mL of *C. butyricum* was also be expounded in PPAR γ activation/inhibition model based on Caco-2 cells and via ChIP technique. It was found that *C. butyricum* promoted the binding of PPAR γ to the PPAR binding site (chr19: 8362157-8362357, located upstream of the transcriptional start site of *angptl4*) of the *angptl4* gene in Caco-2 cells. However, the PPAR γ was not the only way for *C. butyricum* to stimulate ANGPTL4 production. Taken together, PPAR γ played a role in the regulation of ANGPTL4 synthesis by *C. butyricum* in Caco-2 cells.

Keywords: *Clostridium butyricum*, peroxisome proliferator activated receptor γ , angiopoietin-like protein 4, Caco-2 cell

УДК 575.113.5:575.117.2

EBF1 ПОВЫШАЕТ ЧУВСТВИТЕЛЬНОСТЬ КЛЕТОК РАКА ШЕЙКИ МАТКИ К ЦИСПЛАТИНУ ПОСРЕДСТВОМ АКТИВАЦИИ ТРАНСКРИПЦИИ FBN1¹

© 2023 г. N. N. Shen^a, J. H. Lin^b, P. P. Liu^{b, *}

^aDepartment of Pharmacy, Ganzhou Women and Children's Health Care Hospital, Ganzhou, Jiangxi, 341000 P.R. China

^bDepartment of Pharmacy, the First Affiliated Hospital of Gannan Medical University, Ganzhou Jiangxi, 341000 P.R. China

*e-mail: liupeipei198603@163.com

Поступила в редакцию 18.02.2022 г.

После доработки 07.06.2022 г.

Принята к публикации 05.07.2022 г.

Цисплатин широко используется в химиотерапии рака шейки матки, четвертого по частоте злокачественного заболевания у женщин в мире. Однако у некоторых пациенток развивается резистентность к препарату, что ведет к неэффективности химиотерапии, рецидиву опухоли и плохому прогнозу. Поэтому разработка стратегий идентификации регуляторных механизмов, лежащих в основе развития рака шейки матки, и способов повышения чувствительности опухолевых клеток к цисплатину поможет повысить выживаемость пациентов. Нами изучены механизмы EBF1-зависимой регуляции FBN1, которая повышает чувствительность клеток рака шейки матки к химиотерапии. В тканях рака шейки матки, резистентных или чувствительных к химиотерапии, а также в линиях клеток, резистентных или чувствительных к цисплатину (SiHa и SiHa-DDP), определена экспрессия EBF1 и FBN1. Клетки SiHa-DDP трансдуцировали лентивирусами, кодирующими EBF1 или FBN1, для оценки влияния этих двух белков на жизнеспособность клеток, экспрессию *MDR1* и *MRP1*, и агрессивность клеток. Более того, предсказано и показано взаимодействие между EBF1 и FBN1. Наконец, для дальнейшей верификации механизма EBF1/FBN1-зависимой регуляции чувствительности клеток рака шейки матки к цисплатину клетки SiHa-DDP, трансдуцированные лентивирусами, несущими ген *EBF1* и экспрессирующими шпилечную РНК, специфичную к FBN1, трансплантировали мышам и получили модель рака шейки матки. Показано снижение экспрессии EBF1 и FBN1 в тканях и клетках рака шейки матки, в частности в устойчивых к химиотерапии. Трансдукция клеток SiHa-DDP лентивирусами, кодирующими EBF1 или FBN1, ведет к снижению жизнеспособности, IC₅₀, пролиферативной активности, способности к образованию колоний, агрессивности и повышению уровня апоптоза клеток. Показано, что EBF1 активирует транскрипцию FBN1, связываясь с промоторной областью гена *FBN1*. Кроме того, установлено, что сайленсинг FBN1 отменяет стимулирующее действие сверхэкспрессии EBF1 на чувствительность клеток рака шейки матки к химиотерапии *in vivo*. Таким образом, EBF1 стимулирует чувствительность клеток рака шейки матки к химиотерапии путем активации транскрипции *FBN1*.

Ключевые слова: рак шейки матки, фактор 1 ранних В-клеток, фибриллин-1, факторы транскрипции, цисплатин, химиотерапия

DOI: 10.31857/S0026898423030102, EDN: CHKZXJ

EBF1 Promotes the Sensitivity of Cervical Cancer Cells to Cisplatin via Activating FBN1 Transcription

N. N. Shen¹, J. H. Lin², and P. P. Liu^{2, *}

¹Department of Pharmacy, Ganzhou Women and Children's Health Care Hospital, Ganzhou, Jiangxi, 341000 P.R. China

²Department of Pharmacy, the First Affiliated Hospital of Gannan Medical University, Ganzhou Jiangxi, 341000 P.R. China

*e-mail: liupeipei198603@163.com

Cisplatin (DDP) is widely used in the chemotherapy of cervical cancer (CC), the fourth most common female malignancy worldwide. However, some patients progress to chemotherapy resistance, which leads to chemo-

¹ Статья представлена авторами на английском языке.

therapy failure, tumor recurrence, and poor prognosis. Therefore, strategies to identify the regulatory mechanisms underlying CC development and increase tumor sensitivity to DDP will help improve patient survival. This research was designed to ascertain the mechanism of EBF1-dependent regulation of FBN1 which promotes chemosensitivity of CC cells. The expression of EBF1 and FBN1 was measured in CC tissues resistant or sensitive to chemotherapy and in DDP-sensitive or -resistant cells (SiHa and SiHa-DDP cells). SiHa-DDP cells were transduced with lentiviruses encoding EBF1 or FBN1 to evaluate the influence of these two proteins on cell viability, expression of MDR1 and MRP1, and cell aggressiveness. Moreover, the interaction between EBF1 and FBN1 was predicted and demonstrated. Finally, to further verify the EBF1/FBN1-dependent mechanism of DDP sensitivity regulation in CC cells a xenograft mouse model of CC was established using SiHa-DDP cells transduced with lentiviruses carrying EBF1 gene and shRNA directed to FBN1. EBF1 and FBN1 showed decreased expression in CC tissues and cells, particularly in those resistant to chemotherapy. Transduction of SiHa-DDP cells with lentiviruses encoding EBF1 or FBN1 lead to decreased viability, IC₅₀, proliferation capacity, colony formation ability, aggressiveness, and increased cell apoptosis. We have shown that EBF1 activates FBN1 transcription by binding to FBN1 promoter region. Additionally, it was revealed that FBN1 silencing reversed the promoting effect of EBF1 overexpression on chemosensitivity of CC cells *in vivo*. EBF1 facilitated chemosensitivity in CC cells by activating FBN1 transcription.

Keywords: cervical cancer, early B cell factor 1, fibrillin-1, transcription factor, chemosensitivity, cisplatin

УДК 577.22;578.282

ПОЛУЧЕНИЕ СЕРИИ НЕЙТРАЛИЗУЮЩИХ НАНОТЕЛ К S-БЕЛКУ SARS-CoV-2

© 2023 г. В. А. Жучков^а, С. В. Иванов^а, Ю. Е. Кравченко^а, С. П. Чумаков^а, *

^аИнститут биоорганической химии им. академиков М. М. Шемякина и Ю. А. Овчинникова Российской академии наук,
Москва, 117997 Россия

*e-mail: hathkul@gmail.com

Поступила в редакцию 03.06.2022 г.

После доработки 06.10.2022 г.

Принята к публикации 06.10.2022 г.

Противодействие распространению новых респираторных инфекций и снижение ущерба, наносимого ими обществу, требует разработки способов быстрого создания средств направленной терапии, таких как моноклональные антитела и подобные им структуры. Нанотела — антигенраспознающие фрагменты особых антител Верблюдовых, состоящие из тяжелых цепей, — обладают рядом характеристик, которые делают их применение наиболее удобным. Высокие темпы распространения пандемии COVID-19, вызванной новым коронавирусом — SARS-CoV-2, — указывают на то, что ключевым фактором в разработке средств терапии является быстрота получения высокоэффективных блокирующих антител, а также разнообразие эпитопов, с которыми эти антитела связываются. Мы провели оптимизацию процесса селекции нанотел из генетического материала животных семейства Верблюдовых и получили панель вариантов, обладающих сродством к белку шипа (S-белку) SARS-CoV-2 в нижнем наномолярном и пикомолярном диапазоне и высокой специфичностью связывания. В экспериментах *in vitro* и *in vivo* были отобраны нанотела, блокирующие взаимодействие между S-белком и клеточным рецептором SARS-CoV-2 — ангиотензинконвертирующим ферментом-2 (ACE2). Эпитопы, с которыми связываются нанотела, располагаются в рецепторсвязывающем домене (RBD) S-белка и слабо перекрываются, поэтому смесь нанотел может сохранять потенциальную терапевтическую эффективность и в отношении новых вариантов S-белка. Смесь нанотел, обладающих компактным размером и высокой стабильностью, может стать основой для создания терапевтического препарата, доставляемого к пораженному органу в виде аэрозоля.

Ключевые слова: SARS-CoV-2, нанотела, белок шипа, ангиотензинконвертирующий фермент-2, нейтрализующие антитела, моноклональные антитела, иммунотерапия

DOI: 10.31857/S0026898423030230, **EDN:** CIBRDB

Пандемия COVID19, вызванная вирусом SARS-CoV-2, наглядно продемонстрировала, что для эффективного противодействия возникающим новым биологическим угрозам необходимы отработанные пути создания новых вакцин и специфических терапевтических агентов [1]. В контексте противодействия вирусным инфекциям на роль последних лучше всего подходят различные варианты специфических антител [2]. Однако, в отличие от быстрой сертификации и массового применения новых вакцин, терапевтические моноклональные антитела к белкам SARS-CoV-2 не были массово внедрены в медицинскую практику по ряду причин. Среди них — относительно высокая себестоимость производства и результаты исследований, в которых показано, что для достижения терапевтического эффекта такие препараты необходимо применять на ранних стадиях развития заболевания [3]. Дополнительная про-

блема, препятствующая внедрению противовирусных моноклональных антител, — возникновение новых, устойчивых к ним вариантов SARS-CoV-2 с мутациями в поверхностном белке шипа (S), на который и вырабатываются нейтрализующие антитела, в том числе моноклональные [4, 5].

Большинство перечисленных ограничений основано на особенностях структуры моноклональных антител: антигенная специфичность в них определяется участками двух белковых цепей, что усложняет получение подходящих по характеристикам вариантов. Для их производства подходят только эукариотические системы, что сильно влияет на стоимость производства и, следовательно, терапии, а поэтому не подходит для широкого применения. Кроме того, крупный размер молекул иммуноглобулинов ограничивает тканевую проницаемость, что не позволяет эффективно использовать их на ранних стадиях или в каче-

стве профилактического средства. Периодическое появление новых вариантов SARS-CoV-2 дополнительно усугубляет эти ограничения и снижает эффективность терапии с использованием моноклональных антител [4].

Как минимум, часть перечисленных ограничений можно преодолеть путем применения антител альтернативной структуры, например фрагментов особых антител, характерных для представителей семейства Верблюдовых и состоящих из гомодимера тяжелых цепей, — нанотел [6]. За счет того, что антигенная специфичность таких антител определяется последовательностями, расположенными на одной полипептидной цепи, селекция высокоэффективных вариантов из библиотек технически проще и производится быстрее. Кроме того, нанотела могут быть синтезированы в прокариотических экспрессионных системах и, в отличие от фрагментов моноклональных антител, для повышения их стабильности и растворимости не требуется дополнительная модификация [7]. Малый размер нанотел обеспечивает им лучшую тканевую проницаемость, а высокая стабильность структуры облегчает их использование не только как препарата для инъекций, но и для местного или аэрозольного применения.

Скорость и эффективность отбора подходящих последовательностей однодоменных антител определяется рядом параметров: иммуногенностью целевого белка в организме животного, выбранной методикой проведения селекции и системами тестирования кандидатных вариантов. Поверхностный белок шипа (spike, S) SARS-CoV-2 распознает клеточный рецептор — ангиотензин-конвертирующий фермент-2 (ACE2), — связывается с ним и тем самым обеспечивает проникновение вируса в клетку. Важно заметить, что этот гликопротеин SARS-CoV-2 обладает высокой гомологией с соответствующим белком другого вируса семейства *Coronaviridae*, — коронавируса ближневосточного респираторного синдрома, MERS-CoV [8]. Этот вирус попал в человеческую популяцию от животных семейства Верблюдовые, преодолев межвидовой барьер. На основании этого логично предположить высокую иммуногенность белков этого типа в организмах лам и альпак.

Увеличить долю последовательностей, обладающих терапевтическим потенциалом в итоговой библиотеке антител можно следующими способами: 1) за счет обогащения исходных В-клеточных популяций [9] фракциями, несущими распознающие целевой белок рецепторами; 2) применением на стадии отбора фрагмента S-белка, ответственного за взаимодействие с клеточным рецептором, — рецепторсвязывающего домена (RBD); 3) использованием рекомбинантного лиганда для выделения последовательностей нанотел, конкурирующих с ним за общий сайт связывания. Ре-

зультатом таких действий должно стать получение высокообогащенной библиотеки, содержащей множество вариантов последовательностей нанотел, обладающих высоким сродством к мишени, специфичностью взаимодействия и способностью блокировать связывание S-белка и ACE2. Антитела, отобранные из такой библиотеки, могут применяться либо в виде монотерапевтического средства, нацеленного на конкретный вариант S-белка, либо в виде коктейля, обладающего более широким действием за счет множества распознаваемых эпитопов.

Целью данной работы было проведение оптимизации условий получения библиотек последовательностей нанотел и последующей селекции вариантов, обладающих потенциальной терапевтической эффективностью, а также конструирование панели блокирующих нанотел к S-белку SARS-CoV-2 и тестирование их активности в условиях *in vitro* и *in vivo*.

ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

Культивирование клеток. Используемые в работе линии клеток HEK-293 и HEK-293T (культуры трансформированных клеток человека), а также H1299 (немелкоклеточный рак легкого) были получены из Американской коллекции типовых клеточных культур (ATCC, США). Их культивировали в атмосфере с 5% CO₂ при 37°C в среде DMEM-F12 (“PAA”, Австрия) с добавлением 10%-ной эмбриональной сыворотки крупного рогатого скота (FSB; “Gibco”, США), 2 mM аланил-глутамин (“ПанЭко”, Россия), 20 mM HEPES, 100 мкг/мл пенициллина и 100 мкг/мл стрептомицина (“ПанЭко”).

Выделение РНК, обратная транскрипция и количественная ПЦР. Выделение РНК из клеток проводили на колонках с использованием набора RNeasy Mini kit (“Qiagen”, США) или в препаративных количествах с применением реагента ExtractRNA (“Евроген”, Россия) в соответствии с рекомендациями производителей.

Очищенную РНК использовали для синтеза кДНК при помощи ProtoScript II First Strand cDNA Synthesis Kit (“NEB”, США). Реакционная смесь содержала 500 нг РНК и компоненты вышеприведенного набора с использованием праймера d(T)₂₃. Для получения целевой кДНК использовали специфичный праймер CH2_IgG_sp rev (5'-GGTACGTGCTGTTGAAGTGTCC-3'), инкубировали смесь в течение 1 ч при 42°C, затем 5 мин при 80°C.

Получение генетических конструкций для экспрессии ACE2, эктодоменов S-белка и RBD. кДНК, полученную из клеток линии H1299, амплифицировали с использованием высокоточной полимеразы Tersus (“Евроген”) и праймеров: ACE2_XbaI_dir

(5'-TATATCTAGAGCCACCATGTCAAGCTCTTCCTGGCTC-3') и ACE2 8His SalI_rev (5'-CTCTCTCTGTCGACTTATCAATGGTGGTGGTGGTGGTGGTGGTGGGAAACAGGGGGCTGGTTAGGAG-3') либо ACE2_SalI_rev (5'-ATATGTTCGACCTAAAGGAGGTCTGAACATCATCAGTG-3'). Ампликоны разделяли электрофорезом в 1%-ном агарозном геле, очищали методом твердофазной экстракции и, используя лигазу T4 ("Евроген"), клонировали по сайтам рестрикции XbaI и SalI ("NEB") в лентивирусный экспрессионный вектор pLCMV-PL4-puro.

Для получения последовательности тримеризирующегося эктодомена S-белка и его RBD использовали плазмиды pCAGGS-Spike и pCAGGS-RBD соответственно [10]. Последовательность, кодирующую эктодомен S-белка, амплифицировали с праймерами: S_CO_BsmBI_AGCT_dir (5'-AGAGAGC-

GTCTCAAGCTTAGGCCACCATGTTCGTGTTTC-3') и S_CO_XbaI_rev (5'-AGAGAGTCTAGATTAGTGATGATGATGATGATGTCCCAAGAAGGTCGACAG-3'); а последовательность, кодирующую RBD, — с праймерами S_CO_BsmBI_AGCT_dir и RBD_CO_XbaI_rev (5'-AGAGAGCGTCTCAAGCTTAGGCCACCATGTTCGTGTTTC-3'). Полученные конструкции клонировали в вектор pSI-PL1-10His-puro по сайтам HindIII и XbaI.

Получение упаковочной плазмиды pMD2-S. Для конструирования плазмиды, кодирующей последовательность полноразмерного S-белка, использовали кДНК генома вируса SARS-CoV-2. Фрагменты последовательности, кодирующей S-белок, амплифицировали и соединяли при помощи ПЦР с амплификацией перекрывающихся участков (OE-PCR) с праймерами:

SC2_S_dir (5'-GAACAATGTTTGTCTTTCTTGTTTATTGCC-3'),
SC2_S_dir2 (5'-CCACTAGTCTCTAGTCAGTGTGTTAATC-3'),
SC2_S_rev (5'-TTATGTGTAATGTAATTTGACTCCTT TGAGC-3'),
SC2_S_RT (5'-GTAGCATCCTTGATTTACCTTGCTTC-3'),
SC2_RBD_D1 (5'-AAATCGCTCCAGGGCAA- ACT-3'),
SC2_RBD_D6 (5'-TGCACTTGACCCTCTCTCAG-3'),
SC2_RBD_R1 (5'-TGTGGATCACGGACAGCATC-3'),
SC2_RBD_R6 (5'-ACACCTGTGCCTGTAAACCA-3'),
SC2_Sm_D1 (5'-CACCAGCAACTGTTTGTGGAC-3'),
SC2_Sm_R1 (5'-GCACCAAAGGTCCAACCAGA-3') и
SC2_St_D1 (5'-ACGGCCTTACTGTTTTGCCA-3') — и клонировали в вектор pMD2-G по сайту рестрикции EcoRI.

Получение клеточных линий, гиперэкспрессирующих ACE2, и линий-продуцентов рекомбинантных белков. Плазмиды pSI-S-puro и pSI-RBD-puro использовали для химической трансфекции клеток линии HEK-293T. Трансфекцию проводили по протоколу, адаптированному из опубликованной ранее работы [11], в среде Opti-MEM ("Gibco", США). Через 24 ч среду меняли на свежую с добавлением 1 мкг/мл пурамицина и продолжали культивирование до получения устойчивой к селективному антибиотику культуры.

Плазмиды pLCMV-ACE2-puro и pLCMV-ACE2-His-puro использовали для лентивирусной трансдукции клеток HEK-293 и HEK-293T. Для приготовления препарата лентивируса 1×10^6 клеток HEK-293T высевали на культуральную чашку диаметром 10 см. На следующий день плазмиду pLCMV-ACE2-puro или pLCMV-ACE2-His-puro смешивали в весовом соотношении 5 : 3 : 2 (суммарная масса — 20 мкг) с упаковывающими плазмидами psPAX2 (#12260 Addgene plasmid; <http://n2t.net/addgene:12260>; RRID:Addgene_12260) и pMD2.G (#1225912259 Addgene plasmid; <http://n2t.net/addgene:12259>; RRID:Addgene_12259) (обе плазми-

ды были любезно предоставлены Didier Trono). Через сутки среду заменяли на бессывороточную — DMEM-F12 — с добавлением Serum replacement solution и Lipid mixture ("Peprotech", США) и 4 мМ кофеина ("Sigma-Aldrich", США). Через 48 ч обогащенную вирусом среду собирали, фильтровали, добавляли полибрен ("Sigma-Aldrich") до концентрации 10 мкг/мл и использовали для заражения клеток линии HEK-293T либо HEK-293, высаженных на чашку диаметром 10 см в количестве 5×10^5 клеток. Через сутки среду заменяли на обычную. Через 72 ч после заражения в среду добавляли пурамицин ("Sigma-Aldrich") до концентрации 1 мкг/мл и продолжали культивировать в течение 10 сут.

Наработка, выделение и очистка рекомбинантных белков. Для наработки рекомбинантных растворимых фрагментов S-белка (тримера эктодомена и RBD) и ACE2 клетки высевали на 6 культуральных чашек диаметром 15 см (в расчете 2.5×10^7 клеток/чашка) со средой DMEM-F12, в которую добавляли Serum replacement solution и Lipid mixture, и культивировали в стандартных условиях. Культуральную жидкость, содержащую

целевой белок, собирали через каждые четверо суток в течение 16 дней. После сбора из культуральной жидкости либо немедленно выделяли целевой белок, либо хранили ее при 4°C для объединения со следующей партией. Очистку рекомбинантных белков проводили с использованием HIS Mag Sepharose Excel (“Cytiva”, США) по протоколу производителя. После очистки образцы переводили в фосфатно-солевой буфер (PBS) на колонках PD-10 (“Cytiva”). Концентрацию белка измеряли на спектрофотометре Nanodrop OneC (“Thermo Fisher Scientific”, США) при длине волны 280 нм (OD_{280}). Качество препаратов оценивали по результатам электрофореза в полиакриламидном геле (10%-ный ПААГ в присутствии SDS) с окрашиванием Coomassie Blue G-250 по стандартному протоколу.

Постановка количественной ПЦР. Экспрессию ACE2 в клетках HEK-293-ACE2 и HEK-293T-ACE2 оценивали методом количественной ПЦР с использованием полимеразы HS-Taq (“Евроген”) и праймеров ACE2_qPCR_d1 (5'-GGGATCA-GAGATCGGAAGAAGAAA-3') и ACE2_qPCR_r1 (5'-GGAGGTCTGAACATCATCAGTG-3') (“ДНК-синтез”, Россия). Праймеры подбирали с использованием Primer-BLAST (NIH NCBI, США). Для анализа вторичной структуры праймеров использовали OligoAnalyzer Tool (IDT, США). Все эксперименты проводили в 6 независимых повторах. Представленность в образцах целевых транскриптов сравнивали с использованием метода порогового цикла (C_T). В качестве внутреннего контроля при проведении количественной ПЦР использовали уровень экспрессии гена *GAPDH*, относительно которого нормализовали показатели экспрессии целевых генов (ΔC_T). Изменения уровня экспрессии гена ACE2 в образцах рассчитывали по формуле $2^{-\Delta C_{Te} - \Delta C_{Tk}}$, где ΔC_{Te} и ΔC_{Tk} — разницы значений пороговых циклов для ACE2 и *GAPDH* соответственно [12, 13]. ПЦР проводили на приборе MyiQ Single-Color Real-Time PCR Detection System (“Bio-Rad”, США).

Иммунизация животных и получение генетического материала. В работе использовали животных семейства Верблюдовые — *Vicugna pacos* (альпаки). Для первичной иммунизации использовали гомотенную суспензию, содержащую внеклеточную часть тримера S-белка или полный RBD в концентрации 600 мкг в 500 мкл PBS и 500 мкл полного адьюванта Фрейнда (“Pierce”, США). Для бустирования использовали 400 мкг того же рекомбинантного белка в 500 мкл PBS, в смеси с 500 мкл неполного адьюванта Фрейнда. Всего было проведено три бустирования с интервалом в две недели; антигены вводили подкожно и внутримышечно в бедро животного. При каждой иммунизации у животных брали по 10 мл венозной крови, отделяли сыворотку и хранили при темпе-

ратуре 4°C для последующего определения сывороточной концентрации антител к S-белку.

Забор крови для выделения генетического материала проводили через 5 суток после третьего бустирования. Фракции мононуклеарных клеток периферической крови получали из 100 мл свежесобранной крови центрифугированием в фиколе плотностью 1.077 г/см³ (“ПанЭко”) согласно стандартному протоколу.

Предварительный отбор В-лимфоцитов. Препарат тримерного эктодомена S-белка конъюгировали с NHS-LC-biotin (“Covachem”, США) по стандартному протоколу. Биотинилированный препарат использовали для магнитной сепарации В-лимфоцитов при помощи EasySep Biotin Positive Selection Kit II (“Stemcell Technologies”, Канада) согласно протоколу производителя. Для выделения РНК из селектированной фракции клеток использовали соосадитель Satellite red (“Евроген”).

Сборка библиотек последовательностей, кодирующих нанотела. Для приготовления библиотек использовали кДНК, полученную в 8 независимых реакциях — в каждой из них использовали по 500 нг мРНК. Амплификацию целевых последовательностей проводили с использованием высокоточной полимеразы Tersus (“Евроген”) и праймеров: AlpVHH3_uni_fwd (5'-CCACCATGTCTAGASAGKTGCAGSTSGTRGAGTCTGKGGGAGG-3'), AlpVHH_R1 (5'-AATCCGGATCCGGGGGTCTTCGCTGTGGTGCG-3') и AlpVHH_R2 (5'-AATCCGGATCCGGTTGTGGTTTTGGTGTCTTGGG-3') (последние два сконструированы на основании последовательностей, опубликованных ранее [14]) — по следующей программе: 30 с при 95°C, 30 с при 61°C, 30 с при 72°C (30 циклов). Фагмидную библиотеку получали клонированием амплифицированных VHH-последовательностей в фагмиду pHEN2-XB с использованием эндонуклеаз XbaI и BamHI-HF и ElectroLigase (“NEB”). Трансформацию клеток *Escherichia coli*, штамм TG-1, проводили с помощью электропоратора Genepulser (“Bio-Rad”). Селекцию проводили по опубликованному ранее протоколу [15] с использованием антигена, иммобилизованного на иммунологических пробирках Maxibinding (“SPL Lifesciences”, Корея). Для элюции связавшихся с антигеном фаговых частиц использовали 100 мМ водный раствор триэтиламина. В элюат сразу вносили 1/2 объема 1 М Трис-HCl (pH 7.4). Амплификацию выделенных фаговых частиц, а также оставшихся связанными с иммобилизованным антигеном проводили путем последующего заражения экспоненциально растущей культуры клеток *E. coli* TG-1.

Экспрессия и очистка биотинилированных вариантов кандидатных нанотел. Последовательности VHH-фрагментов из препаратов фагмид, полученных из индивидуальных клонов после селек-

ции, клонировали по сайтам XbaI и BamHI в экспрессионный вектор pET-BAD. Полученные конструкции трансформировали в клетки штамма BL21DE3-BirA, экспрессирующие биотинлигазу. Индивидуальные колонии использовали для получения растворимого белка в жидкой среде в соответствии с описанным ранее протоколом [16] с добавлением в среду для культивации 0.2 mM D-биотина ("Covachem"). Растворимые нанотела выделяли из периплазмы путем лизиса осажденных бактерий в буфере, содержащем 50 mM Трис-HCl (pH 7.4), 150 mM NaCl, 0.1% Triton X-100 и 10 мг/мл лизоцима, с добавлением PMSF, трехкратной обработкой на ультразвуковом дезинтеграторе Soniprep-100 ("Soniprep", США) с последующей очисткой на HIS Mag Sepharose Excel ("Cytiva") и переводом в PBS.

Иммуноферментный анализ (ИФА). Содержание антител в сыворотках животных определяли методом ИФА по стандартному протоколу. Для этого 96-луночные планшеты ("SPL Lifesciences") сенсibilизировали полученным ранее тримерным эктодоменом S-белка SARS-CoV-2 в концентрации 5 мкг/мл, вносили образцы сывороток животных и обрабатывали поликлональными овечьими антителами, специфичными к IgG альпаки (получены ранее в нашей лаборатории). Для детекции использовали меченные пероксидазой хрена (HRP) поликлональные ослиные антитела против IgG овцы ("Santa Cruz", США).

Оценку обогащения библиотек целевыми вариантами нанотел в ходе селекции проводили по опубликованному ранее протоколу [15]. Лунки планшета сенсibilизировали тримерным эктодоменом S-белка (5 мкг/мл), для детекции использовали HRP-меченые моноклональные антитела к фалу M13 11973-MM05T-H ("Sino Biological", Китай).

При проведении прямого ИФА с нанотелами планшеты также покрывали S-белком и обрабатывали нанотелами в концентрации 1 мкг/мл. Детекцию проводили с использованием биотинилированных антител к эпиту Mус ("СайСторЛаб", Россия) и стрептавидина, конъюгированного с HRP ("R&D Systems", США), а для биотинилированных нанотел — напрямую с HRP-конъюгированным стрептавидином. Непрямой ИФА с нанотелами проводили на планшетах, сенсibilизированных биотинилированным нанотелом-1 в концентрации 5 мкг/мл, а для детекции использовали биотинилированное нанотело-2 в концентрации 1 мкг/мл с последующей обработкой HRP-меченым стрептавидином.

Для постановки конкурентного ИФА лунки планшета сенсibilизировали рекомбинантным эктодоменом ACE2 в концентрации 5 мкг/мл и вносили смесь анализируемых нанотел (в различных концентрациях) с биотинилированным три-

мерным эктодоменом S-белка (конечная концентрация 1 мг/мл) в пропорции 1 : 1. После инкубации в течение 1 ч при комнатной температуре в лунки планшета вносили HRP-меченый стрептавидин. В качестве субстрата использовали готовый раствор тетраметилбензидина ("Sigma-Aldrich"). Результаты регистрировали при длине волны 450 нм на микропланшетном сканере Triad ("Dy nex", США).

Измерение активности нанотел с помощью псевдовирального репортера. Для приготовления псевдовиральных частиц использовали векторную плазмиду pLCMV-luc-риго, кодирующую последовательность люциферазы светлячка. Приготовление лентивирусных частиц, псевдотипированных S-белком SARS-CoV-2, проводили по протоколу, описанному выше для лентивирусной трансдукции, за исключением того, что вместо упаковочной плазмиды pMD2.G использовали плазмиду pMD2-S. После сбора и фильтрации вирусосодержащей среды вирусные частицы концентрировали центрифугированием при 35000 g в течение 3 ч. Полученный осадок перерастворяли в среде DMEM-F12 ("PAA"), смешивали в пропорции 1 : 1 с исследуемыми нанотелами, растворенными в различных концентрациях в такой же питательной среде, и инкубировали смесь в течение 20 мин. После этого в смесь добавляли полибрен ("Sigma-Aldrich") до концентрации 5 мкг/мл и вносили в клетки линии HEK-293-ACE2 или HEK-293T-ACE2, которые за 24 ч до эксперимента высевали на 96-луночный планшет (1.2×10^4 клеток/лунка). Планшеты центрифугировали при 850 g в течение 3 ч при 37°C, заменяли псевдовирусосодержащую среду на обычную и культивировали клетки в течение 48 ч. Люминесцентный сигнал регистрировали с использованием Luciferase Assay System ("Promega", США) на микропланшетном сканере Triad ("Dy nex"). За 100% принимали величину люминесцентного сигнала в лунках с псевдовирусом без нанотел.

Биослойная интерферометрия. Для определения константы диссоциации (K_d) взаимодействия нанотел с тримерным эктодоменом S-белка использовали очищенные препараты биотинилированных нанотел и препарат экспонированного тримера S-белка. Определение проводили при помощи биосенсоров Streptavidin (SA) и прибора BLItz ("ForteBio", США). Для оценки перекрытия эпипептидов различных нанотел использовали пары "биотинилированное нанотело-1—небиотинилированное нанотело-2" и препарат рекомбинантного RBD. На биосенсор сорбировали биотинилированное нанотело-1, погружали биосенсор в раствор препарата RBD (1 мг/мл), отмывали и помещали в раствор нанотела-2 (1 мг/мл). Критерием, определяющим взаиморасположение эпипептидов, служила высота плато сигнала после добавления

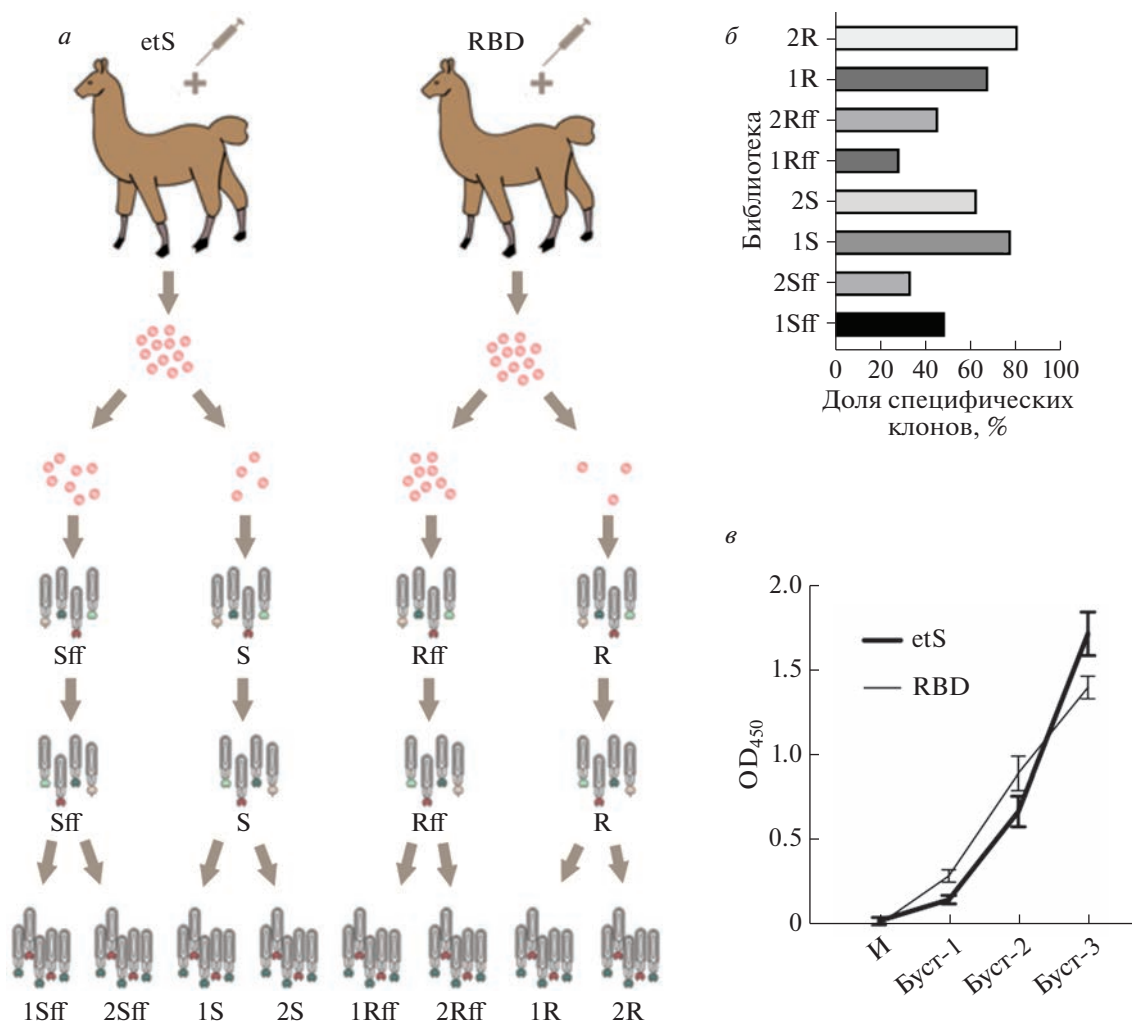


Рис. 1. Селекция нанотел из иммунных библиотек. *а* – Схема приготовления библиотек и проведения селекции. Sff, Rff – библиотеки, полученные из тотальных фракций периферических мононуклеаров животных; S и R приготовлены из селектированных с тримерным эктодоменом S-белка (etS) В-клеток. Под цифрой 1 (1S, 1R, 1Sff, 1Rff) приведены библиотеки, полученные после второго раунда селекции против RBD; под цифрой 2 (2S, 2R, 2Sff, 2Rff) – библиотеки после второго раунда селекции против S-белка. *б* – Доля специфических антител в каждой из библиотек после проведения селекции, определенная по результатам ИФА на единичных клонах. *в* – Специфическая активность сывороточных антител в крови животных на момент каждого из раундов иммунизации. И – иммунизация.

RBD и после отмытки от нанотела-2. В случае, когда оба плато были одинаковой высоты, считали, что эпитопы перекрываются; когда второе плато было выше первого, считали, что эпитопы отделены друг от друга. Для каждой пары нанотел эксперимент повторяли в прямой и обратной постановке.

Статистический анализ. Статистическую обработку результатов проводили с помощью пакета статистических программ Prism 8 (GraphPad Software, США).

РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ

Для конструирования иммунных библиотек использовали генетический материал, получен-

ный от двух иммунизированных альпак (*Vicugna pacos*). Для иммунизации одного из животных использовали тримерный эктодомен S-белка SARS-CoV-2 (Уханьский штамм), второго иммунизировали рекомбинантным RBD S-белка. После проведения четырехкратной иммунизации и достижения высоких уровней сывороточных антител против S-белка по данным ИФА (рис. 1в) у животных проводили забор периферической крови и выделяли субфракции мононуклеарных клеток, связывающихся с имобилизованным на микроносителях S-белком. Помимо отобранных с помощью селекции фракций В-клеток для создания библиотек использовали и тотальные фракции периферических мононуклеаров. Дальнейшую селекцию проводили в два раунда, при-

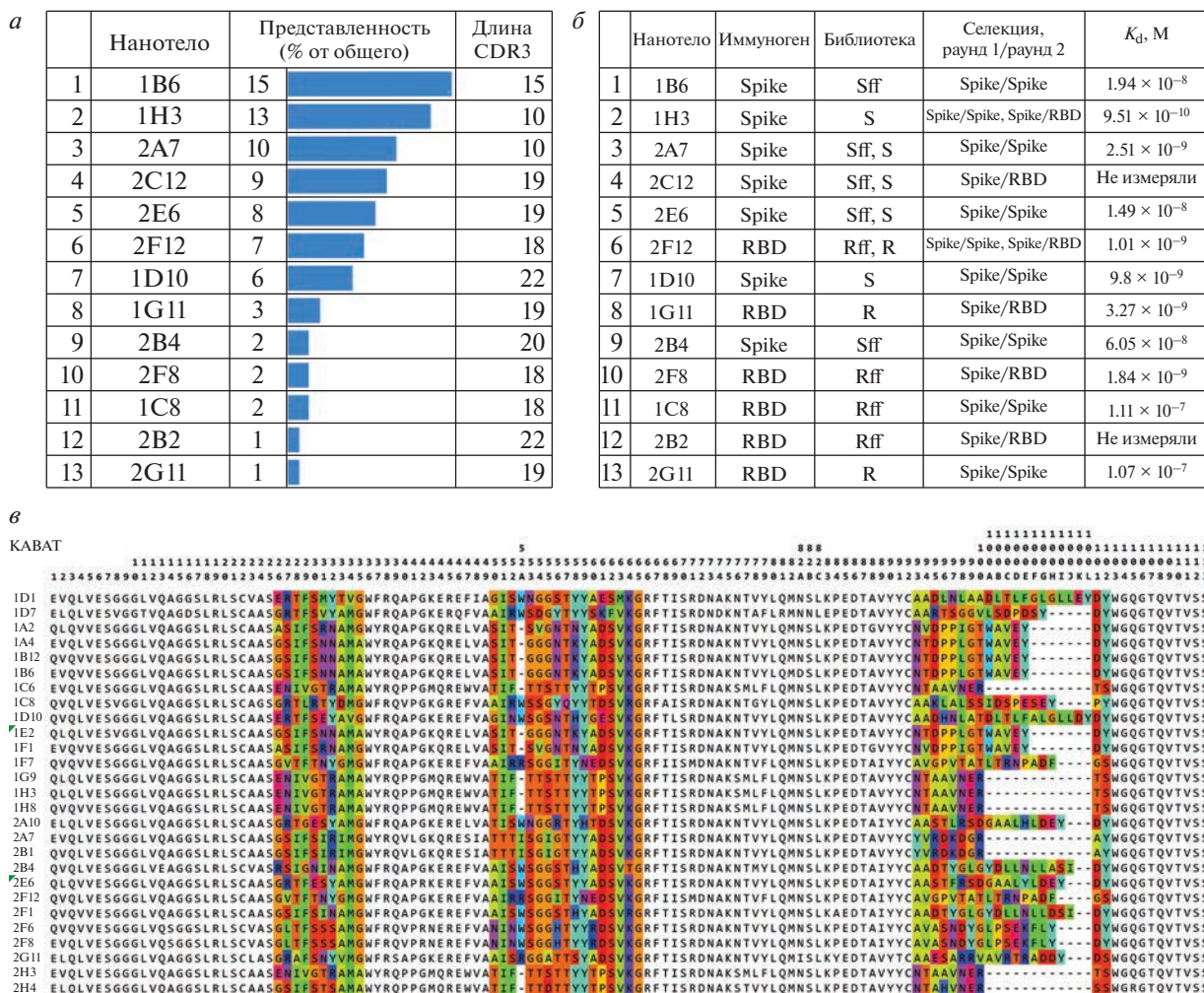


Рис. 2. Характеристики отобранных нанотел. *а* – Перечень наиболее распространенных последовательностей нанотел. Длина варибельных областей (CDR) приведена в виде числа аминокислотных остатков. *б* – Общая характеристика полученных нанотел: иммуноген, в ответ на который сформировалось нанотело; исходная библиотека, в которой оно было обнаружено; мишени, использованные для селекции на первом и втором раундах, и константа диссоциации для растворимого варианта нанотела. *в* – Выравнивание аминокислотных последовательностей нанотел. Сверху – нумерация по Кабату; цветом выделены аминокислоты, входящие в CDR.

чем для второго раунда использовали в параллели либо тримерный эктодомен S-белка, либо RBD. Схема процесса селекции приведена на рис. 1а. По окончании селекции в каждой из библиотек оценивали долю клонов, специфически связывающихся с эктодомом S-белка. Установлено, что наибольшая доля специфических клонов выявлена в библиотеках 2R и 1S. Наименьшее число релевантных вариантов нанотел содержали библиотеки, полученные из тотальных фракций периферических мононуклеаров (рис. 1б). Это означает, что предварительная селекция В-клеток на иммобилизованном антигене позволяет избираться от нецелевых последовательностей, которые иначе продолжают контаминировать смесь бактериофагов даже после проведения двух раундов селекции. Кроме того, использование в качестве мишени для селекции альтернативной фор-

мы целевого белка (эктодомена S-белка в случае RBD-библиотеки или RBD для библиотеки, полученной от животного, иммунизированного эктодомом S-белка) позволяет добиться большей представленности релевантных последовательностей за счет обеднения вариантов, распознающих недоступные эпитопы.

Рестрикционное картирование индивидуальных клонов нанотел, связывающихся с S-белком, и последующее выборочное секвенирование представителей каждого типа показало, что в библиотеках 1S, 2S, 1R и 2R доминируют 13 семейств последовательностей. Представленность каждого из них варьировала от единственной копии до 15% от всех проанализированных клонов, а длина наиболее варибельных областей (CDR3) – от 10 до 22 аминокислотных остатков (рис. 2а). Структура нанотел была характерной для пред-

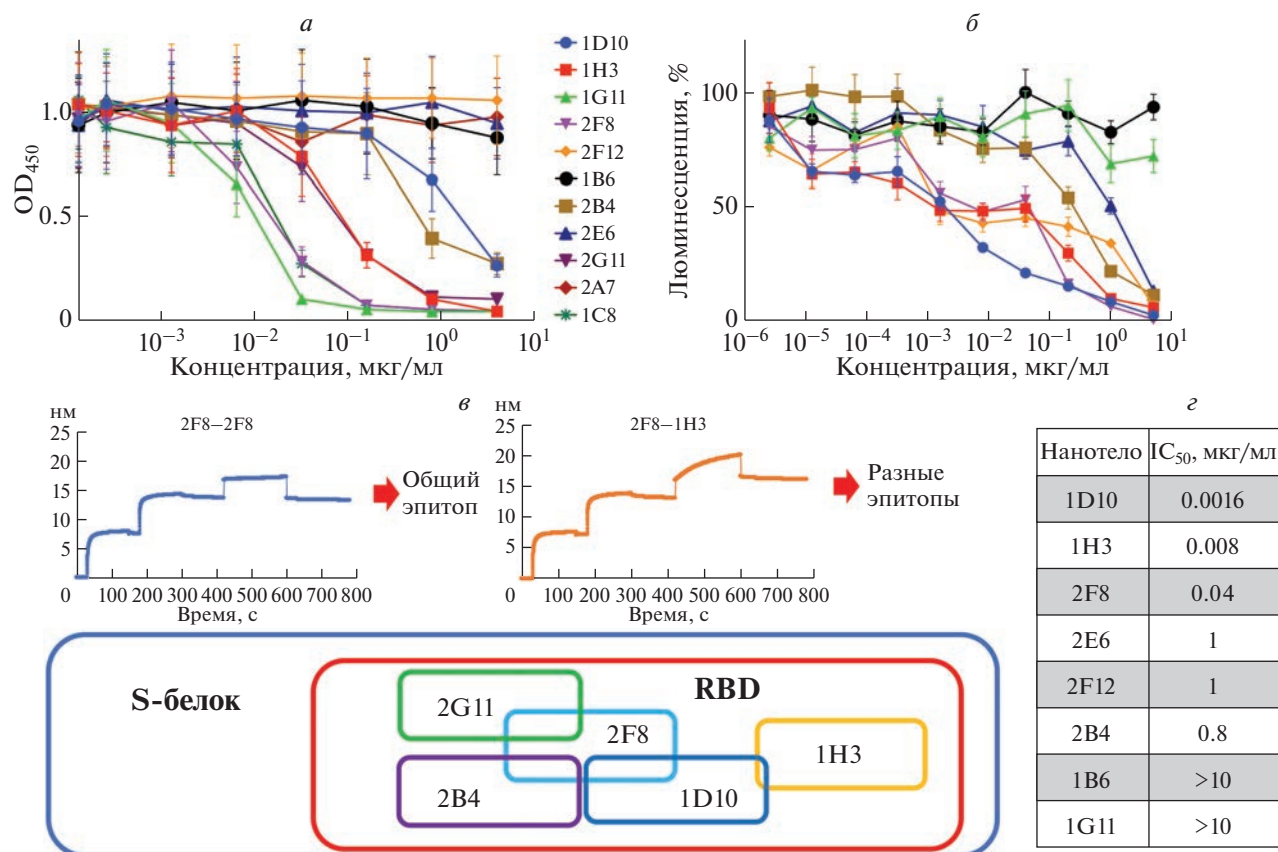


Рис. 3. Функциональная активность отобранных нанотел и эпитопное картирование. *а* – Результаты конкурентного ИФА на эктодомен ACE2 с отобранными нанотелами; детекция – по биотинилированному эктодомену S-белка. *б* – Ингибирование нанотелами трансдукции чувствительных клеток репортерными псевдовirusами. За 100% принят люминесцентный сигнал в контрольных лунках (без добавления нанотел). *в* – Пример анализа на перекрывание эпитопов связывания для двух нанотел (2F8–2F8 – общий эпитоп; 2F8–1H3 – различные эпитопы) и схематическое изображение взаимного расположения эпитопов, распознаваемых нанотелами. *г* – Ингибирование нанотелами трансдукции клеточной линии HEK-293T-ACE2 репортерными псевдовirusами.

ставителей Верблюдовых Нового света; однако наличие дополнительной дисульфидной связи, как отличительной особенности состоящих из тяжелых цепей антител Верблюдовых, было выявлено только у одного варианта – 2B2 (50С–100С). Сравнение последовательностей различных семейств нанотел приведено на рис. 2в.

Представители каждого из отобранных семейств нанотел были экспрессированы в свободной форме и исследованы методом биослойной интерферометрии для определения степени сродства к S-белку и специфичности такого взаимодействия (рис. 2б). Наиболее сильное связывание было зафиксировано для варианта 1H3 ($K_d = 9.51 \times 10^{-10}$ М), а наименьшее – для варианта 2G11 ($K_d = 1.07 \times 10^{-7}$ М). Все проанализированные варианты характеризовались минимальной фоновой активностью в тесте с клеточными лизатами.

Для анализа ингибиторной активности отобранных вариантов в отношении взаимодействия S-белка с ACE2 был использован конкурентный

формат ИФА. При оценке активности нанотел в диапазоне концентраций от 5 мкг/мл до 2.5 пг/мл показано, что варианты 1G11, 1C8, 2F8, 2G11, 1H3, 2B4 и 1D10 блокируют взаимодействие S-белка с ACE2, причем наименьшие концентрации требовались для вариантов 1G11, 1C8 и 2F8, а наибольшая – для 1D10 (рис. 3а). Все другие проанализированные нанотела (2A7, 2E6, 2F12, 1B6) заметной ингибиторной активностью не обладали. С целью верифицировать релевантность полученных результатов на модели, более приближенной к естественным условиям вирусной инфекции, мы получили клеточную линию HEK-293T-ACE2, экспрессирующую на плазматической мембране белок ACE2, и конструкцию pMD2-Swt, кодирующую полноразмерную последовательность S-белка. Используя эту конструкцию, мы получили лентивирусные частицы, псевдотипированные S-белком и несущие репортерный ген люциферазы. Варианты нанотел, сработавшие в конкурентном ИФА, были проанализированы на способность подавлять инфекцию чувствитель-

ных клеток линии НЕК-293Т-ACE2 репортерсодержащими псевдовиральными частицами. Показано, что варианты 1D10, 1H3, 2F8, 2F12, 2B4 и 2E6 блокируют трансдукцию клеток псевдовиральными. Лучшим ингибитором в этом случае оказался вариант 1D10 — с 50%-ной концентрацией ингибирования (IC_{50}) 0.0016 мкг/мл, а самыми слабыми были варианты 2E6 и 2F12 (IC_{50} = 1 мкг/мл). Нанотела 1G11 и 1B6 не ингибировали трансдукцию (рис. 3б и 3г).

С целью определить, связываются ли отобранные нанотела с каким-то одним эпитопом в составе S-белка либо их области связывания не перекрываются, мы провели эксперимент с парным эпитопным картированием. Для этого на биосенсор BLI сорбировали первое нанотело и регистрировали сигнал, после чего сенсор обрабатывали эктодоменом S-белка либо RBD, а затем вторым нанотелом. В том случае, когда обработка вторым нанотелом приводила к увеличению сигнала, считали, что все три белка сохраняют связь с сенсором. Следовательно, анализируемые нанотела связываются с разными эпитопами S-белка. В том случае, когда сигнал сохранялся на прежнем уровне либо снижался, считали, что эпитопы исследуемых нанотел перекрываются. На основании полученных данных построена схематичная карта расположения эпитопов нейтрализующих нанотел. Все варианты нанотел связывались с доменом RBD, причем эпитоп нанотела 2F8 частично перекрывался с эпитопами 2G11, 2B4 и 1D10, которые при этом взаимно не перекрывались; а эпитоп варианта 1H3 незначительно перекрывался только с эпитопом 1D10 (рис. 3в).

ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

При сравнении эффективности проведения селекции из библиотек, полученных с использованием и без использования процедуры обогащения В-клеточной фракции целевыми последовательностями, можно заключить, что ее введение повышает выход релевантных последовательностей и снижает контаминацию выходных пулов неспецифическими вариантами нанотел. Для проведения обеих стадий селекции использовали более интенсивные, чем обычно, режимы отмывки, что теоретически должно снижать влияние селекции В-клеток за счет быстрого удаления несвязавшихся бактериофагов [17], однако различия в содержании целевых клонов между библиотеками сохранились на достаточно высоком уровне. На основании этого можно сделать вывод, что обогащение целевыми В-клетками перед приготовлением библиотеки позволяет снизить трудозатраты и быстро получить высокоспецифичные нанотела против мишени. Вместе с тем, в библиотеках, полученных из тотальной фракции мононуклеаров, помимо примесных, по-видимо-

му, сохранялось большее разнообразие редких целевых последовательностей нанотел [18], что позволило получить из них варианты, обладающие высоким сродством к S-белку и хорошей нейтрализующей активностью (например, 2F8). Сравнивая результаты проведенной нами селекции с полученными другими авторами [19–24], можно отметить их хорошую сходимость. Так, при селекции нанотел к SARS-CoV-2 из синтетических библиотек [22], либо с использованием глубокого секвенирования последовательностей между раундами селекции [23], либо с применением бактериального дисплея [24] максимальные значения аффинности находились в нижнем наномолярном или субнаномолярном диапазоне, что примерно соответствует характеристикам нанотел, полученных нами. Это подтверждает, что использование обогащения В-клеточной фракции позволяет ускорить получение высокоспецифичных нанотел без снижения качества по сравнению с другими методами селекции.

Оба варианта использованных антигенов обладали высокой иммуногенностью в организме альпак, а высокое содержание специфически связывающихся с ними клонов нанотел уже после двух раундов селекции может быть обусловлено тем, что в ходе отбора чередовали использование S-белка и RBD. При анализе структуры отобранных нанотел следует обратить внимание на низкую частоту возникновения характерной для тяжелоцепочечных антител Верблюдовых дополнительной дисульфидной связи, что может быть связано с особенностями наиболее иммуногенных эпитопов в структуре белка. Это может затруднить использование этих нанотел в качестве лекарственных средств, для которых важны такие показатели как устойчивость к агрегированию и высокая стабильность в широком диапазоне температур [25].

Сравнивая результаты тестирования отобранных нанотел в конкурентном ИФА и в тесте с репортерными псевдовиральными частицами, мы пришли к заключению, что применение ИФА позволяет с высокой долей вероятности предсказать наличие биологической активности нанотел и в тестах *in vivo*, так как этот метод нечувствителен к побочным факторам и определяет только взаимодействие между S-белком и его мишенью — ACE2. При этом следует учитывать, что конкурентный ИФА не позволяет выявлять нанотела, нейтрализующие вирус иным образом, например за счет блокирования конформации S-белка в неактивированной форме, что препятствует слиянию мембран вируса и клетки [26]. Мы предполагаем, что именно этим эффектом можно объяснить результаты, полученные при измерении активности нанотел 2E6 и 2F12: неактивные в конкурентном ИФА, они блокировали трансдукцию чувствительных клеток репортерными псев-

довirusными частицами. В целом, по результатам проведенной работы, для дальнейших исследований — с целью создания терапевтического препарата — логично выбрать наиболее активные нанотела: 1D10, 1H3 и 2F8, — для которых получены наименьшие значения IC_{50} как в конкурентном ИФА, так и в тесте с псевдовirusными частицами.

Эпитопы, с которыми связываются отобранные нанотела, различаются и перекрываются на небольших областях. Это повышает вероятность того, что как минимум часть идентифицированных нанотел сохранит свою активность и в отношении мутантных вариантов S-белка, в частности B.1.617.2 и B.1.1.529. Мы считаем, что наибольшей терапевтической эффективностью будет обладать коктейль из нескольких вариантов нанотел [27]. Это позволит перекрыть широкий спектр эпитопов SARS-CoV-2 и нейтрализовать появляющиеся эволюционные варианты вируса без необходимости изменять формулу лекарственного средства [28]. Малый размер нанотел, их физико-химические свойства и возможность быстрого получения — эти факторы позволяют надеяться на применение такого коктейля в ингаляционной форме [29]. Это важно для быстрой доставки препарата к цели — в клетки легочного эпителия [30]. Такой способ введения обеспечит быстрое достижение терапевтического эффекта на ранней стадии заражения, а также позволит применять препарат в профилактических целях.

При помощи селекции В-клеток перед сборкой фагмидных библиотек и чередования антигенов в ходе отбора нам удалось получить панель нейтрализующих нанотел к S-белку SARS-CoV-2, которые эффективно ингибируют взаимодействие между S-белком и его клеточным рецептором — ACE2. Полученные нанотела с высокой аффинностью и специфичностью связываются с разными эпитопами в области RBD S-белка SARS-CoV-2. Терапевтическое применение коктейля таких нанотел можно рассматривать как потенциальное противовирусное средство, эффективное в отношении вновь появляющихся вариантов вируса SARS-CoV-2, быстро эволюционирующего в человеческой популяции.

Работа выполнена при поддержке гранта РФФИ № 20-04-60488.

Все манипуляции, проведенные с животными, соответствовали этическим стандартам, сформулированным в Хельсинкской декларации и принятой практике проведения таких процедур.

Авторы заявляют об отсутствии конфликта интересов.

СПИСОК ЛИТЕРАТУРЫ

1. Pidiyar V., Kumraj G., Ahmed K., Ahmed S., Shah S., Majumder P., Verma B., Pathak S., Mukherjee S.

- (2022) COVID-19 management landscape: a need for an affordable platform to manufacture safe and efficacious biotherapeutics and prophylactics for the developing countries. *Vaccine*. **40**, 5302–5312.
2. Zhou D., Zhou R., Chen Z. (2022) Human neutralizing antibodies for SARS-CoV-2 prevention and immunotherapy. *Immunother. Adv.* **2**, Itab027.
3. Miguez-Rey E., Choi D., Kim S., Yoon S., Sandulescu O. (2022) Monoclonal antibody therapies in the management of SARS-CoV-2 infection. *Expert Opin. Investig. Drugs*. **31**, 41–58.
4. Liu L., Iketani S., Guo Y., Chan J.F., Wang M., Liu L., Luo Y., Chu H., Huang Y., Nair M.S., Yu J., Chik K.K., Yuen T.T., Yoon C., To K.K., Chen H., Yin M.T., Sobieszczyk M.E., Huang Y., Wang H.H., Sheng Z., Yuen K.Y., Ho D.D. (2022) Striking antibody evasion manifested by the Omicron variant of SARS-CoV-2. *Nature*. **602**, 676–681.
5. Planas D., Saunders N., Maes P., Guivel-Benhassine F., Planchais C., Buchrieser J., Bolland W.H., Porrot F., Staropoli I., Lemoine F., Pere H., Veyer D., Puech J., Rodary J., Baele G., Dellicour S., Raymenants J., Gorissen S., Geenen C., Vanmechelen B., Wawina-Bokalanga T., Marti-Carreras J., Cuypers L., Seve A., Hocqueloux L., Prazuck T., Rey F.A., Simon-Loriere E., Bruel T., Mouquet H., Andre E., Schwartz O. (2022) Considerable escape of SARS-CoV-2 Omicron to antibody neutralization. *Nature*. **602**, 671–675.
6. Hamers-Casterman C., Atarhouch T., Muyldermans S., Robinson G., Hamers C., Song E.B., Bendahman N., Hamers R. (1993). Naturally occurring antibodies devoid of light chains. *Nature*. **363**, 446–448.
7. Тиллиб С.В. (2020) Перспективы использования однодоменных антител в биомедицине. *Молекулярная биология*. **54**(3), 362–373.) <https://doi.org/10.31857/S0026898420030167>
8. Huo J., Le Bas A., Ruza R.R., Duyvesteyn H.M.E., Mikolajek H., Malinauskas T., Tan T.K., Rijal P., Dumoux M., Ward P.N., Ren J., Zhou D., Harrison P.J., Weckener M., Clare D.K., Vogirala V.K., Radecke J., Moynie L., Zhao Y., Gilbert-Jaramillo J., Knight M.L., Tree J.A., Buttigieg K.R., Coombes N., Elmore M.J., Carroll M.W., Carrique L., Shah P.N.M., James W., Townsend A.R., Stuart D.I., Owens R.J., Naismith J.H. (2020) Neutralizing nanobodies bind SARS-CoV-2 spike RBD and block interaction with ACE2. *Nat. Struct. Mol. Biol.* **27**, 846–854.
9. Zost S.J., Gilchuk P., Chen R.E., Case J.B., Reidy J.X., Trivette A., Nargi R.S., Sutton R.E., Suryadevara N., Chen E.C., Binshtein E., Shrihari S., Ostrowski M., Chu H.Y., Didier J.E., MacRenaris K.W., Jones T., Day S., Myers L., Eun-Hyung Lee F., Nguyen D.C., Sanz I., Martinez D.R., Rothlauf P.W., Bloyet L.M., Whelan S.P.J., Baric R.S., Thackray L.B., Diamond M.S., Carnahan R.H., Crowe J.E., Jr. (2020) Rapid isolation and profiling of a diverse panel of human monoclonal antibodies targeting the SARS-CoV-2 spike protein. *Nat. Med.* **26**, 1422–1427.
10. Amanat F., Stadlbauer D., Strohmaier S., Nguyen T.H.O., Chromikova V., McMahon M., Jiang K., Arunkumar G.A., Jurczyszak D., Polanco J., Bermudez-Gonzalez M., Kleiner G., Aydllo T., Miorin L., Fierer D.S., Lugo L.A., Kojic E.M., Stoeve J., Liu S.T.H., Cun-

- ningham-Rundles C., Felgner P.L., Moran T., Garcia-Sastre A., Caplivski D., Cheng A.C., Kedzierska K., Vapalahti O., Hepojoki J.M., Simon V., Krammer F. (2020) A serological assay to detect SARS-CoV-2 seroconversion in humans. *Nat. Med.* **26**, 1033–1036.
11. Fukumoto Y., Obata Y., Ishibashi K., Tamura N., Kikuchi I., Aoyama K., Hattori Y., Tsuda K., Nakayama Y., Yamaguchi N. (2010) Cost-effective gene transfection by DNA compaction at pH 4.0 using acidified, long shelf-life polyethylenimine. *Cytotechnology*. **62**, 73–82.
 12. Ginestier C., Cervera N., Finetti P., Esteyries S., Esterni B., Adelaide J., Xerri L., Viens P., Jacquemier J., Charafe-Jauffret E., Chaffanet M., Birnbaum D., Bertucci F. (2006) Prognosis and gene expression profiling of 20q13-amplified breast cancers. *Clin. Cancer Res.* **12**, 4533–4544.
 13. Pfaffl M.W. (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* **29**, e45.
 14. Maass D.R., Sepulveda J., Pernthaner A., Shoemaker C.B. (2007) Alpaca (*Lama pacos*) as a convenient source of recombinant camelid heavy chain antibodies (VHHs). *J. Immunol. Methods*. **324**, 13–25.
 15. Benhar I., Reiter Y. (2002) Phage display of single-chain antibody constructs. *Curr. Protoc. Immunol.* Chapter 10, Unit 10 19B.
<https://doi.org/10.1002/0471142735.im1019bs48>
 16. Studier F.W. (2005) Protein production by auto-induction in high density shaking cultures. *Protein Expr. Purif.* **41**, 207–234.
 17. Ch'ng A.C.W., Ahmad A., Konthur Z., Lim T.S. (2019) A high-throughput magnetic nanoparticle-based semi-automated antibody phage display biopanning. *Methods Mol. Biol.* **1904**, 377–400.
 18. Yuan T.Z., Garg P., Wang L., Willis J.R., Kwan E., Hernandez A.G.L., Tuscano E., Sever E.N., Keane E., Soto C., Mucker E.M., Fouch M.E., Davidson E., Doranz B.J., Kailasan S., Aman M.J., Li H., Hooper J.W., Saphire E.O., Crowe J.E., Liu Q., Axelrod F., Sato A.K. (2022). Rapid discovery of diverse neutralizing SARS-CoV-2 antibodies from large-scale synthetic phage libraries. *mAbs*. **14**, 2002236.
 19. Favorskaya I.A., Shcheblyakov D.V., Esmagambetov I.B., Dolzhikova I.V., Alekseeva I.A., Korobkova A.I., Voronina D.V., Ryabova E.I., Derkaev A.A., Kovyryshina A.V., Iliukhina A.A., Botikov A.G., Voronina O.L., Egorova D.A., Zubkova O.V., Ryzhova N.N., Aksanova E.I., Kunda M.S., Logunov D.Y., Naroditsky B.S., Gintsburg A.L. (2022) Single-domain antibodies efficiently neutralize SARS-CoV-2 variants of concern. *Front. Immunol.* **13**, 822159.
 20. Ye G., Gallant J., Zheng J., Massey C., Shi K., Tai W., Odle A., Vickers M., Shang J., Wan Y., Du L., Aihara H., Perlman S., LeBeau A., Li F. (2021) The development of *Nanosota-1* as anti-SARS-CoV-2 nanobody drug candidates. *Elife*. **10**, e64815.
<https://doi.org/10.7554/eLife.64815>
 21. Maeda R., Fujita J., Konishi Y., Kazuma Y., Yamazaki H., Anzai I., Watanabe T., Yamaguchi K., Kasai K., Nagata K., Yamaoka Y., Miyakawa K., Ryo A., Shirakawa K., Sato K., Makino F., Matsuura Y., Inoue T., Imura A., Namba K., Takaori-Kondo A. (2022) A panel of nanobodies recognizing conserved hidden clefts of all SARS-CoV-2 spike variants including Omicron. *Commun. Biol.* **5**, 669.
 22. Custodio T.F., Das H., Sheward D.J., Hanke L., Pazicky S., Pieprzyk J., Sorgenfrei M., Schroer M.A., Gruzinov A.Y., Jeffries C.M., Graewert M.A., Svergun D.I., Dobrev N., Remans K., Seeger M.A., McInerney G.M., Murrell B., Hallberg B.M., Low C. (2020) Selection, biophysical and structural analysis of synthetic nanobodies that effectively neutralize SARS-CoV-2. *Nat. Commun.* **11**, 5588.
 23. Hanke L., Vidakovic Perez L., Sheward D.J., Das H., Schulte T., Moliner-Morro A., Corcoran M., Achour A., Karlsson Hedestam G.B., Hallberg B.M., Murrell B., McInerney G.M. (2020) An alpaca nanobody neutralizes SARS-CoV-2 by blocking receptor interaction. *Nat. Commun.* **11**, 4420.
 24. Valenzuela Nieto G., Jara R., Watterson D., Modhiran N., Amarilla A.A., Himelreichs J., Khromykh A.A., Salinas-Rebolledo C., Pinto T., Cheuquemilla Y., Margolles Y., Lopez Gonzalez Del Rey N., Miranda-Chacon Z., Cuevas A., Berking A., Deride C., Gonzalez-Moraga S., Mancilla H., Maturana D., Langer A., Toledo J.P., Muller A., Uberti B., Krall P., Ehrenfeld P., Blesa J., Chana-Cuevas P., Rehren G., Schwefel D., Fernandez L.A., Rojas-Fernandez A. (2021) Potent neutralization of clinical isolates of SARS-CoV-2 D614 and G614 variants by a monomeric, sub-nanomolar affinity nanobody. *Sci. Rep.* **11**, 3318.
 25. Kunz P., Zinner K., Mucke N., Bartoschik T., Muyltermans S., Hoheisel J.D. (2018) The structural basis of nanobody unfolding reversibility and thermoresistance. *Sci. Rep.* **8**, 7934.
 26. Rossey I., Gilman M.S., Kabeche S.C., Sedeyn K., Wrapp D., Kanekiyo M., Chen M., Mas V., Spitaels J., Melero J.A., Graham B.S., Schepens B., McLellan J.S., Saelens X. (2017) Potent single-domain antibodies that arrest respiratory syncytial virus fusion protein in its prefusion state. *Nat. Commun.* **8**, 14158.
 27. Touret F., Baronti C., Pastorino B., Villarreal P.M.S., Ninove L., Nougaiere A., de Lamballerie X. (2022) *In vitro* activity of therapeutic antibodies against SARS-CoV-2 Omicron BA.1, BA.2 and BA.5. *Sci. Rep.* **12**, 12609.
 28. Hwang Y.C., Lu R.M., Su S.C., Chiang P.Y., Ko S.H., Ke F.Y., Liang K.H., Hsieh T.Y., Wu H.C. (2022) Monoclonal antibodies for COVID-19 therapy and SARS-CoV-2 detection. *J. Biomed. Sci.* **29**(1), 1.
 29. Piepenbrink M.S., Park J.G., Oladunni F.S., Deshpande A., Basu M., Sarkar S., Loos A., Woo J., Lovanti P., Sloan D., Ye C., Chiem K., Bates C.W., Burch R.E., Erdmann N.B., Goepfert P.A., Truong V.L., Walter M.R., Martinez-Sobrido L., Kobia J.J. (2021) Therapeutic activity of an inhaled potent SARS-CoV-2 neutralizing human monoclonal antibody in hamsters. *Cell Rep. Med.* **2**, 100218.
 30. Lu J., Yin Q., Pei R., Zhang Q., Qu Y., Pan Y., Sun L., Gao D., Liang C., Yang J., Wu W., Li J., Cui Z., Wang Z., Li X., Li D., Wang S., Duan K., Guan W., Liang M., Yang X. (2022) Nasal delivery of broadly neutralizing antibodies protects mice from lethal challenge with SARS-CoV-2 delta and omicron variants. *Virol. Sin.* **37**, 238–247.

Development of a Series of Neutralizing Nanobodies against SARS-CoV-2 Spike Protein

V. A. Zhuchkov¹, S. V. Ivanov¹, J. E. Kravchenko¹, and S. P. Chumakov^{1, *}

¹*Shemyakin–Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, Moscow, 117997 Russia*

**e-mail: hathkul@gmail.com*

Countering the spread of new respiratory infections and reducing the damage they cause to society requires efficient strategies for rapid development of targeted therapeutics, such as monoclonal antibodies. Nanobodies, defined as variable fragments of heavy-chain camelid antibodies, have a set of characteristics that make them particularly convenient for this purpose. The speed at which the SARS-CoV-2 pandemic had spread has confirmed that a key factor in the development of therapeutics is obtaining highly effective blocking agents as soon as possible, as well as the diversity of epitopes to which these agents bind. We have optimized the process of selection of blocking nanobodies from the genetic material of camelids and obtained a panel of nanobody structures with affinity to spike protein in the lower nanomolar and picomolar ranges and high binding specificity. The subset of nanobodies that demonstrate the ability to block the interaction between the spike protein and the cellular ACE2 receptor was selected in experiments *in vitro* and *in vivo*. It has been established that the epitopes bound by the nanobodies are located in the RBD domain of the spike protein and have little overlap. The diversity of binding regions may allow the mixture of nanobodies to retain potential therapeutic efficacy towards new variants of the spike protein, and the structural features of nanobodies, in particular, their compact size and high stability, indicate the possibility of their utilization in the form of aerosols.

Keywords: nanobodies, SARS-CoV-2, spike protein, RBD, ACE2, neutralizing antibodies, monoclonal antibody, immunotherapy

СТРУКТУРНО-ФУНКЦИОНАЛЬНЫЙ АНАЛИЗ БИОПОЛИМЕРОВ И ИХ КОМПЛЕКСОВ

УДК 577.113.7

ЛИГАНДЫ, СПЕЦИФИЧНЫЕ К ОПРЕДЕЛЕННЫМ ПОСЛЕДОВАТЕЛЬНОСТЯМ ПАР ОСНОВАНИЙ ДНК. XIX. СИНТЕЗ, СПЕКТРАЛЬНЫЕ ХАРАКТЕРИСТИКИ, ВИРУСОЛОГИЧЕСКИЕ И БИОХИМИЧЕСКИЕ ИССЛЕДОВАНИЯ ФЛУОРЕСЦЕНТНЫХ СИММЕТРИЧНЫХ ДИМЕРНЫХ ТРИСБЕНЗИМИДАЗОЛОВ $DB_3(n)$

© 2023 г. А. Ф. Арутюнян^а, А. А. Костюков^б, С. П. Королёв^{с, д}, М. Б. Готтих^{с, д}, О. Ю. Сусова^е,
Д. Н. Калюжный^а, А. Л. Жузе^{а, *}

^аИнститут молекулярной биологии им. В.А. Энгельгардта Российской академии наук,
Москва, 119991 Россия

^бИнститут биохимической физики им. Н.М. Эмануэля Российской академии наук,
Москва, 119334 Россия

^сХимический факультет Московского государственного университета им. М.В. Ломоносова,
Москва, 119991 Россия

^дНаучно-исследовательский институт физико-химической биологии им. А.Н. Белозерского
Московского государственного университета им. М.В. Ломоносова, Москва, 119991 Россия

^еИнститут канцерогенеза Национального медицинского исследовательского центра онкологии им. Н.Н. Блохина,
Москва, 115478 Россия

*e-mail: zhuze@eimb.ru

Поступила в редакцию 02.08.2022 г.

После доработки 02.11.2022 г.

Принята к публикации 03.11.2022 г.

Синтезирована серия новых флуоресцентных узкобороздочных лигандов $DB_3(n)$ и охарактеризованы свойства этих лигандов. Соединения $DB_3(n)$ на основе димерных трисбензимидазолов обладают способностью связываться с АТ-участками ДНК. Синтез $DB_3(n)$, трисбензимидазольные фрагменты которых связаны олигометиленовыми линкерами разной длины ($n = 1, 5, 9$), основан на конденсации мономерного трисбензимидазола MB_3 с α, ω -алкилдикарбоновыми кислотами. $DB_3(n)$ в субмикромольной концентрации (0.20–0.30 мкМ) оказались эффективными ингибиторами каталитической активности интегразы ВИЧ-1. Обнаружено, что $DB_3(n)$ ингибирует каталитическую активность ДНК-топоизомеразы I в низких микромолярных концентрациях.

Ключевые слова: ДНК, специфичный к последовательности узкобороздочный лиганд, флуоресцентные димерные трисбензимидазолы $DB_3(n)$, спектральные методы, интегразы ВИЧ-1, ДНК-топоизомераза I

DOI: 10.31857/S0026898423030023, EDN: CGXJLT

ВВЕДЕНИЕ

Создание и изучение низкомолекулярных соединений, сайт-специфически узнающих нуклеотидные последовательности в дцДНК, остается одной из актуальных проблем биоорганиче-

ской химии и молекулярной биологии, поскольку химиотерапевтическая активность большинства из существующих противоопухолевых препаратов зависит от аффинности и селективности их взаимодействия с ДНК. В перспективе подобные молекулярные инструменты могут использоваться для изучения и контроля экспрессии конкретных генов. Создание таких мишень-направленных низкомолекулярных соединений представляет важную фармакологическую задачу.

Низкомолекулярные сайт-специфические соединения, способные образовывать нековалентные комплексы с ДНК, можно разделить на две группы в зависимости от механизма их взаимодействия с ДНК: интеркаляторы и соединения,

Сокращения: топо-I – ДНК-топоизомераза I; ВИЧ-1 – вирус иммунодефицита человека типа I; ссДНК – суперскрученная ДНК; релаксирДНК – релаксированная ДНК; PBS – фосфатно-солевой буфер; IC_{50} – 50%-ная ингибирующая концентрация; HBTU – (2-(1*H*-бензотриазол-1-ил)-1,1,3,3-тетраметилурония гексафторфосфат; Boc-GlyOH – *N*-третбутилоксикрбонилглицин; *i*-BuOC(O)Cl – изобутилхлорформат; NMM – *N*-метилморфолин; DMF – *N,N*-диметилформамид; DIPEA – *N,N*-диизопропилэтиламин.

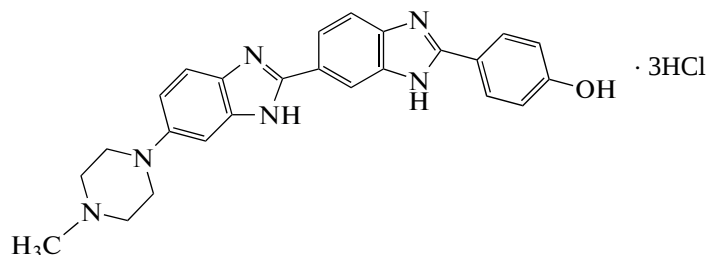


Рис. 1. Химическая структура Hoechst 33258.

связывающиеся в широкой либо в узких бороздках двойной спирали. Наша работа посвящена синтезу и изучению свойств лигандов, связывающихся с дцДНК по узкой бороздке, к которым относятся узкобороздочные лиганды, в том числе антибиотики нетропсин и дистамицин А, флуоресцентные красители Hoechst 33258 и DAPI, а также беренил и подобные диамидины [1, 2]. Существует прямая зависимость между биологической активностью этих лигандов и прочностью их комплекса с АТ-последовательностями в узкой бороздке. Узкая бороздка ДНК важна для функционирования большого числа ДНК-зависимых ферментов и аллостерической регуляции транскрипционных факторов. Поэтому она представляет собой наиболее удобную мишень для проявления биологической активности синтетических терапевтических препаратов и соединений, способных специфически взаимодействовать с ДНК [3]. К числу таких перспективных соединений относятся производные бензимидазола. Бензимидазол, структурно родственный пуриновым основаниям ДНК, обнаружен в биологически значимом природном соединении – витамине В₁₂ [4]. Производные бензимидазола проявляют широкий диапазон биологической активности [5–8]. Кроме того, бисбензимидазолы интенсивно изучают в качестве узкобороздочных агентов [9, 10], а замещенные производные бензимидазола нашли применение в различных терапевтических областях. Кегі и соавт. [11] опубликован всесторонний обзор текущих разработок соединений на основе бензимидазола во всем спектре приложений медицинской химии в качестве противоопухолевых, антибактериальных, противогрибковых, противовоспалительных, обезболивающих, анти-ВИЧ, антиоксидантных, противосудорожных, противотуберкулезных, противодиабетических, противолейшманиозных, антигистаминных, противомаларийных и других лекарственных средств.

Наша работа посвящена синтезу и изучению АТ-специфичных узкобороздочных лигандов на основе бисбензимидазольного красителя Hoechst 33258 (рис. 1), широко используемого в цитологии в качестве ДНК-специфичной флуоресцентной метки [12].

Так называемые узкобороздочные лиганды в значительной степени лишены недостатков, свойственных традиционным биологически активным препаратам на основе алкилирующих агентов и интеркаляторов. Они, в частности, не повреждают ДНК, не вызывают значительного искажения ее пространственной структуры и не проявляют мутагенного эффекта.

Нами осуществлен синтез флуоресцентного мономерного трисбензимидазола МВ₃. На его основе получены димерные трисбензимидазолы DB₃(*n*), где *n* – число метиленовых групп в линкере, соединяющем два трисбензимидазольных блока (рис. 2), и изучены их физико-химические и биологические свойства.

Структура линкеров, содержащих олигометиленовые группы, конформационно подвижна. Она обладает низким сродством к сахарофосфатному остову и основаниям ДНК. Это позволяет молекулам DB₃(*n*) принимать форму, изоспиральную узкой бороздке ДНК, благодаря чему они способны взаимодействовать с ДНК как бидентантные лиганды, связываясь с двумя сайтами (каждый из трех АТ-пар), разделенных двумя и более парами любых нуклеотидов (рис. 3).

ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

Материалы и методы. В работе использовали: HBTU, NMM, *N*-метилморфолин, DMSO (“Alfa Aesar”, США); диизопропилэтиламин (“Fluka”, Германия); ундекандиовую кислоту, 4-аминобензонитрил (“Merck”, Германия); пимелиновую кислоту, изобутилхлорформиат, 1-метилпиперазин (“Acros Organics”, Бельгия); Вос-глицин, DMF, AcOH, AcOEt, EtOH, MeOH, *i*PrOH (“Рехим”, Россия). Растворы веществ в органических растворителях сушили над Na₂SO₄. Растворители упаривали на ротаторном испарителе в вакууме водоструйного насоса, как правило, при 40–50°C. Вещества высушивали в вакууме над P₂O₅ и NaOH. Температуры плавления определяли на приборе Boethius (Германия) и не исправляли. Гидрирование проводили над 10% Pd/C (“Merck”) при атмосферном давлении и комнатной температуре до прекращения поглощения водорода.

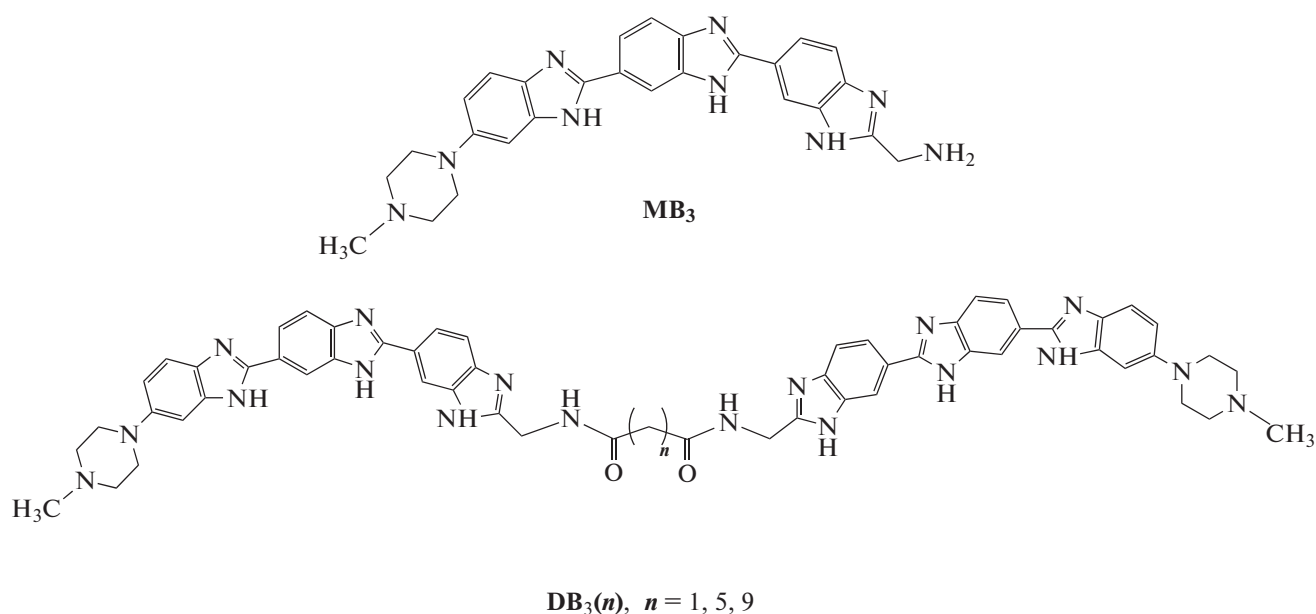


Рис. 2. Химическая структура мономерного трисбензимидазола **MB₃** и димерных трисбензимидазолов **DB₃(n)**.

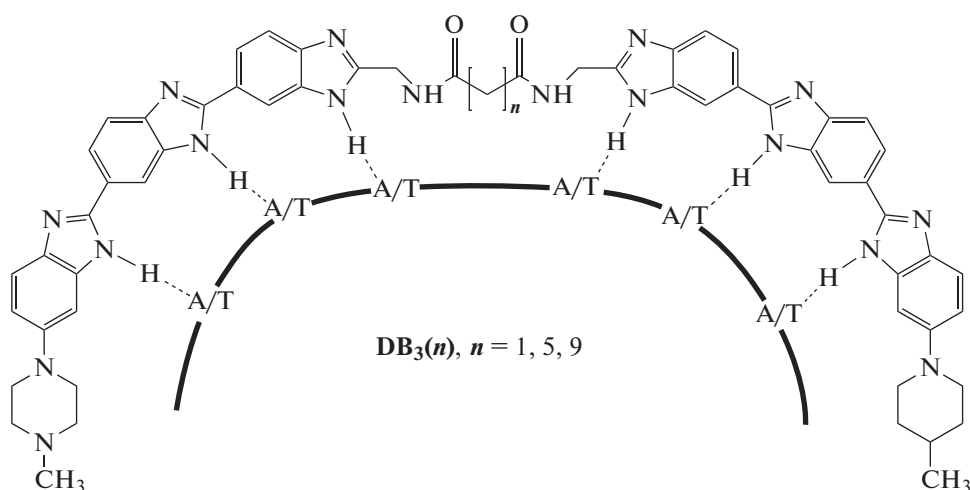


Рис. 3. Схематическое изображение комплекса **DB₃(n)** с двумя сайтами ДНК, состоящими из трех подряд расположенных АТ-пар. Пунктирные линии — водородные связи. Толстой линией изображена ДНК.

Чистоту полученных соединений контролировали с помощью ТСХ на пластинках Kieselgel 60 F₂₅₄ (“Merck”) в системах **A**, MeOH–конц. NH₄OH (25 : 1); **B**, MeOH–TFA–H₂O (5 : 1 : 2). Вещества на хроматограммах обнаруживали в УФ-свете по поглощению при 254 нм и/или по флуоресценции при 365 нм.

Спектры ¹H-ЯМР регистрировали на спектрометре Avance 3 300 МГц (“Bruker”, Германия) в DMSO-*d*₆ (δ, м.д.; J, Гц) при 30°C; в качестве внутреннего стандарта использовали сигнал остаточных протонов растворителя.

Масс-спектры регистрировали методом MALDI-TOF на приборе 4800 Plus (“AB Sciex”,

США) в режиме регистрации положительных ионов; матрица — 2,5-дигидроксibenзойная кислота; лазер Nd:YAG, 355 нм.

Спектральные исследования DB₃(n). Спектры поглощения регистрировали на спектрофотометре Shimadzu UV-3101 (Япония). Препарат — ДНК из спермы лосося (Деринат, Россия).

Спектры флуоресценции регистрировали на спектрофлуориметре Fluorate Rapogama (Россия); возбуждение при 320 нм. Препарат — ДНК из спермы лосося (Деринат).

Спектры КД регистрировали на спектрополяриметре Jasco J715 (“Jasco”, Япония). Препарат —

жидкокристаллическая дисперсия, приготовленная из фрагментированной ДНК спермы лосося.

Химический синтез. Синтез 4-[6-(4-метилпиперазин-1-ил)-1*H*-1,3-бензодиазол-2-ил]-2-нитроанилина (III). К раствору 5-(4-метилпиперазин-1-ил)бензол-1,2-диамина (I) [13] в 50 мл AcOH, полученного из 3.74 г (15.6 мМ) исходного 5-(4-метилпиперазин-1-ил)-2-нитрофениламина, добавили 3.83 г (15.6 мМ) дихлоргидрата имидоэфира (II) [14], реакционную смесь кипятили в течение 1 ч. Выпавший осадок отфильтровывали, промывали на фильтре AcOH до отсутствия окраски фильтрата и растворяли полученное желтое твердое вещество в минимальном количестве горячей воды. К водному раствору приливали концентрированную NH_4OH до получения сильно щелочной среды, выпавший осадок отфильтровывали и высушивали. Выход соединения (III) в виде кирпично-красного порошка 2.53 г (46%), R_f 0.60 (A), т. пл. 152°C. $^1\text{H-NMR}$ ($\text{DMSO}-d_6$): 2.26 (3H, s, N-CH₃), 3.13 (4H, m, pip), 7.01 (3H, m, ArH), 7.39 (1H, dd, $J = 8.6$, $J = 34.7$, ArH), 7.74 (2H, brs, -NH₂), 8.14 (1H, m, ArH), 8.75 (1H, dd, $J = 1.8$, $J = 9.9$, ArH), 12.59 (1H, s, BimNH).

Синтез 4-[6-(4-метилпиперазин-1-ил)-1*H*-1,3-бензодиазол-2-ил]бензол-1,2-диамина (IV). Суспендировали 25 г 10% Pd/C в 20 мл AcOH и при перемешивании пропускали интенсивный ток водорода до полного прекращения поглощения. Затем приливали раствор 2.53 г (7.18 мМ) соединения (III) в 30 мл AcOH и через получившуюся смесь начинали пропускать водород. Через 2.5 ч (по данным ТСХ в системе этанол/ NH_4OH (25 : 1)) исходного вещества не осталось. Реакционную массу отфильтровывали от катализатора, фильтрат упаривали при пониженном давлении. К образовавшемуся маслу добавляли 50 мл концентрированной NH_4OH и оставляли на ночь при 4°C. На следующий день твердый осадок отфильтровывали и высушивали. Выход целевого продукта в виде серого порошка 2.12 г (92%), R_f 0.52 (A), т. пл. 156–159°C. $^1\text{H-NMR}$ ($\text{DMSO}-d_6$): 2.24 (3H, s, N-CH₃), 3.09 (4H, m, pip), 4.62 (2H, br.s, -NH₂), 4.88 (2H, br.s, -NH₂), 6.59 (1H, d, $J = 7.95$, ArH), 6.84 (2H, dd, $J = 1.8$, 8.7 Hz, ArH), 7.16 (1H, dd, $J = 1.2$, 8.1 Hz, ArH), 7.34 (2H, d, $J = 1.65$, ArH), 12.06 (1H, s, BimNH).

Синтез 4-{6-[6-(4-метилпиперазин-1-ил)-1*H*-1,3-бензодиазол-2-ил]-1*H*-1,3-бензодиазол-2-ил}-2-нитроанилина (V). Растворяли 2.12 г (6.6 мМ) диамина (IV) и 1.6 г (6.6 мМ) дихлоргидрата имидоэфира (II) [14] в 80 мл смеси EtOH/2,2,2-трифторэтанола (1 : 1), реакционную массу кипятили с обратным холодильником в течение 3 ч. Выпавший осадок отфильтровывали и заливали 50 мл конц. NH_4OH . Выпавшее твердое вещество оранжевого цвета было отфильтровано и высушено.

Выход соединения (V) 1.26 г (41%), R_f 0.41(A), т. пл. 240°C. $^1\text{H-NMR}$ ($\text{DMSO}-d_6$): 2.24 (3H, s, N-CH₃), 3.12 (4H, m, pip), 6.93 (1H, dd, $J = 1.5$, 9 Hz, ArH), 7.01 (1H, s, ArH), 7.20 (1H, d, $J = 8.6$, ArH), 7.44 (1H, d, $J = 8.7$, -NH₂), 7.67 (1H, d, $J = 8.4$, ArH), 7.84 (2H, brs, Ar-NH₂), 8.02 (1H, d, $J = 8.5$, ArH), 8.25 (1H, dd, $J = 1.5$, 8.9 Hz, ArH), 8.30 (1H, s, ArH), 8.88 (1H, d, $J = 8.5$, ArH).

Синтез 4-{6-[6-(4-метилпиперазин-1-ил)-1*H*-1,3-бензодиазол-2-ил]-1*H*-1,3-бензодиазол-2-ил}бензол-1,2-диамина (VI). Суспендировали 0.12 г 10% Pd/C в 20 мл AcOH и при перемешивании пропускали интенсивный ток водорода до полного прекращения поглощения. Затем прилили раствор 1.26 г (2.7 мМ) 2-нитроанилина (V) в 30 мл AcOH и начали пропускать водород через получившуюся смесь. Через 2.5 ч (по данным ТСХ в системе EtOH/ NH_4OH (25 : 1)) исходного вещества не осталось. Реакционный раствор отфильтровывали от катализатора, фильтрат упаривали при пониженном давлении. Образовавшееся масло было залито 50 мл концентрированной NH_4OH и оставлено на ночь при 4°C. Твердый осадок отфильтровывали и высушили. Выход соединения (VI) в виде серого порошка 1.13 г (95%), R_f 0.45 (A), т. пл. 234°C. Масс-спектр: 438.49, рассчитано: 438.52 ($\text{C}_{25}\text{H}_{26}\text{N}_8$), $^1\text{H-NMR}$ ($\text{DMSO}-d_6$): 2.24 (3H, s, N-CH₃), 3.13 (4H, m, pip), 4.69 (2H, br.s, -NH₂), 5.01 (2H, br.s, Bim-NH₂), 6.63 (1H, d, $J = 8.1$, ArH), 6.92 (1H, d, $J = 7.9$, ArH), 6.99 (1H, m, ArH), 7.27 (1H, dd, $J = 7.9$, 0.6 Hz, Bim-NH₂), 7.44 (1H, d, $J = 1.8$, ArH), 7.49 (2H, m, ArH), 7.94 (1H, d, $J = 8.1$, ArH), 8.20 (1H, s, ArH), 12.55 (2H, brs, BimNH).

Синтез трет-бутил-N-[(2-амино-5-{6-[6-(4-метилпиперазин-1-ил)-1*H*-1,3-бензодиазол-2-ил]-1*H*-1,3-бензодиазол-2-ил}фенил)карбонил]метилкарбамата (VII). К 0.49 г (2.8 ммоль) Boc-Gly-OH приливали 10 мл абс. DMF и 0.5 мл NMM. Смесь охлаждали при перемешивании до -25°C, затем добавляли 0.35 г (2.57 ммоль) *i*-BuOCOCi. Через 10 мин в реакционную смесь вносили предварительно охлажденный до -30°C раствор 1.13 г (2.57 ммоль) диамина (VI) в 15 мл абс. DMF. Реакционную массу выдержали при -30°C в течение 1.5 ч, дали нагреться до комнатной температуры, а затем упарили. Полученное в виде масла соединение (VII) использовали в дальнейшем без очистки.

Синтез мономерного трисбензимидазола MB₃. Масло (VII), полученное из 1.13 г диамина (VI), растворили в 30 мл AcOH и кипятили с обратным холодильником в течение 10 ч. AcOH упарили при пониженном давлении, образовавшееся масло заливали 30 мл концентрированной NH_4OH и оставляли на 4 дня при 4°C. Полученный серый порошок без дополнительной очистки растворяли в конц. HCl и кипятили с обратным холодильником в течение 5 ч, затем реакционную массу при

перемешивании вливали в 100 мл изопропанола. Выпавший осадок отфильтровывали и очищали методом колоночной хроматографии на силикагеле 40–63 мкм, элюент MeOH/NH₄OH (25 : 1). Выход **MB₃** в виде светло-коричневого порошка 0.41 г (33.46% после 3 стадий), *R_f* 0.14 (A); *R_f* 0.68 (B), т. пл. >350°C. Масс-спектр: 477.59, рассчитано: 477.56 (C₂₇H₂₇N₉). ¹H-NMR (CD₃OD): 2.29 (3H, s, N–CH₃), 2.54 (4H, m, pip), 3.16 (4H, m, pip), 4.10 (2H, s, CH₂), 6.93 (1H, dd, *J* = 2.0, 8.8 Hz, ArH), 7.04 (1H, s, ArH), 7.45 (1H, d, *J* = 8.4, ArH), 7.70 (2H, m, ArH), 8.02 (1H, d, *J* = 6.7, ArH), 8.08 (1H, dd, *J* = 0.75, *J* = 8.3, ArH), 8.32 (1H, s, ArH), 8.39 (1H, s, ArH).

Общий метод синтеза DB₃(n). К раствору α,ω-алкилдикарбоновой кислоты (0.1 мМ) в 2 мл абс. ДМФА прибавляли HBTU (0.25 мМ), DIPEA (0.50 мМ) и перемешивали при комнатной температуре в течение 30 мин. К полученному раствору добавляли 0.10 г (0.2 мМ) **MB₃**, перемешивали в течение еще 1 ч и оставляли реакционную массу на ночь. Растворитель упаривали при пониженном давлении, получившееся масло затирали с абс. EtOH. В образовавшуюся суспензию добавляли 0.5 мл 35% HCl в диоксане, отфильтровывали выпавший осадок в виде желтого порошка и высушивали в вакууме над NaOH/P₂O₅. По данным ТСХ в системе MeOH–TFA–H₂O (5 : 1 : 2) продукт реакции – димерный трисбензмидазол **DB₃(n)** – был гомогенным.

DB₃(1)·8HCl: Выход 121 мг (92%), т. пл. >350°C, Масс-спектр, *m/z*: 1024.27 [M + H]⁺, рассчитано M 1023.16 (C₅₇H₅₄N₁₈O₂), ¹H-NMR (DMSO-*d*₆): δ 2.88 (4H, m, pip), 2.91 (6H, s, M–Me), 3.07 (2H, s, –CO–CH₂CO–), 3.20 (4H, m, pip), 3.27 (8H, m, pip), 4.53 (4H, s, 2Bim–CH₂–N), 6.95 (2H, d, *J* = 1.3 Hz, ArH), 7.03 (2H, dd, *J* = 4.5, 1.3 Hz, ArH), 7.47 (2H, d, *J* = 4.5 Hz, ArH), 7.74 (2H, d, *J* = 1.9, 0.4 Hz, ArH), 7.81 (2H, dd, *J* = 5.1, 0.4 ArH), 8.04 (2H, d, *J* = 1.3, Hz, ArH), 8.17 (6H, m, ArH).

DB₃(5)·8HCl: Выход 122 мг (89%), т. пл. >350°C. Масс-спектр, *m/z*: 1080.32 [M + H]⁺, рассчитано M 1079.26 (C₆₁H₆₂N₁₈O₂), ¹H-NMR (DMSO-*d*₆): δ 1.51 (2H, m, CH₂), 1.59 (4H, m, 2CH₂), 2.22 (4H, t, *J* = 6 Hz, 2CH₂CO–), 2.88 (4H, m, pip), 2.91 (6H, s, N–Me), 3.20 (4H, m, pip), 3.27 (8H, m, pip), 4.53 (4H, s, 2Bim–CH₂–N), 6.95 (2H, d, *J* = 1.3 Hz, ArH), 7.03 (2H, dd, *J* = 4.5, 1.3 Hz, ArH), 7.47 (2H, d, *J* = 4.5 Hz, ArH), 7.74 (2H, d, *J* = 1.9, 0.4 Hz, ArH), 7.81 (2H, dd, *J* = 5.1, 0.4 Hz, ArH), 8.04 (2H, d, *J* = 1.3, Hz, ArH), 8.17 (6H, m, ArH).

DB₃(9)·8HCl: Выход 124 мг (87%), т. пл. >350°C. Масс-спектр, *m/z*: 1136.44 [M + H]⁺, рассчитано M 1135.37 (C₆₅H₇₀N₁₈O₂), ¹H-NMR (DMSO-*d*₆): δ 1.26 (10H, m, (CH₂)₅), 1.64 (4H, m, 2CH₂), 2.23 (4H, t, *J* = 6 Hz, 2CH₂CO–), 2.88 (4H, m, pip), 2.91

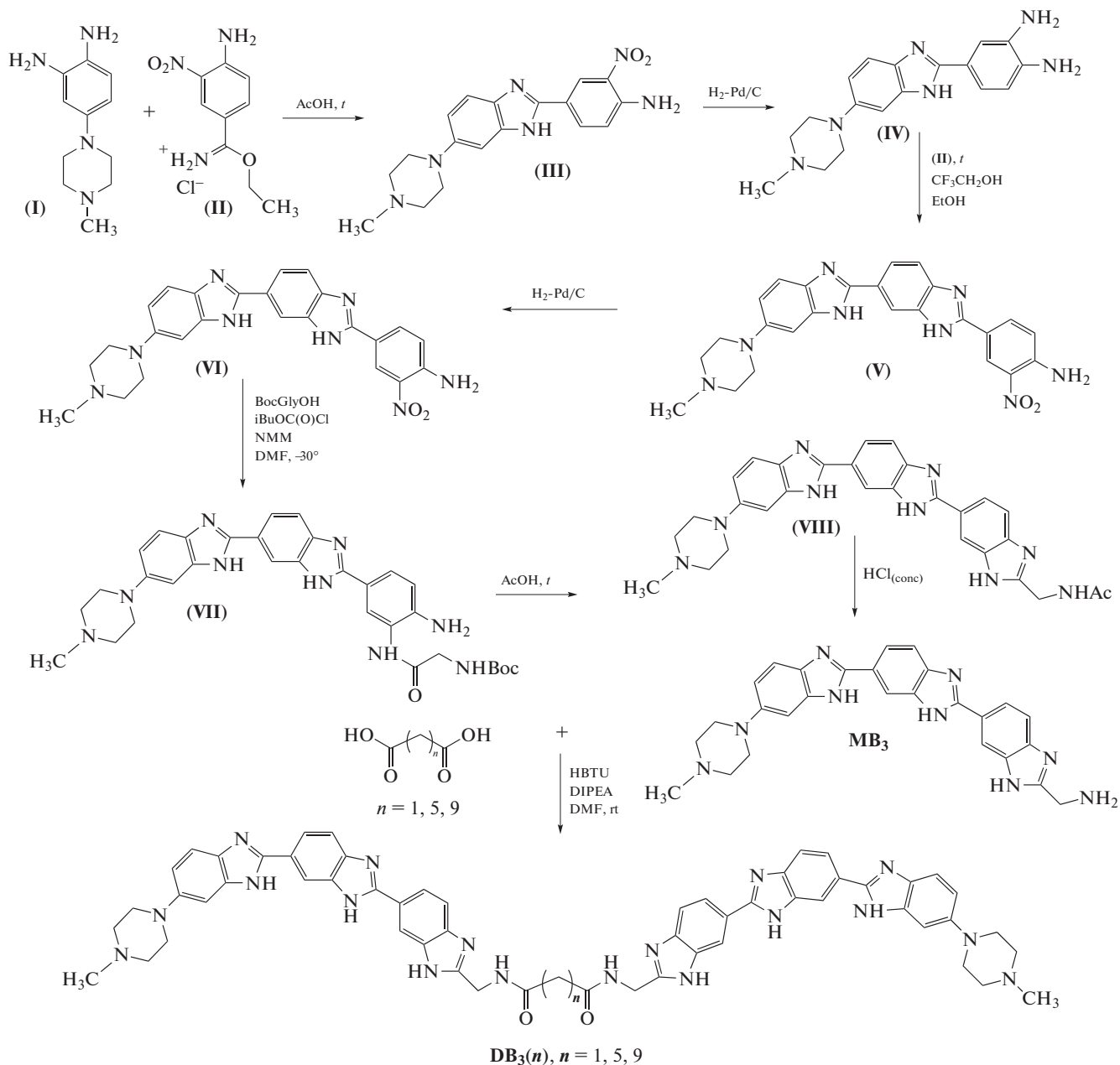
(6H, s, N–Me), 3.20 (4H, m, pip), 3.27 (8H, m, pip), 4.53 (4H, s, 2Bim–CH₂–N), 6.95 (2H, d, *J* = 1.3 Hz, ArH), 7.03 (2H, dd, *J* = 4.5, 1.3 Hz, ArH), 7.47 (2H, d, *J* = 4.5 Hz, ArH), 7.74 (2H, d, *J* = 1.9, 0.4 Hz, ArH), 7.81 (2H, dd, *J* = 5.1, 0.4 ArH), 8.04 (2H, d, *J* = 1.3, Hz, ArH), 8.16 (6H, m, ArH).

Антивирусная активность DB₃(n) и MB₃ в отношении интегразы ВИЧ-1. Рекомбинантную интегразу ВИЧ-1 экспрессировали в *Escherichia coli* штамм BL21 (DE3) Codon Plus и очищали в присутствии ионов Zn²⁺ в отсутствие детергентов согласно [15].

Ингибирование 3'-концевого процессинга, катализируемого интегразой ВИЧ-1. дцДНК-субстрат U5 интегразы ВИЧ-1 (5 нМ), состоящий из 21-звенных олигодезоксирибонуклеотидов U5A (5'-ACTGCTAGAGATTTTCCACAC-3') и флуоресцентно меченного U5B-FAM (5'-FAM-TGTG-GAAAATCTCTAGCAGT-3'), инкубировали с 100 нМ интегразой в 50 мкл буфера (20 мМ HEPES, pH 7.2, 7.5 мМ MgCl₂, 1 мМ DTT) в присутствии возрастающих концентраций ингибитора (0–50 мкМ) при 37°C в течение 100 мин. Реакцию останавливали, добавляя 15 мкл стоп-раствора (30 мМ EDTA, 1.2 М NaOAc, 40 мМ Трис-HCl, pH 8.0, 0.5 мг/мл гликоген). ДНК-дуплекс осаждали, добавляя 200 мкл этанола и центрифугируя при 10000 g в течение 10 мин, анализировали в 20%-ном полиакриламидном геле в присутствии 7 М мочевины. Гель анализировали в многофункциональном сканере GE Healthcare Typhoon FLA 9500. За прохождением реакции следили по появлению полосы, соответствующей укороченному в результате 3'-процессинга меченому олигонуклеотиду U5B-FAM. Эффективность прохождения реакции определяли с использованием программы Quantity One™ 4.6.6. Значения параметров IC₅₀, соответствующих концентрациям ингибиторов, при которых эффективность реакции снижается вдвое, рассчитывали с использованием программы GraphPad prism 7.03.

Ингибирование топо-I. Топо-I – одна из наиболее широко используемых мишеней ряда противоопухолевых средств. Влияние синтезированных соединений на каталитическую активность фермента сравнивали с помощью теста для оценки способности веществ ингибировать топо-I (Topogen, TG1018-2).

Исследуемые соединения добавляли в различных концентрациях в смесь, содержащую плазмидную ссДНК pHot, воду и реакционный буфер. Смесь инкубировали в течение 30 мин при 37°C для комплексообразования и добавляли топо-I. Далее реакцию проводили согласно методике производителя. Продукты реакции разделяли с использованием электрофореза в 1%-ном агарозном геле, с последующим окрашиванием бромистым этидием. ДНК визуализировали на тран-

Рис. 4. Схема синтеза димерного трисбензимидазола $DB_3(n)$.

силлюминаторе при длине волны 312 нм (Vilber Lourmat transilluminator, Франция). Эффект ингибирования топо-I оценивали по способности тестируемых соединений задерживать релаксацию ссДНК. Эксперимент повторяли дважды.

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Синтез $DB_3(n)$

Схема синтеза димерных трисбензимидазолов $DB_3(n)$ представлена на рис. 4.

Мономерный трисбензимидазол MB_3 был синтезирован исходя из 5-(4-метилпиперазина-1-

ил)бензол-1,2-диамина (I) [13]. Он был сконденсирован с хлоргидратом имидоэфира (II), полученного по реакции Пиннера из соответствующего нитрила [14]. Полученное соединение (III) было восстановлено каталитическим гидрированием в диамин (IV), который при конденсации с имидоэфиром (II) образовывал нитроамин (V). При его гидрировании над палладиевым катализатором получен диамин (VI). Бисбензимидазол (VI) синтезирован в реакции диамина (VI) с Вос-Gly-OH методом смешанных ангидридов с последующим кипячением в уксусной кислоте с образованием N-ацетильного производного (VIII). При кипяче-

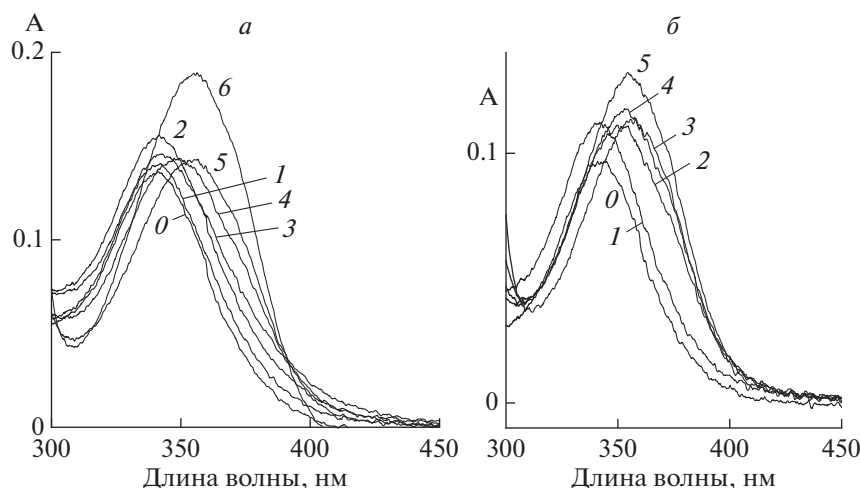


Рис. 5. *а* – Спектр поглощения **DB₃(1)** в отсутствие (1.38×10^{-6} М, кривая 0) и в присутствии (кривые 1–6) ДНК в концентрации 0.42 (1); 2.1 (2); 6.18 (3); 29.17 (4); 58.4 (5); 97.7 (6) $\times 10^{-6}$ М п.н. соответственно. Буфер 10 мМ PBS (рН 7.4), 22°C. Оптический путь 10 мм. *б* – Спектр поглощения **DB₃(9)** (1.4×10^{-6} М, кривая 0) и в присутствии (кривые 1–5) ДНК; концентрации ДНК 0.42 (1); 2.1 (2); 6.18 (3); 29.17 (4); 58.4 (5) $\times 10^{-6}$ М п.н. соответственно. Буфер 10 мМ PBS (рН 7.4), 22°C. Оптический путь 10 мм.

нии этого производного в концентрированной HCl получен мономерный трисбензимидазол **MB₃**. Димерные трисбензимидазолы **DB₃(n)** получены с выходом 89–92% при конденсации **MB₃** с α, ω -алкилдикарбоновыми кислотами (в соотношении 2 : 1) в присутствии HBTU. Все синтезированные соединения обладали хорошей флуоресценцией. Структура этих соединений подтверждена методами $^1\text{H-NMR}$ -спектроскопии и спектрометрии MALDI-TOF.

Физико-химические исследования

Взаимодействие димерных трисбензимидазолов **DB₃(1, 5, 9)** с ДНК охарактеризовано методами спектрофотометрии и спектрофлуориметрии. Зарегистрированы спектры поглощения **DB₃(1, 5, 9)** в присутствии дцДНК из спермы лосося (Деринат). На рис. 5 в качестве примера представлен спектр **DB₃(1)** и **DB₃(9)** в присутствии ДНК в разных концентрациях (кривые 1–6 и 1–5). При увеличении концентрации ДНК в растворе наблюдался небольшой сдвиг полосы поглощения в длинноволновую область спектра. Такой характер изменения спектров поглощения **DB₃(n)** свидетельствует об образовании комплексов **DB₃(n)**–ДНК.

Спектры флуоресценции **DB₃(1, 5, 9)** зарегистрированы в отсутствие и в присутствии дцДНК. Все они также имели схожий характер. На рис. 6 в качестве иллюстрации приведены спектры флуоресценции **DB₃(1)** и **DB₃(9)**. Добавление ДНК приводило к смещению положения максимума флуоресценции **DB₃(1)** от 486 к 470 нм и **DB₃(9)** от 486 к 472 нм, при этом интенсивность флуорес-

ценции **DB₃(1)** в присутствии ДНК увеличивалась почти в 15 раз.

Локализацию **DB₃(n)** в комплексе с ДНК определяли методом кругового дихроизма (КД), основанном на использовании холестерических жидкокристаллических ДНК (ХЖКД) (рис. 7).

Сопоставление кривых (рис. 7, заполненные кружки) показало, что в спектрах КД комплексов **DB₃(n)** с ХЖКД ДНК наблюдается появление не большой положительной полосы в области поглощения лигандов (300–400 нм). Положительные значения КД в области поглощения хромофоров **DB₃(n)** свидетельствуют о том, что угол между плоскостью хромофоров лиганда и спиральной осью молекулы ДНК составляет менее 54° , что возможно только в случае локализации данных хромофоров в одной из бороздок ДНК [16, 17]. Поскольку согласно данным рентгеноструктурного анализа Hoechst 33258 связывается с ДНК в узкой бороздке [18, 19], а молекулы **DB₃(n)** являются димерами производного Hoechst 33258, то мы полагаем, что они также располагаются именно в узкой бороздке ДНК, т.е. представляют собой узкобороздочные лиганды.

Биологические исследования: антивирусная активность **DB₃(n)** в отношении интегразы ВИЧ-1

Интегразы ВИЧ-1 осуществляет интеграцию кДНК вируса в геном человека и поэтому считается привлекательной мишенью для терапии инфекции ВИЧ-1 [20].

Ранее в качестве ингибиторов активности интегразы ВИЧ-1 мы изучали различные димерные

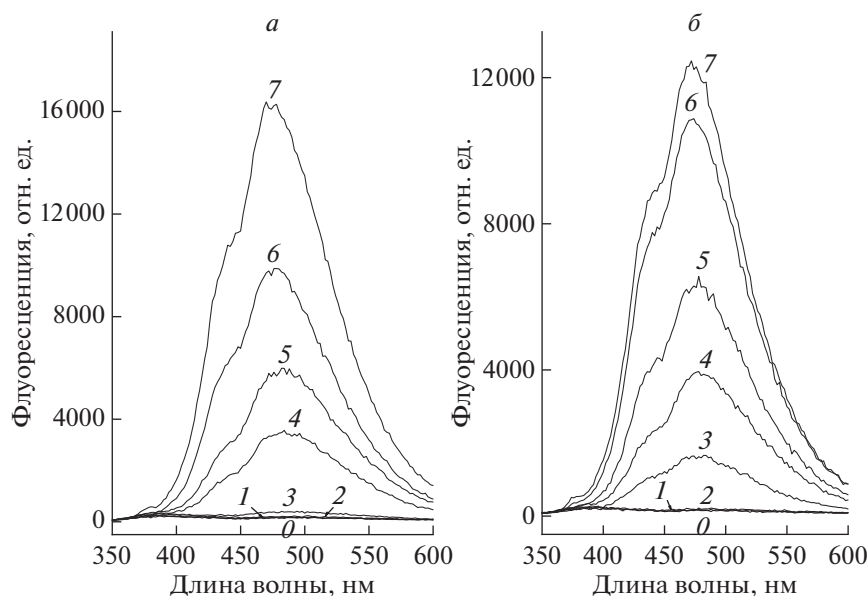


Рис. 6. *a* – Спектр флуоресценции буфера (0), **DB₃(1)** в отсутствие (1) и в присутствии ДНК (2–7). Концентрация **DB₃(1)** 1.34×10^{-6} М; концентрация ДНК: 1.0×10^{-6} (2); 2.1×10^{-6} (3); 41.8×10^{-6} (4); 0.21×10^{-5} (5); 0.42×10^{-5} (6); 0.84×10^{-5} М п.н. (7). Буфер 10 мМ PBS (pH 7.4), 22°C. $\lambda_{\text{exc}} = 340$ нм, оптический путь 10 мм. *б* – Спектр флуоресценции буфера (0), **DB₃(9)** в отсутствие (1) и в присутствии ДНК (2–7). Концентрация **DB₃(9)** 1.34×10^{-6} М; концентрация ДНК: 1.0×10^{-6} (2); 2.1×10^{-6} (3); 41.8×10^{-6} (4); 0.21×10^{-5} (5); 0.42×10^{-5} (6); 0.84×10^{-5} М п.н. (7). Буфер 10 мМ PBS (pH 7.4), 22°C. $\lambda_{\text{exc}} = 340$ нм, оптический путь 10 мм.

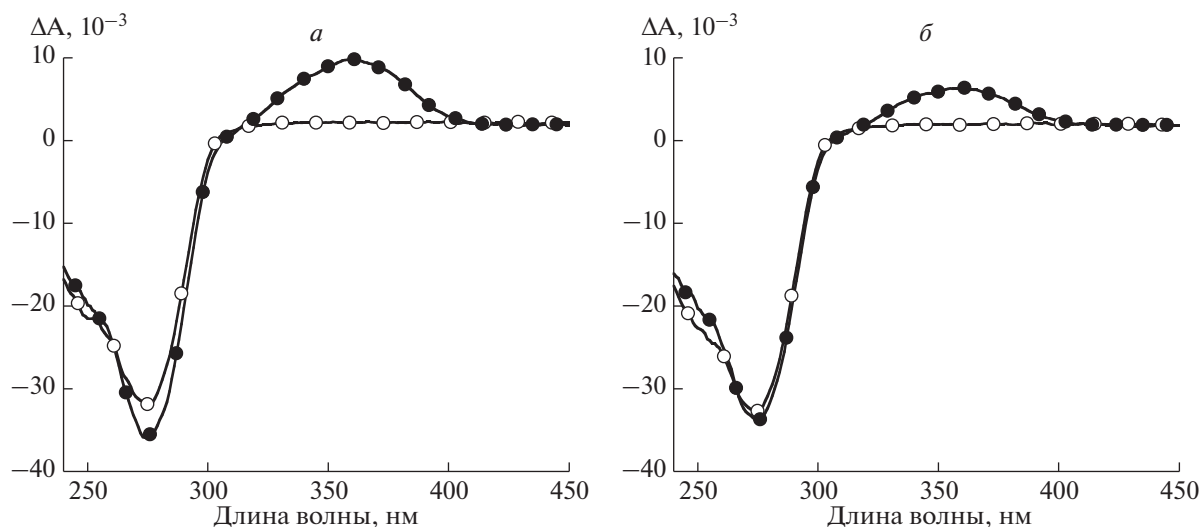


Рис. 7. Спектры КД холестерической жидкокристаллической дисперсии фрагментированной ДНК из спермы лосося в отсутствие (открытые кружки) и в присутствии **DB₃(1)** (*a*) и **DB₃(9)** (*б*) (заполненные кружки). Буфер 0.3 М NaCl + 0.002 М Na₃PO₄ pH 6.85 в присутствии ПЭГ (4000) 170 мг/мл. Длина оптического пути 10 мм. Концентрация ДНК 10^{-4} М п.н., концентрация **DB₃(1)** (*a*) и **DB₃(9)** (*б*) 4 мкМ.

производные бисбензимидазолов [1, 13, 14]. В настоящей работе мы решили расширить данные исследования и использовать синтезированные димерные трисбензимидазолы. В первую очередь

изучили активность мономерного трисбензимидазола **MB₃**. Оказалось, что это соединение способно ингибировать каталитическую активность интегразы ВИЧ-1 с $IC_{50} = 1.2$ мкМ. Димерные

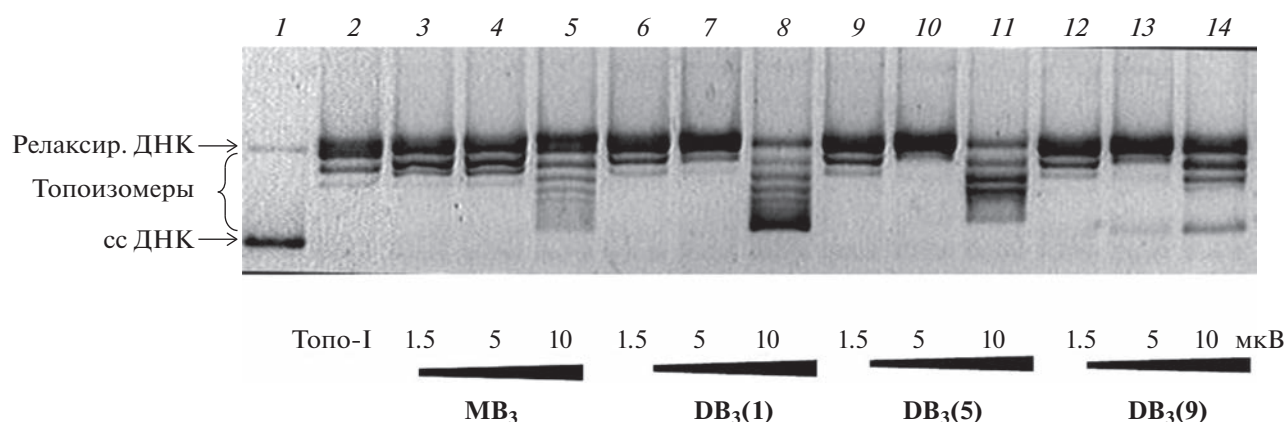


Рис. 8. Сравнение ингибирования топо-I мономерным трисбензимидазолом **MB₃** и тремя его димерными аналогами **DB₃(1)**, **DB₃(5)** и **DB₃(9)**.

трисбензимидазолы оказались еще более активными, чем **MB₃** (табл. 1). При этом соединение **DB₃(1)** с самым коротким метиленовым линкером между трисбензимидазольными фрагментами было примерно в 1.5 раза менее активным, чем соединения **DB₃(5)** и **DB₃(9)**, имеющие более длинные линкеры.

Полученные нами результаты интересно сравнить с данными работы [14], в которой нами исследовано влияние красителя Hoechst 33258 на активность интегразы. Данный недимеризованный бисбензимидазол Hoechst в отличие от мономерного трисбензимидазола **MB₃** оказался неактивным вплоть до концентрации 100 мкМ. При этом димеризация молекулы Hoechst 33258 с использованием метиленового линкера длиной пять или семь звеньев (**bis-Ht(5)** и **bis-Ht(7)**) приводила к появлению активности на уровне 3.2 и 2.6 мкМ соответственно. Таким образом, димерные трисбензимидазолы **DB₃(n)** оказались примерно в 10–15 раз более активными, чем димерные бисбензимидазолы **bis-Ht(n)** с аналогичной структурой. Более того, **DB₃(n)** обладают таким важным преимуществом, как более высокая растворимость в воде.

Таблица 1. Ингибирование реакции 3'-процессинга, катализируемой интегразой ВИЧ-1, использованными соединениями

Соединение	IC ₅₀ , мкМ
DB₃(1)	0.30 ± 0.07
DB₃(5)	0.19 ± 0.04
DB₃(9)	0.21 ± 0.06
MB₃	1.2 ± 0.3

*Биохимические эксперименты: **MB₃** и **DB₃(n)** ингибируют каталитическую активность эукариотической ДНК-топоизомеразы I в реакции релаксации суперскрученной ДНК*

Топо-I — одна из наиболее широко используемых мишеней ряда противоопухолевых средств. Способность соединений новой серии **DB₃(n)** ингибировать топо-I из тимуса теленка протестировали в бесклеточной системе. Реакцию релаксации ДНК топо-I проводили в присутствии исследуемых лигандов, взятых в низких (микромольных) концентрациях. В отсутствие ингибиторной активности топо-I приводит к полной релаксации плазмидной ссДНК (рис. 8, дорожка 2). Мономерный трисбензимидазол **MB₃** (рис. 2), содержащий в своей структуре только один трисбензимидазольный блок, практически не ингибировал топо-I.

Ингибирование топо-I мономером **MB₃** наблюдается, начиная с концентрации 10 мкМ, но на начальной стадии происходит образование топоизомеров без накопления ссДНК. Все новые димерные трисбензимидазолы обладали лучшей, чем у **MB₃**, способностью ингибировать топо-I, но различия в их эффективности были небольшими. При концентрации всех соединений 1, 5 и 5 мкМ ДНК находилась в релаксированной форме, что свидетельствует об отсутствии ингибирования топо-I. В случае димеров при небольшой разнице **DB₃(1)** в концентрации 10 мкМ лучше, чем **DB₃(5)** и **DB₃(9)**, ингибирует топо-I, о чем свидетельствует накопление ссДНК. Таким образом, наиболее функциональный размер линкера имеет **DB₃(1)**.

ЗАКЛЮЧЕНИЕ

Малые молекулы, образующие нековалентные комплексы, сайт-специфично связывающиеся с

узкой бороздкой ДНК, способны регулировать специфические эндогенные пути экспрессии генов [21]. Вероятно, регулирование работы гена происходит либо с прямой пространственной блокадой связывания ДНК-зависимого фермента или фактора транскрипции, либо за счет аллостерического изменения топологии ДНК. Исследование водорастворимых флуоресцентных димерных трисбензимидазолов **DB₃(n)** выявило широкий спектр их биологической активности. Будучи АТ-специфичными узкобороздочными лигандами, **DB₃(n)** проявили ингибиторную активность в отношении интегразы ВИЧ-1, ингибируя *in vitro* 3'-процессинг катализируемый интегразой ВИЧ-1 в субмикромольных концентрациях (0.20–0.30 мкМ). Кроме того, **DB₃(n)** способны ингибировать топо-I в бесклеточной системе из тимуса теленка в низких (микромольных) концентрациях. Мы полагаем, что новые АТ-сайт-специфичные соединения – димерные трисбензимидазолы **DB₃(n)** могут найти применение в молекулярной биологии и медицине.

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (№ 20-33-90287).

Настоящая статья не содержит описания выполненных авторами исследований с участием людей или использованием животных в качестве объектов.

Авторы заявляют об отсутствии конфликта интересов.

СПИСОК ЛИТЕРАТУРЫ

- Koval V.S., Arutyunyan A.F., Salyanov V.I., Kostyukov A.A., Melkina O.E., Zavilgelsky G.B., Klimova R.R., Kushch A.A., Korolev S.P., Agapkina Yu.Yu., Gottikh M.B., Vaiman A.V., Rybalkina E.Yu., Susova O.Yu., Zhuze A.L. (2020) DNA sequence-specific ligands. XVIII. Synthesis, physico-chemical properties; genetic, virological, and biochemical studies of fluorescent dimeric bisbenzimidazoles DBPA(n). *Bioorg. Med. Chem.* **28**, 115378.
- Bhaduri S., Ranjan N., Arya Dev P. (2018) An overview of recent advances in duplex DNA recognition by small molecules. *Beilstein J. Org. Chem.* **14**, 1051–1086.
- Rahman A., O'Sullivan P., Rozas I. (2019) Recent developments in compounds acting in the DNA minor groove. *Med. Chem. Comm.* **10**, 26–40.
- Battersby A.R. (1988) Synthetic and biosynthetic studies on vitamin B₁₂. *J. Nat. Prod.* **51**, 643–661.
- Gudmundsson K.S., Freeman G.A., Drach J.C., Townsend L.B. (2000) Synthesis of fluorosugar analogues of 2,5,6-trichloro-1-(β-D-ribofuranosyl)benzimidazole as antivirals with potentially increased glycosidic bond stability. *J. Med. Chem.* **43**, 2473–2478.
- Skalitzky D.J., Marakovits J.T., Maegley K.A., Ekker A., Yu X.H., Hostomsky Z., Webber S.E., Eastman B.W., Almasy R., Li J., Curtin N.J., Newell D.R., Calvert A.H., Griffin R.J., Golding B.T. (2003) Tricyclic benzimidazoles as potent poly(ADP-ribose) polymerase-1 inhibitors. *J. Med. Chem.* **46**, 210–213.
- He C.Y., Yang J., Wu B.-G., Risen L., Swayze E.E. (2004) Synthesis and biological evaluations of novel benzimidazoles as potential antibacterial agents. *Bioorg. Med. Chem. Lett.* **14**, 1217–1220.
- Seth P.P., Miyaji A., Jefferson E.A., Kristin S.-L., Osgood S.A., Propp S.S., Ranken R., Massire C., Sampath R., Ecker D.J., Swayze E.E., Griffey R.H. (2005) SAR by MS discovery of a new class of RNA-binding small molecules for the hepatitis C virus: internal ribosome entry site IIA subdomain. *J. Med. Chem.* **48**, 7099–8102.
- Tanious F.A., Hamelberg D., Bailly C., Czarny A., Boykin D.W., Wilson W.D. (2004) DNA sequence dependent monomer–dimer binding modulation of asymmetric benzimidazole derivatives. *J. Am. Chem. Soc.* **126**, 143–153.
- Bailly C., Chessari G., Carrasco C., Joubert A., Mann J., Wilson W.D., Neidle S. (2003) Sequence-specific minor groove binding by bis-benzimidazoles: water molecules in ligand recognition. *Nucl. Acids Res.* **31**, 1514–1524.
- Keri R.S., Hiremathad A., Budagumpi S.I., Nagaraja B.M. (2015) Comprehensive review in current developments of benzimidazole-based medicinal chemistry. *Chem. Biol. Drug Des.* **86**, 19–65.
- Latt S.A., Stetten G., Juergens L.A., Buchanan G.R., Gerald P.S. (1975) Comprehensive review in current developments of benzimidazole-based medicinal chemistry. *J. Histochem. Cytochem.* **23**, 493–505.
- Громыко А.В., Саянов В.И., Стрельцов С.А., Олейников В.А., Королев С.П., Готтих М.Б., Жузе А.Л. (2007) Лиганды, специфичные к определенным последовательностям пар оснований ДНК. XIII. Новые димерные молекулы Хехста 33258 – ингибиторы интегразы ВИЧ-1 *in vitro*. *Биоорганическая химия*. **33**, 613–623.)
- Королев С.П., Ташлицкий В.Н., Смолов М.А., Громыко А.В., Жузе А.Л., Агапкина Ю.Ю., Готтих М.Б. (2010) Ингибирование интегразы ВИЧ-1 димерными бисбензимидазолами с различной структурой линкера. *Молекулярная биология*. **44**, 718–727.
- Leh H., Brodin P., Bischerour J., Deprez E., Tauc P., Brochon J.C., LeCam E., Coulaud D., Auclair C., Mouscadet, J.F. (2000) Determinants of Mg-dependent activities of recombinant human immunodeficiency virus type 1 integrase. *Biochemistry*. **39**, 9285–9294.
- Евдокимов Ю.М. Саянов В.И., Нечипуренко Ю.Д., Скуридин С.Г., Захарова М.А., Спенер Ф., Палумбо М. (2003) Молекулярные конструкции (суперструктуры) с регулируемыми свойствами на основе двухцепочечных нуклеиновых кислот. *Молекулярная биология*. **37**, 340–355.
- Yevdokimov Yu.M., Skuridin S.G., Nechipurenko Y.D., Zakharov M.A., Salyanov V.I., Kurnosov A.A.,

- Kuznetsov V.D., Nikiforov V.N. (2005) Nanoconstructions based on double-stranded nucleic acids. *Int. J. Biol. Macromol.* **36**, 103–115.
18. Pjura P.E., Grzeskowiak K., Dickerson R.E. (1987) Binding of Hoechst 33258 to the minor groove of B-DNA. *J. Mol. Biol.* **197**, 257–271.
19. Teng M.K., Usman N., Frederick C.A., Wang A.H. (1988) The molecular structure of the complex of Hoechst 33258 and the DNA dodecamer d(CGC-GAATTCGCG). *Nucl. Acids Res.* **16**, 2671–2690.
20. Pommier Y., Johnson A.A., Marchand C. (2005) Integrase inhibitors to treat HIV/AIDS. *Nat. Rev. Drug Discov.* **4**, 236–248.
21. Leung C.H., Chan D.S.-H., Ma V.P.-Y., Ma D.-L. (2013) DNA-binding small molecules as inhibitors of transcription factors. *Med. Res. Rev.* **33**, 823–846.

DNA Sequence-Specific Ligands. XIX. Synthesis, Spectral Properties, Virological and Biochemical Studies of Fluorescent Dimeric Trisbenzimidazoles $\text{DB}_3(n)$

A. F. Arutyunyan¹, A. A. Kostyukov², S. P. Korolev^{3, 4}, M. B. Gottikh^{3, 4},
O. Yu. Susova⁵, D. N. Kaluzhny¹, and A. L. Zhuze^{1, *}

¹Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Moscow, 119991 Russia

²Emanuel Institute of Biochemical Physics, Russian Academy of Science, Moscow, 119334 Russia

³Department of Chemistry, Lomonosov Moscow State University, Moscow, 119991 Russia

⁴Belozersky Institute of Physicochemical Biology, Lomonosov Moscow State University, Moscow, 119991 Russia

⁵Research Institute of Carcinogenesis, Blokhin National Medical Research Center of Oncology, Moscow, 115478 Russia

*e-mail: zhuze@eimb.ru

In this work, we synthesized and characterized the properties of a series of new fluorescent narrow-groove ligands $\text{DB}_3(n)$. $\text{DB}_3(n)$ compounds based on dimeric trisbenzimidazoles have the ability to bind to the AT regions of DNA. The synthesis of $\text{DB}_3(n)$, trisbenzimidazole fragments of which are linked by oligomethylene linkers of different lengths ($n = 1, 5, 9$), is based on the condensation of monomeric trisbenzimidazole MB_3 with α, ω -alkyldicarboxylic acids. $\text{DB}_3(n)$ proved to be effective inhibitors of the catalytic activity of HIV-1 integrase at submicromolar concentrations (0.20–0.30 μM). $\text{DB}_3(n)$ was found to inhibit the catalytic activity of DNA topoisomerase I at low micromolar concentrations.

Keywords: DNA, sequence-specific minor-groove ligand, fluorescent dimeric trisbenzimidazoles $\text{DB}_3(n)$, spectral methods, HIV-1 integrase, DNA topoisomerase I

СТРУКТУРНО-ФУНКЦИОНАЛЬНЫЙ АНАЛИЗ БИОПОЛИМЕРОВ И ИХ КОМПЛЕКСОВ

УДК 577.323.3

ОКИСЛЕНИЕ ПОРФИРИНОМ ZnP1 G4-СТРУКТУР В ПОСЛЕДОВАТЕЛЬНОСТЯХ ПРОМОТОРОВ ГЕНОВ *MYC* И *TERT*

© 2023 г. Г. В. Чашина^а, Д. Н. Калужный^{а, *}

^аИнститут молекулярной биологии им. В.А. Энгельгардта Российской академии наук,
Москва, 119991 Россия

*e-mail: uzhny@mail.ru

Поступила в редакцию 10.11.2022 г.

После доработки 04.12.2022 г.

Принята к публикации 06.12.2022 г.

Формирование G4-структур в двойной спирали ДНК конкурирует с комплементарной нитью. Локальное окружение в ДНК может изменять равновесные G-квадруплексы, изучаемые на одонитевых моделях. Для обнаружения квадруплексов в промоторных областях генома необходимы методы их локализации в составе протяженных ДНК природной последовательности. Порфириновое производное ZnP1 избирательно связывается и приводит к фотоиндуцируемому окислению гуанина в G-квадруплексах в модельных системах с одноцепочечной и двухцепочечной ДНК. На примере нативной последовательности промоторных участков онкогенов *MYC* и *TERT* нами показано окислительное действие ZnP1 в областях, потенциально способных формировать G4-структуры. Установлены и отнесены к нуклеотидной последовательности одонитевые разрывы в G-богатой нити двойной спирали, вызванные окислением гуанина под действием ZnP1 и последующим расщеплением цепи ДНК Fpg гликозилазой. Обнаруженные сайты разрывов соответствовали последовательностям, потенциально способным формировать G-квадруплексы. Таким образом, продемонстрирована возможность использования порфирина ZnP1 для идентификации и локализации G-квадруплексов в составе протяженных участков генома. Получены новые данные о возможности сворачивания G4-структур в присутствии комплементарной нити в контексте двойной спирали ДНК природной последовательности.

Ключевые слова: G-квадруплексы, порфирин, фотосенсибилизаторы, окисление ДНК

DOI: 10.31857/S0026898423030035, **EDN:** CGYDDDB

Богатые гуанином последовательности ДНК могут складываться в неканонические вторичные структуры, называемые G-квадруплексами (G4). Появляется все больше доказательств, что G4-структуры нуклеиновых кислот участвуют в различных биологических процессах, таких как транскрипция, репликация и трансляция; а также влияют на стабильность генома и обеспечивают контроль экспрессии генов [1–3]. В результате полногеномного секвенирования 12 видов организмов выявлено более 3.5 млн участков ДНК, потенциально способных образовывать G4-структуры. Сильное обогащение такими последовательностями выявлено в промоторных областях и точках начала транскрипции у высших организмов: например, человека и мыши [4, 5].

Много данных по стабильности и конформации G4-структур получено биофизическими, спектральными и структурными методами (ядерный магнитный резонанс (ЯМР) и рентгеноструктурный анализ (РСА)). Информация о G4-струк-

турах и их комплексах с лигандами систематизирована в различных базах данных [6]. Традиционно в таких исследованиях используют относительно короткие синтетические одно- (оц) или двухцепочечные (дц) фрагменты ДНК. Внесение мутаций в модельные олигонуклеотиды позволяет с высокой достоверностью охарактеризовать одну G4-структуру. Главный недостаток таких методов и подходов — отсутствие информации о влиянии локального окружения. Неизвестно, какое влияние оказывает комплементарная нить при формировании G4-структур в промоторных областях генома. Природные последовательности могут содержать несколько перекрывающихся G4-участков, способных образовывать множество вариантов складывания неканонических структур. На текущий момент до конца не изучено, какие именно G4-структуры образуются в присутствии комплементарной нити ДНК. Более термодинамически стабильные G4-структуры исключают образование менее стабильных [7]. Однако до-

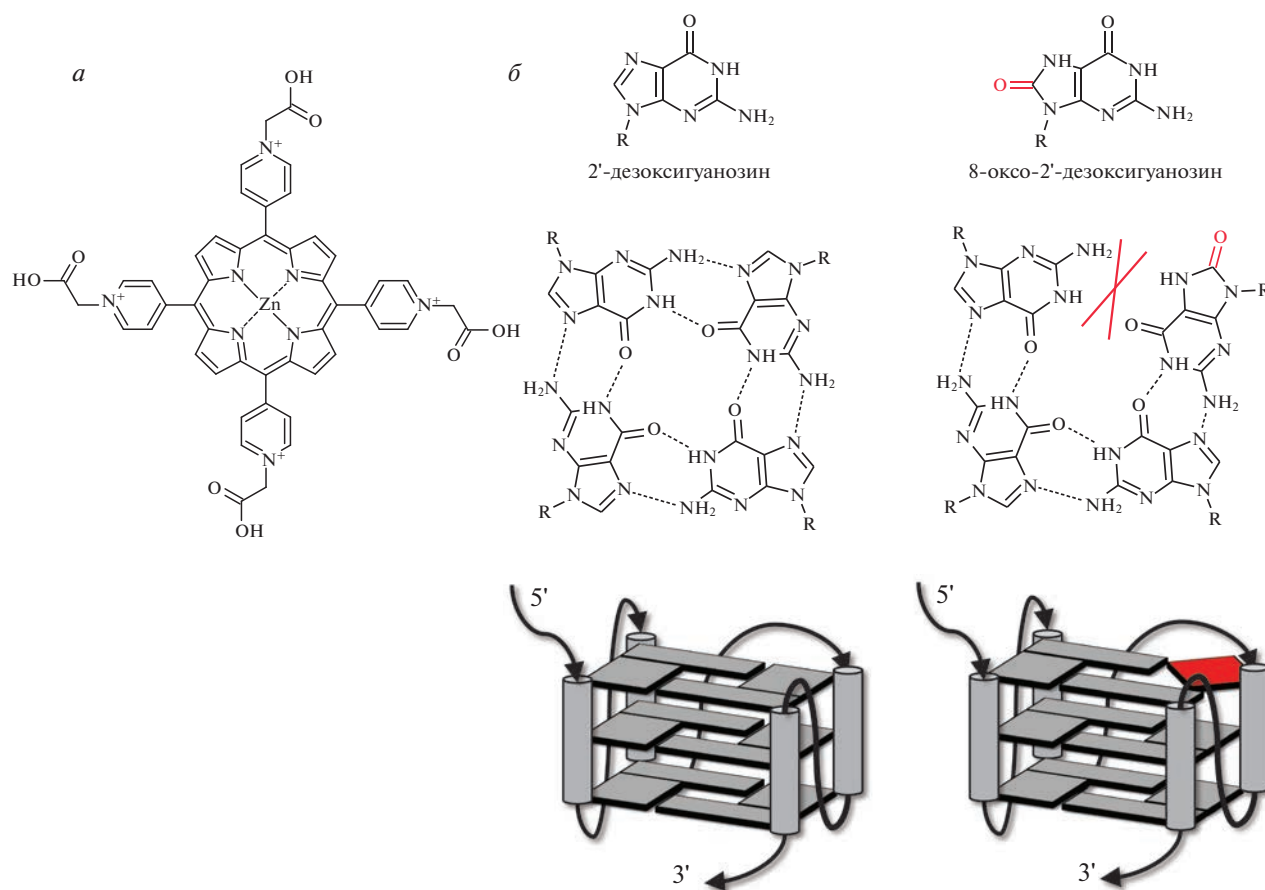


Рис. 1. Схематическое изображение исследуемых веществ. *a* – Структура порфирина ZnP1. *б* – Окисление одного из гуанинов в составе G-квадруплекса приводит к дестабилизации структуры в целом. R – дезоксирибоза.

полнительные взаимодействия между структурами внутри протяженной дцДНК могут изменять стабильность G4-структуры, образованной одной последовательностью, в пользу соседней.

Разработка методов обнаружения неканонических структур и их локализация в протяженных дцДНК особенно актуальна для исследования участков генома, потенциально вовлеченных в регуляцию биологических процессов. Классическим подходом к химическому зондированию ДНК, в том числе G4-структур, считается модификация с использованием диметилсульфата (DMS) [8]. Гуанины, вовлеченные в образование G4-структур, не доступны для модификации DMS. Использование этой методики позволяет обнаружить стабильные участки, полностью образующие структуры ДНК, альтернативные двойной спирали. При частичном или динамическом формировании G4-структур в оцДНК или дцДНК не удается регистрировать короткоживущие альтернативные структуры [9]. Для детекции таких динамически образующихся G4-структур необходимы агенты, направленные именно на неканонически свернутые нуклеиновые кислоты.

Использование порфиринов в качестве структурного зонда имеет ряд преимуществ перед DMS-зондированием. Фотоиндуцируемое окисление ДНК под действием порфиринов способно распознать даже незначительную фракцию G4-структур и слабо чувствительно к дцДНК [10, 11]. Порфириновое производное ZnP1 (рис. 1*a*) связывается с малой бороздкой ДНК [12], при этом аффинность ZnP1 выше к G4-структурам, чем к дцДНК. Ранее показано [10], что связывание ZnP1 с ДНК минимально искажает G4-структуру и поэтому может быть использовано как структурный зонд. Под воздействием синего света ZnP1 окисляет азотистое основание гуанин, участвующее в образовании G4-структуры. Окисление гуанина до 8-оксогуанина и других производных приводит к разрушению хугстиновской связи между двумя основаниями, участвующими в образовании квартета (рис. 1*б*), что нарушает его структуру и снижает стабильность. Это приводит к конформационному переходу G-квадруплекса в двойную спираль ДНК, как показано на олигонуклеотидных моделях с одной G4-структурой в дуплексе [13]. Учитывая эти свойства, мы выбра-

Таблица 1. Последовательности праймеров

Ген	Праймеры для амплификации ^a , 5'→3'	Флуоресцентно меченные праймеры ^b , 5'→3'
<i>MYC</i>	Forward: GAGGAGCAGCAGAGAAAGGG Reverse: TCCCTCCGTTCTTTTCCCG	FAM-TCCTAGAGCTAGAGTGCTCGG
<i>TERT</i>	Forward: GGCCGATTCGACCTCTCT Reverse: CAGCGCTGCCTGAAACTC	FAM-CTTCCAGCTCCGCCTCCTCC

^aForward — прямой праймер; Reverse — обратный праймер.

^bFAM — флуоресцеин.

ли соединение ZnP1 с высоким окислительным потенциалом, большой избирательностью связывания, при этом минимально искажающим структуру ДНК [10, 12, 14].

В представленной работе мы исследовали длинные участки (~400 п.н.) промоторов двух генов человека: *MYC* и *TERT*. В их состав входит несколько последовательностей, потенциально образующих G4-структуры. Для промотора *TERT* в литературе описаны модели, предполагающие формирование трех следующих друг за другом параллельных квадруплексов [15–17]. Для промотора *MYC* человека также известно несколько G4-структур, описанных в литературе [18–20].

Таким образом, целью данной работы было определение положения G4-структур, образующихся в контексте двойной спирали ДНК, в участках промоторов *MYC* и *TERT*. С помощью производного порфирина ZnP1 мы провели специфическое окисление гуанинов G4-структур и установили локализацию окисленных форм гуанинов методом удлинения праймера.

ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

Олигонуклеотиды. Праймеры для ПЦР были синтезированы в ЗАО “Евроген” (Россия), флуоресцентно меченные праймеры — в ООО “ДНК-синтез” (Россия). Последовательности праймеров представлены в табл. 1.

ПЦР-амплификация. Амплификацию проводили по ранее описанной методике [11]. Стандартный буфер для проведения ПЦР был заменен на не содержащий ионы калия — для снижения стабильности потенциальных G4-структур в матричной цепи ДНК.

Очистка ПЦР-продуктов. К реакционной смеси объемом 50 мкл добавляли 200 мкл 0.3 М ацетата натрия (pH 5.3) и 600 мкл 96%-ного этанола и оставляли при –20°C на 30 мин, после чего осадок отделяли центрифугированием при 14000 × g в течение 10 мин. После промывания 70%-ным этанолом образцы растворяли в воде. Для удаления праймеров использовали набор магнитных частиц CleanMag DNA (ЗАО “Евроген”) в соответствии с рекомендациями фирмы-производите-

ля. Концентрацию очищенных образцов ДНК определяли на спектрофотометре NanoDrop1000 (“Thermo Fisher Scientific”, США) при длине волны 260 нм.

Формирование G4-структур. Полученные ПЦР-продукты в концентрации 10 нМ отжигали (нагревали 5 мин при 95°C и медленно охлаждали при комнатной температуре в течение 12 ч) в 10 мМ фосфатном буфере (pH 7.8), содержащем 100 мМ KCl и 40% ПЭГ₂₀₀. Формирование квадруплексов в полученных образцах регистрировали по изменению электрофоретической подвижности в 8%-ном ПААГ.

Фотоиндуцируемое разрушение G4-структур. Порфирин ZnP1 (рис. 1a) был получен по ранее описанной методике [12, 21]. После отжига к образцам добавляли порфирин ZnP1 до конечной концентрации 1 мкМ, облучали синим светом (465 нм) в течение 5 мин для специфического окисления гуаниновых оснований, образующих G4-структуру. Анализ окисленных G-квадруплексных структур проводили в нативном 8%-ном ПААГ. Гель окрашивали красителем SYBR Gold (“Invitrogen”, США) и визуализировали на флуоресцентном сканере Typhoon 9500 (“GE Healthcare”, США).

Внесение одонитевых разрывов в местах окисленных форм гуанина. К облученному образцу объемом 25 мкл добавляли 1 U формамидопиримидин-ДНК-*N*-гликозилазы (Fpg) *Escherichia coli* (“NEB”, США), 1× NEB buffer1 и инкубировали 30 мин при 37°C. Реакцию останавливали прогреванием при 70°C в течение 10 мин. ДНК осаждали спиртом (как написано выше) и растворяли в 5 мкл воды.

Удлинение флуоресцентно меченного праймера. Реакцию удлинения праймера проводили в 25 мкл смеси, содержащей 20 мМ Трис-HCl (pH 9.5), 2.5 мМ MgCl₂, 5% DMSO, 20 мМ CsCl, 0.1 мМ dNTPs, 0.5 U SNPdetect-полимеразы (ЗАО “Евроген”), 5 мкл ДНК-матрицы, 100 нМ флуоресцентно меченный праймер.

Реакционная смесь для секвенирования по Сэнгеру (25 мкл) содержала те же компоненты и 2 нМ ПЦР-продукт в качестве ДНК-матрицы.

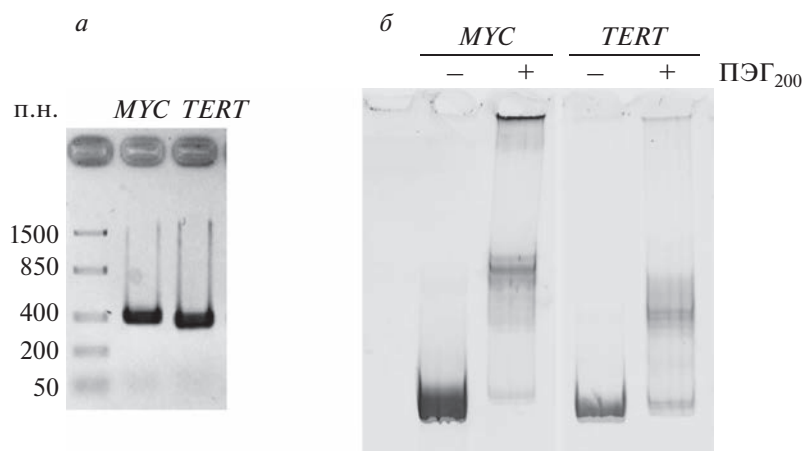


Рис. 2. Электрофоретическая подвижность изучаемых образцов. *a* – ПЦР-продукт *MYC* (401 п.н.) и *TERT* (378 п.н.). *б* – Отжиг ПЦР-продуктов *MYC* и *TERT* в присутствии ПЭГ₂₀₀.

В реакционную смесь вносили ddG, ddA, ddT или ddC до конечной концентрации 0.1, 0.1, 0.1 и 0.01 мкМ соответственно. Полученные реакционные смеси инкубировали при 95°C в течение 3 мин, а затем проводили 20 циклов амплификации: денатурация 30 с при 95°C, отжиг в течение 30 с, удлинение праймера при 72°C в течение 60 с. Температуру отжига подбирали для каждого праймера с помощью программы OligoAnalyzer Tool (www.idtdna.com) и корректировали опытным путем.

По окончании реакции в каждую смесь вносили по 200 мкл 0.3 М ацетата натрия (pH 5.3) и 600 мкл 96%-ного этанола и оставляли при –20°C на 30 мин, после чего осадок отделяли центрифугированием при 14000 × *g* в течение 10 мин. Осадок промывали 70%-ным спиртом, сушили при комнатной температуре и растворяли в буфере, содержащем 30 мМ Na-фосфат (pH 7.8), 80% формамида, 0.025% бромфенолового синего и 0.025% ксиленианола. Денатурацию проводили при 95°C в течение 5 мин и наносили смесь на 10%-ный ПААГ, содержащий 7 М мочевины и 1× TBE-буфер. Гель анализировали на флуоресцентном сканере Typhoon 9500 (“GE Healthcare”). Для построения денситометрических профилей использовали программу ImageJ.

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Формирование квадруплексных структур в протяженной дцДНК

Для изучения G4-структур в протяженной дцДНК мы выбрали промоторные участки онкогенов человека *MYC* и *TERT*, содержащие последовательности, потенциально способные образовывать квадруплексы. Объектом исследования служили фрагменты длиной 401 п.н. (*MYC*) и

378 п.н. (*TERT*), полученные ПЦР-амплификацией с геномной ДНК человека (рис. 2*a*). Участки ДНК были выбраны так, чтобы положение потенциальных G4-структур приходилось на центральную часть двойной спирали и исключало краевые эффекты от конформационных перестроек ДНК при формировании квадруплекса. Амплификацию проводили в условиях, подобранных в работе Г. Чашчиной (G. Chashchina) и соавт. [11], в которой было показано, что ионы калия в стандартных буферах термофильных полимераз затрудняют синтез участков ДНК, матричная нить которых содержит потенциальные G4-последовательности.

ПЦР-продукты всегда представляют собой дцДНК независимо от наличия или отсутствия квадруплексных участков вследствие последовательного синтеза с образованием классических пар Уотсона–Крика. Для формирования альтернативных структур в дцДНК необходимо преодолеть энергетический и кинетический барьер конформационного перехода, что предполагает расплетание двойной спирали.

Для имитации клеточных условий часто используют инертные наполнители, например фикоколл, декстран или ПЭГ. Последний имитирует клеточную среду, так называемые условия “молекулярной скученности”. При количественном анализе влияния молекулярной скученности на термодинамику образования G-квадруплексов было показано, что G4-структура стабилизируется в условиях молекулярного скопления. Это коррелирует с понижением активности воды и, как следствие, приводит к дестабилизации двойной спирали ДНК, образованной парами Уотсона–Крика [22].

Zheng с соавт. [23] показали, что отжиг G4-последовательностей в составе двойной спирали в присутствии ПЭГ₂₀₀ индуцирует образование

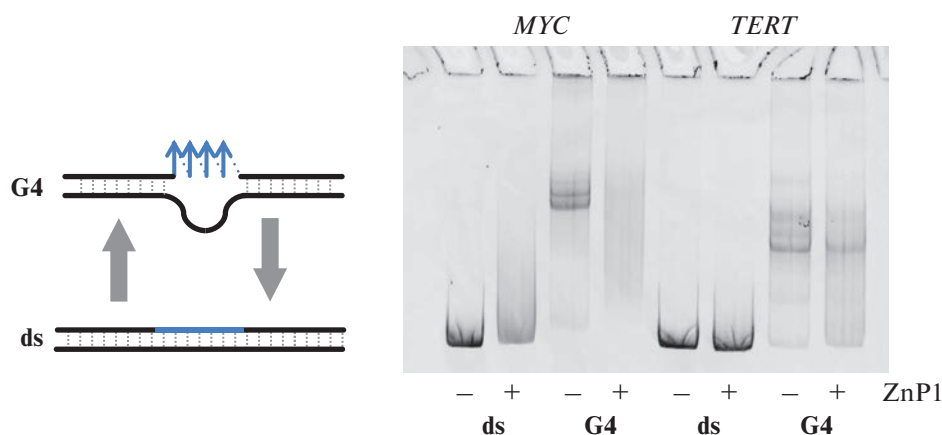


Рис. 3. Отжиг ПЦР-продуктов промоторных областей *MYC* (401 п.н.) и *TERT* (378 п.н.) в присутствии ПЭГ₂₀₀ (а) и последующее окисление полученных дцДНК (ds) порфирином ZnP1 (б).

квадруплексов в контексте дцДНК. В условиях молекулярной скученности G-квадруплексы не только образуются, но и сохраняют стабильную структуру [23].

Ранее нами предложен альтернативный способ формирования G4-структуры в дуплексе, основанный на стабилизации лигандом пиридогистамина (PDS) [13]. Благодаря планарной структуре PDS преимущественно связывается с крайними гуаниновыми квартетами, приводя к значительной стабилизации G4 [24]. PDS обладает низкой аффинностью к двойной спирали ДНК и при связывании сдвигает энергетический баланс в сторону G4. Ранее нам удалось получить стабильные структуры G4 в дуплексе [13] и показать, что комплекс G4 с PDS в меньшей мере подвержен окислительному действию из-за стерических препятствий при связывании ZnP1, то есть PDS защищает гуанины от окисления. Таким образом, в данной работе для формирования G4-структур в составе промоторов двух онкогенов мы использовали отжиг в присутствии ПЭГ₂₀₀.

Медленным отжигом в буфере, содержащем 100 мМ KCl и 40% ПЭГ₂₀₀, были сформированы G4-структуры из очищенных протяженных дцДНК (ПЦР-продуктов). Образование альтернативного сворачивания дцДНК и образования G4-структуры детектировали по снижению электрофоретической подвижности в ПААГ относительно нативных ПЦР-продуктов (рис. 2б). Меньшая подвижность в геле соответствует образованию G4 в дуплексе [13, 23]. Для промоторного участка *MYC* обнаружено, что часть образца соответствует высокомолекулярному комплексу и находится вблизи лунки для нанесения образца. Возможная агрегация G4-структур, сформированных в дуплексе, ранее обсуждалась и была подтверждена методом атомно-силовой микроскопии [25]. Стоит отметить, что такая агрегация образца при от-

жиге не воспроизводится в нескольких повторах, хотя положение полосы с меньшей подвижностью относительно дуплекса воспроизводимо для образцов *MYC* и *TERT*, что может отражать разное количество и разные конформации G4-структур в составе дцДНК выбранных промоторных участков.

Изменение конформации ДНК при фотоиндуцированном окислении G4-структур порфирином ZnP1

Полученные отжигом G4-структуры инкубировали с порфирином ZnP1 и подвергали облучению синим светом. Анализ подвижности в ПААГ показал, что происходит частичное разрушение квадруплексной структуры, о чем свидетельствует уменьшение или размывание полосы с меньшей подвижностью относительно дцДНК (рис. 3). Этот эффект может быть объяснен дестабилизацией G4-структуры из-за окисления гуанина порфирином ZnP1. Эффект перехода G4 в двойную спираль оказался сильнее для последовательности промотора *MYC*, чем для *TERT*. Мы предполагаем, что это может быть связано с незначительной специфичностью ZnP1 к определенной последовательности, конформационным особенностям G4-структуры или недоступностью некоторых квадруплексов для взаимодействия с порфирином. Известно, что дцДНК подвержена окислению ZnP1 в меньшей степени, чем квадруплекс, — аналогично тем случаям, которые наблюдались на коротких олигонуклеотидах [10, 13].

Определение позиции окисленных форм гуанина в нуклеотидной последовательности

Важный аспект в исследовании конформации ДНК посредством таргетного окисления G4-структур — обнаружение образующихся

продуктов. Ранее сайты окисления на коротких ДНК-дуплексах детектировали по электрофоретической подвижности (в денатурирующих условиях — с 7 М мочевиной) оцДНК, полученных после кипячения в формамиде [13]. В случае протяженной ДНК этот метод не применим вследствие возникающих затруднений при разделении схожих по длине фрагментов ДНК.

Здесь мы предлагаем использовать метод удлинения праймера для локализации одонитевых разрывов ДНК на участках, потенциально способных формировать G4-структуры. Для получения одонитевых разрывов на месте окисленных форм гуанина исследуемые образцы ДНК подвергали окислению ZnP1 и обрабатывали гликозилазой *E. coli* Fpg. Используя метод удлинения праймера, мы картировали позицию полученных разрывов. Появление модифицированных оснований гуанина детектировали по неполному синтезу комплементарной цепи ДНК, так как ДНК-полимераза останавливается на разрывах в матричной цепи (рис. 4). Сайты остановок ДНК-полимеразы обнаружены только на предварительно отоженных и облученных образцах. Удлинение праймера на образцах облученных дуплексов, а также квадруплексных структурах, не окисленных порфирином, происходило без обрыва цепи.

Таким образом, обработка промоторных участков генов *MYC* и *TERT* порфирином ZnP1 приводит к модификации ДНК в определенных местах. Точные сайты разрывов получены соотношением длины укороченных продуктов полимеразного синтеза к последовательности матричной нити ДНК. В качестве маркера длин использовали электрофореграмму секвенирования методом Сэнгера. Денситометрический профиль предварительно отоженной G4-структуры и остановок ДНК-полимеразы на окисленных образцах дцДНК приведен на рис. 5.

На последовательности промотора гена *MYC* ДНК-полимераза останавливалась перед участками, структура которых была детально изучена биофизическими методами на одонитевых фрагментах. Наиболее выраженная остановка происходила на GG-динуклеотиде, входящем в 3'-последовательность олигонуклеотида ru24, G4-структура которого в комплексе с порфирином TMRyP4 (PDB ID 2A5R) была исследована методом ЯМР [20]. В составе этого комплекса окисляемый динуклеотид находится в противоположном квартете от связанной молекулы порфирина. В наблюдаемом нами случае, то есть в составе двойной спирали, окисление образованной G4-структуры может быть обусловлено большей доступностью нуклеотидов к растворителю по сравнению с другими гуанинами, либо отличным от TMRyP4 сайтом посадки фотосенсибилизатора ZnP1. Стоит отметить, что фотоокисление в мутированной последовательности промотора гена *MYC* [13], способной образовывать только одну форму G4, имело специфичность к 5'-нуклеотидам каждого из четырех участков подряд идущих гуанинов, образующих G4-структуру. Мы предполагаем, что наличие нескольких гуаниновых участков вблизи известных G4-структур может изменять характер сворачивания ДНК при формировании стабильных одонитевых структур. Возможно, в природной последовательности с несколькими перекрывающимися G4-участками образуются структуры, отличные от изученных ранее.

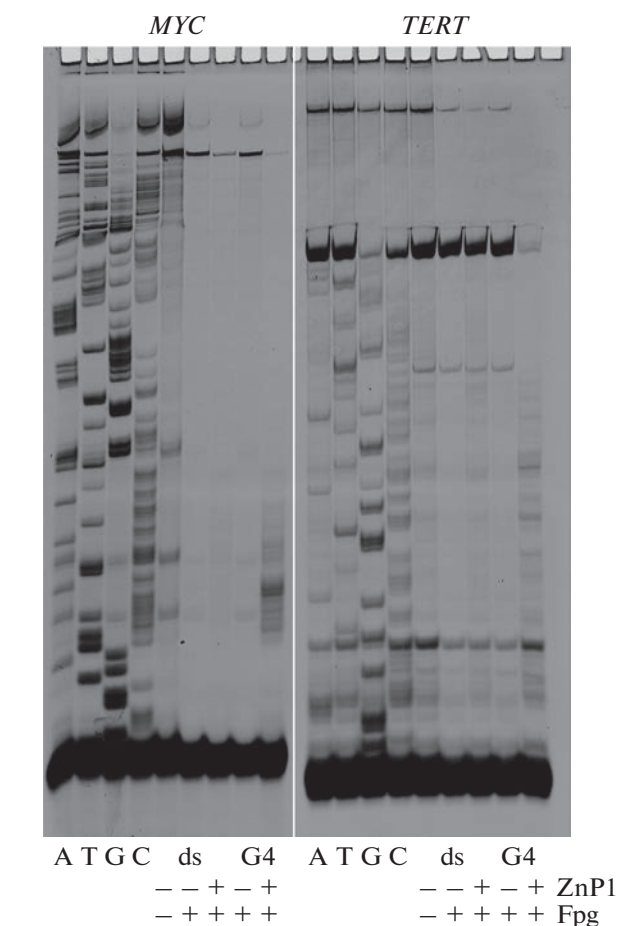


Рис. 4. Продукты удлинения флуоресцентно меченого праймера на квадруплексных последовательностях в промоторных участках *MYC* и *TERT*. Для определения точных позиций, на которых останавливается ДНК-полимераза, проводили секвенирование ДНК методом Сэнгера (первые четыре дорожки). Электрофорез проводили в 10%-ном ПААГ в денатурирующих условиях (7 М мочевины).

затора ZnP1. Стоит отметить, что фотоокисление в мутированной последовательности промотора гена *MYC* [13], способной образовывать только одну форму G4, имело специфичность к 5'-нуклеотидам каждого из четырех участков подряд идущих гуанинов, образующих G4-структуру. Мы предполагаем, что наличие нескольких гуаниновых участков вблизи известных G4-структур может изменять характер сворачивания ДНК при формировании стабильных одонитевых структур. Возможно, в природной последовательности с несколькими перекрывающимися G4-участками образуются структуры, отличные от изученных ранее.

В последовательности промотора гена *TERT* методом ЯМР изучен только один участок, отмеченный на рис. 5 как PQS1 [17], находящийся в

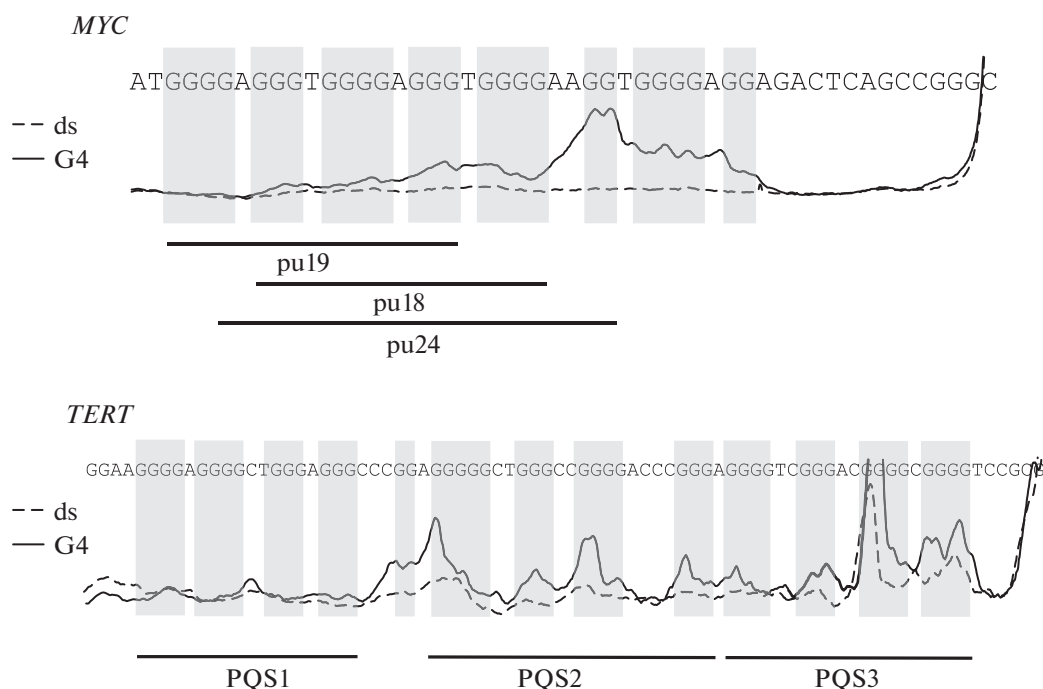


Рис. 5. Денситометрический профиль позиций, на которых останавливается ДНК-полимераза при удлинении праймера, отнесенный к нуклеотидным последовательностям *MYC* и *TERT*. Известные G4-последовательности с подтвержденными структурами квадруплексов [16, 20] показаны черными линиями.

равновесии между параллельной и 3+1 конформацией G4. R. Monsen и соавт. [16] предложили модель формирования G4-структур промоторной последовательности гена *TERT*. Используя ряд биофизических методов, авторы показали, что эти структуры образованы тремя взаимодействующими друг с другом квадруплексами. При анализе позиций окисленных форм гуанина в сформированной G4-структуре выявили, что наиболее эффективно окисление происходит на 5'-нуклеотидах подряд идущих гуанинов, которые входят в состав последовательности PQS2 (рис. 5). Такая локализация окисленных оснований соответствует формированию G4-структуры с параллельной ориентацией гуаниновых участков, что соотносится с предложенной ранее структурой участка PQS2 из промоторной области гена *TERT* [14].

Используя порфирин ZnP1 для окисления гуанинов в составе квадруплексов, мы подтвердили возможность образования, по крайней мере, одной доступной для окисления G4-структуры — PQS2 — в промоторе гена *TERT*. Именно эта область промотора наиболее подвержена мутациям C → T или G → A [26], что позволяет предположить участие окисления G4-структур в появлении мутаций в промоторных областях.

Нами продемонстрирована возможность идентификации G4-структур в протяженной дцДНК

при помощи отжига в присутствии ПЭГ₂₀₀ и ионов калия. Впервые показана возможность окислительного действия порфирина ZnP1 на G4-структуры, образующиеся в участках промоторов генов *MYC* и *TERT*. Методом удлинения праймера определена локализация окисленных форм гуанина. Позиции окисленных форм гуанина в промоторе гена *TERT* соответствуют ранее изученным G4-структурам. В длинных природных последовательностях дцДНК возможно альтернативное сворачивание G4-структур, вызванное наличием комплементарной нити и дополнительных повторов гуанина, что было показано на примере промотора гена *MYC*.

Обнаруженные сайты окисления ДНК могут быть использованы для интерпретации регуляторной роли изменения конформации ДНК и выбора альтернативных модельных структур ДНК при разработке новых фотосенсибилизаторов для направленного действия на G4-структуры промоторов онкогенов.

Исследование выполнено за счет гранта Российского научного фонда (проект № 20-14-00332).

Статья не содержит экспериментов, включающих людей или животных в качестве объектов исследования.

Авторы заявляют об отсутствии конфликта интересов.

СПИСОК ЛИТЕРАТУРЫ

- Besnard E., Babled A., Lapasset L., Milhavet O., Parinello H., Dantec C., Marin J.M., Lemaitre J.M. (2012) Unraveling cell type-specific and reprogrammable human replication origin signatures associated with G-quadruplex consensus motifs. *Nat. Struct. Mol. Biol.* **19**, 837–844.
- Siddiqui-Jain A., Grand C.L., Bearss D.J., Hurley L.H. (2002) Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress *c-MYC* transcription. *Proc. Natl. Acad. Sci. USA*. **99**, 11593–11598.
- Spiegel J., Adhikari S., Balasubramanian S. (2020) The structure and function of DNA G-quadruplexes. *Trends Chem.* **2**, 123–136.
- Marsico G., Chambers V.S., Sahakyan A.B., McCauley P., Boutell J.M., Antonio M.D., Balasubramanian S. (2019) Whole genome experimental maps of DNA G-quadruplexes in multiple species. *Nucleic Acids Res.* **47**, 3862–3874.
- Chambers V.S., Marsico G., Boutell J.M., Di Antonio M., Smith G.P., Balasubramanian S. (2015) High-throughput sequencing of DNA G-quadruplex structures in the human genome. *Nat. Biotechnol.* **33**, 877–881.
- Miskiewicz J., Sarzynska J., Szachniuk M. (2021) How bioinformatics resources work with G4 RNAs. *Brief. Bioinformatics*. **22**, bbaa201.
- Grün J.T., Schwalbe H. (2022) Folding dynamics of polymorphic G-quadruplex structures. *Biopolymers*. **113**, e23477.
- Sun D., Hurley L.H. (2010) Biochemical techniques for the characterization of G-quadruplex structures: EMSA, DMS footprinting, and DNA polymerase stop assay. *Methods Mol. Biol.* **608**, 65–79.
- Guo J.U., Bartel D.P. (2016) RNA G-quadruplexes are globally unfolded in eukaryotic cells and depleted in bacteria. *Science*. **353**(6306), aaf5371.
- Beniaminov A.D., Novikov R.A., Mamaeva O.K., Mitkevich V.A., Smirnov I.P., Livshits M.A., Shchyolkina A.K., Kaluzhny D.N. (2016) Light-induced oxidation of the telomeric G4 DNA in complex with Zn(II) tetracarboxymethyl porphyrin. *Nucleic Acids Res.* **44**, 10031–10041.
- Chashchina G.V., Beniaminov A.D., Kaluzhny D.N. (2019) Stable G-quadruplex structures of oncogene promoters induce potassium-dependent stops of thermostable DNA polymerase. *Biochemistry (Moscow)*. **84**, 562–569.
- Kovaleva O.A., Tsvetkov V.B., Mamaeva O.K., Ol'shevskaya V.A., Makarenkov A.V., Dezhenskova L.G., Semeikin A.S., Borisova O.F., Shtil A.A., Shchyolkina A.K., Kaluzhny D.N. (2014) Preferential DNA photocleavage potency of Zn(II) over Ni(II) derivatives of carboxymethyl tetracationic porphyrin: the role of the mode of binding to DNA. *Eur. Biophys. J.* **43**, 545–554.
- Beniaminov A., Chashchina G., Shchyolkina A., Kaluzhny D. (2021) Oxidative probing of the G4 DNA structure induced in double-stranded DNA by molecular crowding or pyridostatin. *Biochimie*. **191**, 33–36.
- Beniaminov A., Shchyolkina A., Kaluzhny D. (2019) Conformational features of intramolecular G4-DNA constrained by single-nucleotide loops. *Biochimie*. **160**, 122–128.
- Chaires J.B., Trent J.O., Gray R.D., Dean W.L., Buscaglia R., Thomas S.D., Miller D.M. (2014) An improved model for the *hTERT* promoter quadruplex. *PLoS One*. **9**, e115580.
- Monsen R.C., DeLeeuw L., Dean W.L., Gray R.D., Sabo T.M., Chakravarthy S., Chaires J.B., Trent J.O. (2020) The *hTERT* core promoter forms three parallel G-quadruplexes. *Nucleic Acids Res.* **48**, 5720–5734.
- Lim K.W., Lacroix L., Yue D.J.E., Lim J.K.C., Lim J.M.W., Phan A.T. (2010) Coexistence of two distinct G-quadruplex conformations in the *hTERT* promoter. *J. Am. Chem. Soc.* **132**, 12331–12342.
- Ambrus A., Chen D., Dai J., Jones R.A., Yang D. (2005) Solution structure of the biologically relevant G-quadruplex element in the human *c-MYC* promoter. Implications for G-quadruplex stabilization. *Biochemistry*. **44**, 2048–2058.
- Phan A.T., Modi Y.S., Patel D.J. (2004) Propeller-type parallel-stranded G-quadruplexes in the human *c-myc* promoter. *J. Am. Chem. Soc.* **126**, 8710–8716.
- Phan A.T., Kuryavyi V., Gaw H.Y., Patel D.J. (2005) Small-molecule interaction with a five-guanine-tract G-quadruplex structure from the human *MYC* promoter. *Nat. Chem. Biol.* **1**, 167–173.
- Arutyunyan A.F., Tevonyan L.L., Beniaminov A.D., Yegorov Y.E., Kaluzhny D.N. (2021) The phototoxic effect of water-soluble porphyrins on human clear cell renal cell carcinoma line Caki-1. *Biophysics*. **66**, 273–277.
- Miyoshi D., Karimata H., Sugimoto N. (2006) Hydration regulates thermodynamics of G-quadruplex formation under molecular crowding conditions. *J. Am. Chem. Soc.* **128**, 7957–7963.
- Zheng K.W., Chen Z., Hao Y.H., Tan Z. (2010) Molecular crowding creates an essential environment for the formation of stable G-quadruplexes in long double-stranded DNA. *Nucleic Acids Res.* **38**, 327–338.
- Liu L.-Y., Ma T.-Z., Zeng Y.-L., Liu W., Mao Z.-W. (2022) Structural basis of pyridostatin and its derivatives specifically binding to G-quadruplexes. *J. Am. Chem. Soc.* **144**, 11878–11887.
- Severov V., Tsvetkov V., Barinov N., Babenko V., Klinov D., Pozmogova G. (2022) Spontaneous DNA synapsis by forming noncanonical intermolecular structures. *Polymers*. **14**, 2118.
- Hasanau T., Pisarev E., Kisil O., Nonoguchi N., Le Calvez-Kelm F., Zvereva M. (2022) Detection of *TERT* promoter mutations as a prognostic biomarker in gliomas: methodology, prospects, and advances. *Bio-medicines*. **10**, 728.

Oxidative Probing of the G4 DNA Structure by ZnP1 Porphyrin within Sequences of *MYC* and *TERT* Promoters

G. V. Chashchina¹ and D. N. Kaluzhny^{1, *}

¹*Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Moscow, 119991 Russia*

**e-mail: uzhny@mail.ru*

The formation of G4 structures in a DNA double helix competes with the complementary strand, which can change the equilibrium G4 structures studied on single-strand models by classical structural methods. A relevant task is to develop methods for detecting and localizing G4 in extended double-stranded (ds) DNA in the promoter regions of the genome. The porphyrin derivative ZnP1 selectively binds and leads to photo-induced oxidation of guanine in G4 structures on single-stranded (ss) and dsDNA model systems. In this research, we show the oxidative effect of ZnP1 on native sequences of *MYC* and *TERT* oncogene promoters that potentially capable to form G4 structures. Single strand breaks in the guanine rich sequence caused by ZnP1 oxidation and subsequent cleavage of the DNA strand by Fpg glycosylase were identified and assigned to the nucleotide sequence. The detected break sites corresponded to sequences potentially capable of forming G4 structures. New data were obtained on the possibility of folding G4 structures in the presence of a complementary strand in the context of the DNA double helix of the natural sequence.

Keywords: G-quadruplexes, porphyrin, photosensibilisation, DNA oxidation

ПОИСК ПРОГНОСТИЧЕСКИХ МАРКЕРОВ ГЕПАТОЦЕЛЛЮЛЯРНОЙ КАРЦИНОМЫ С ПОМОЩЬЮ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА ТЕКСТОВ И ДАННЫХ¹

© 2023 г. Z. H. Yang^a, S. X. Wang^b, *

^a*School of Criminal Law, East China University of Political Science and Law, Shanghai, 200042 China*

^b*Department of Transfusion, Minhang Hospital, Fudan University, Shanghai, 201199 China*

*e-mail: 2021110083@ecupl.edu.cn

Поступила в редакцию 20.04.2022 г.

После доработки 15.09.2022 г.

Принята к публикации 14.10.2022 г.

Транскатетерная артериальная химиоэмболизация — один из интервенционных методов лечения гепатоцеллюлярной карциномы (ГЦК). Этот метод обычно используют для пациентов с ГЦК от промежуточной до прогрессирующей стадий. Логично предположить, что идентификация генов, связанных с ГЦК, позволит повысить результативность транскатетерной артериальной химиоэмболизации. С целью исследовать роль связанных с ГЦК генов, и предоставить достоверные доказательства лечения с применением транскатетерной артериальной химиоэмболизации мы провели комплексный биоинформатический анализ. С помощью интеллектуального анализа текста (“гепатоцеллюлярная карцинома”) и анализа данных микрочипа (GSE104580) мы получили стандартный набор генов, который проанализировали по генной онтологии и Киотской энциклопедии генов и геномов (Kyoto Encyclopedia of Genes and Genomes). Значимые восемь генов, кластеризованные в сеть белок-белковых взаимодействий, выбрали для дальнейших исследований. При анализе выживаемости обнаружено, что низкая экспрессия ключевых генов тесно связана с показателем выживаемости пациентов, принимавших участие в исследовании. Для оценки корреляции между экспрессией ключевых генов и иммунной инфильтрацией опухоли использовали корреляционный анализ Пирсона. В результате идентифицировано 15 препаратов, нацеленных на семь из восьми исследованных генов, на основании чего их можно рассматривать как потенциальные лекарственные компоненты для транскатетерной артериальной химиоэмболизации при лечении ГЦК.

Ключевые слова: рак печени, гепатоцеллюлярная карцинома, транскатетерная артериальная химиоэмболизация, биоинформатический анализ, прогноз

DOI: 10.31857/S0026898423030187, EDN: CHYJBX

Exploring the Prognostic Features of Hepatocellular Carcinoma via Text Mining and Data Analysis

Z. H. Yang¹ and S. X. Wang², *

¹*School of Criminal Law, East China University of Political Science and Law, Shanghai, 200042 China*

²*Department of Transfusion, Minhang Hospital, Fudan University, Shanghai, 201199 China*

*e-mail: 2021110083@ecupl.edu.cn

Transcatheter arterial chemoembolization is one of the interventional treatments for hepatocellular carcinoma (HCC). This treatment is generally used for patients with intermediate to advanced hepatocellular carcinoma, and identifying the role of HCC-related genes can help improve the efficiency of transcatheter arterial chemoembolization. To investigate the role of HCC-related genes and to provide valid evidence for transcatheter arterial chemoembolization treatment, we performed a comprehensive bioinformatics analysis.

¹ Статья представлена авторами на английском языке.

Through text mining (“hepatocellular carcinoma”) and microarray data analysis (GSE104580), we obtained a standard gene set, which was followed by gene ontology and Kyoto Gene and Genome Encyclopedia analysis. The significant 8 genes clustered in protein-protein interactions network were chosen to be used in the follow-up analysis. Through survival analysis low expression of the key genes were found to be strongly associated with survival in HCC patients in this study. The correlation between the expression of the key genes and tumor immune infiltration was assessed by Pearson correlation analysis. As a result, 15 drugs targeting seven of the eight genes have been identified, and therefore can be considered as potential components for transcatheter arterial chemoembolization treatment of HCC.

Keywords: liver cancer, hepatocellular carcinoma, transcatheter arterial chemoembolization, bioinformatics analysis, prognosis

УДК 57.083.18;57.087.23;57.088.3

РЕКОМБИНАЗНАЯ ПОЛИМЕРАЗНАЯ АМПЛИФИКАЦИЯ ДЛЯ БЫСТРОГО ВЫЯВЛЕНИЯ ВОЗБУДИТЕЛЕЙ БАКТЕРИАЛЬНОЙ ПНЕВМОНИИ ЧЕЛОВЕКА

© 2023 г. С. А. Лапа^а, *, С. А. Суржииков^а, С. А. Благодатских^б, В. Е. Шершов^а, А. В. Чудинов^а

^аИнститут молекулярной биологии им. В.А. Энгельгардта Российской академии наук,
Москва, 119991 Россия

^бГосударственный научный центр прикладной микробиологии и биотехнологии,
Оболensk, Московская обл., 142279 Россия

*e-mail: lapa@biochip.ru

Поступила в редакцию 02.11.2022 г.

После доработки 19.11.2022 г.

Принята к публикации 20.11.2022 г.

Разработана диагностическая система, основанная на рекомбиназной полимеразной амплификации (RPA), для выявления шести бактериальных возбудителей пневмонии человека. Сконструированы и оптимизированы видоспецифичные праймеры для проведения мультиплексной реакции в одном общем объеме. Используются меченые праймеры для надежной дискриминации близких по размеру продуктов амплификации. Идентификацию возбудителя проводят путем визуального анализа электрофореграммы. Аналитическая чувствительность разработанной мультиплексной RPA составила 10^2 – 10^3 копий ДНК. Специфичность метода определяли по отсутствию перекрестной амплификации исследованных образцов ДНК возбудителей пневмонии для каждой пары праймеров, а также для ДНК *Mycobacterium tuberculosis* H37_{rv}. Для исследованных образцов этот показатель составил 100%. Время выполнения анализа — менее часа, включая электрофоретический контроль реакции. Тест-система может найти применение в профильных клинических лабораториях для экспресс-анализа образцов от пациентов с подозрением на пневмонию.

Ключевые слова: рекомбиназная полимеразная амплификация, RPA, бактериальная пневмония

DOI: 10.31857/S0026898423030072, **EDN:** CHGYQF

Инфекционная пневмония относится к социально значимым заболеваниям и представляет собой воспаление респираторных отделов легких. Внебольничная пневмония бактериальной природы встречается гораздо чаще грибковой, паразитарной, аллергической, ожоговой и лучевой, однако часто сочетается с вирусной инфекцией [1, 2]. Ежегодно бактериальной пневмонией болеет около 1000 человек на 100 тыс. населения (данные сильно варьируют для различных возрастных категорий [3, 4]). Бактериальная пневмония характеризуется этиологической неоднородностью и вызывается широким спектром грамположительных и грамотрицательных бактерий. В этой связи крайне важно своевременное выявление возбудителя, поскольку стратегия лечения для различных патогенных бактерий может отличаться [5].

Современная клиническая диагностика базируется на классических методах: зачастую просто на определении клинической картины заболевания [6], на данных радиологического исследова-

ния [7, 8] или культуральных методов [9], — что с учетом быстрой динамики развития патологии далеко не оптимально. В лучшем случае используют серологические тесты или ПЦР, которые в основном ориентированы на определение лишь одного вида возбудителя [10, 11].

Ранее нами разработан метод мультиплексной ПЦР в едином общем объеме для выявления наиболее значимых возбудителей пневмонии человека [12, 13]. В представленной статье описана разработка диагностической экспресс-системы, основанной на применении рекомбиназной полимеразной амплификации (Recombinase Polymerase Amplification, RPA) в мультиплексном формате. Для этой реакции, проходящей в изотермическом режиме, характерна более высокая, по сравнению с ПЦР, скорость накопления продукта и повышенная чувствительность, а также для ее проведения не нужен термоциклер [14, 15]. Преимущества RPA в условиях все большего распространения и доступности реагентов, необходимых для выполнения анализа в профильных клинических

лабораториях, позволяют рассматривать этот метод как перспективную альтернативу широко используемым способам выявления возбудителей пневмонии.

ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

Штаммы. В работе использовали ДНК штаммов из коллекции Государственного научного центра прикладной микробиологии и биотехнологии (“ГНЦ ПМБ”, п. Оболенск, Россия), в том числе клинические и природные изоляты. Перед выделением ДНК изоляты из природных источников, а также клинические образцы культивировали на селективных средах. ДНК из культур выделяли и деконтаминировали с использованием СТАВ-метода, как описано ранее [16, 17]. Работы с изолятами и живыми коллекционными культурами проводили на базе указанного учреждения.

Для контроля стерильности образцов ДНК, выделенных из микроорганизмов, делали высеив из полученных растворов ДНК на плотные питательные среды. В случае штаммов *Pseudomonas aeruginosa* и *Klebsiella pneumoniae* использовали среду № 1 ГРМ (“ГНЦ ПМБ”, Россия), культивирование проводили при 37°C в течение трех суток. В случае штаммов *Haemophilus influenzae* использовали среду ГБМ-АГАР с ростовой добавкой (“ГНЦ ПМБ”), в случае штаммов *Legionella pneumophila* — среду ЛЕГИОНЕЛБАКАГАР (“ГНЦ ПМБ”), в случае штаммов *Streptococcus pneumoniae* и *Staphylococcus aureus* — Blood Agar Base (“ГНЦ ПМБ”) с добавлением 5% крови барана. Инкубацию проводили при 37°C в CO₂-инкубаторе в присутствии 5% CO₂ в течение 3 суток. Через трое суток на питательных средах ни в одном случае роста не наблюдали.

Качество ДНК оценивали по электрофоретической подвижности в 1%-ном агарозном геле, концентрацию измеряли спектрофотометрически с использованием GeneQuant Pro RNA/DNA Calculator (“Amersham Pharmacia Biotech”, Великобритания).

Для контроля специфичности праймеров использовали деконтаминированную очищенную геномную ДНК *Mycobacterium tuberculosis* H37_{rv}.

Праймеры. Нуклеотидные последовательности геномных мишеней выравнивали с помощью алгоритма ClustalW (www.clustal.org). Праймеры для RPA (длиной от 35 до 40 н.) конструировали с использованием сетевого ресурса www.idtdna.com. Проводили тестирование на внутри- и межмолекулярные вторичные структуры. Специфичность анализировали с помощью алгоритма BLAST (NIN, США, <https://blast.ncbi.nlm.nih.gov/Blast.cgi>).

Праймеры конструировали для образования продуктов амплификации различной длины с учетом удобства их визуального контроля по электрофореграмме. Кроме того, с целью повышения надежности дискриминации на 5'-конец праймеров, специфичных к *H. influenzae*, *L. pneumophila* и *S. aureus*, вводили Cy5-метку [17].

Твердофазный синтез олигонуклеотидов проводили на автоматическом синтезаторе ABI 394 DNA/RNA (“Applied Biosystems”, США) по стандартному регламенту. Для очистки целевых продуктов использовали колонку BDS Hypersil C18 (“Thermo Fisher Scientific”, США).

Мультиплексная RPA. Реакционная смесь (50 мкл) содержала компоненты набора TwistAmp (“Twist-DX”, Великобритания) в рекомендуемых производителем концентрациях; раствор 6 пар праймеров, специфичных к бактериальным возбудителям пневмонии (из расчета 10 пмоль на реакцию каждого), а также 1 мкл очищенных полногеномных бактериальных ДНК в концентрации от 10¹ до 10⁶ копий/мкл. Раствор, содержащий 6 пар праймеров, перед внесением в общий реакционный объем прогревали в течение 5 мин при 95°C и сразу же помещали на лед. Реагенты набора, за исключением матрицы (образец ДНК) и ацетата магния, входили в общую смесь, которую после аккуратного перемешивания разливали по реакционным пробиркам объемом 200 мкл. На внутреннюю сторону каждой реакционной пробирки вносили ацетат магния и пробирку закрывали лишь перед окончательным перемешиванием и началом проведения реакции, образец ДНК вводили на верхнюю часть внутренней поверхности пробирки для предотвращения преждевременного контакта с компонентами смеси. Растворы “сбрасывали” коротким центрифугированием, аккуратно перемешивали шестикратным переворачиванием и еще раз “сбрасывали”, после чего немедленно помещали в термостат. Реакцию проводили в течение 20 мин при 40°C на лабораторном термостате для микропробирок TDB-120 (“Biosan”, Латвия). Реакцию останавливали либо внесением в реакционную смесь концентрированного буфера для электрофоретических образцов, либо прогреванием в течение 5 мин при 95°C.

Чувствительность RPA определяли с использованием панели разведений ДНК в интервале 10¹–10³ копий в реакционной смеси (50 мкл). Скорость накопления продукта оценивали при температурном режиме, описанном выше, с добавлением красителя EvaGreen (“Lumiprobe”, Россия) на амплификаторе iCycler iQ5 (“Bio-Rad”, США). Эксперименты выполняли в трех повторениях для каждого исследуемого образца и рассчитывали среднее значение. Качественный анализ

Таблица 1. RPA-праймеры

Бактерия	Праймер			Длина продукта ^b , п.н.
	название	нуклеотидная последовательность ^a , 5'→3'	длина, н.	
<i>Pseudomonas aeruginosa</i>	oprL-f1	GTACAACATGGCTCTGGGCGAGCGTCGTGCCAAGG	35	150
	oprL-r1	TTCTGAGCCCAGGACTGCTCGTCGTGGCCGGTAGC	35	
<i>Haemophilus influenzae</i>	fucK-f1	TGCTAACGGTAATTGGGATCCATCGATTTTAGCCT	35	195
	fucK-r1	CC(A/T)GCACCAGACCCAAACACAGCAAATTGGGTATC	35	
<i>Streptococcus pneumoniae</i>	cpsB-f1	ATGTAGATGACGGTCCCAAGTCAAGAGAGGAAAGC	35	210
	cpsB-r1	ATCTCTGCGCCATAAGCAATGACTAAATCATCTGC	35	
<i>Legionella pneumophila</i>	sidA-f1	GCAGCGATCA(C/T)CCTCGGTGGATTGGCTGCCGCAAT	35	232
	sidA-r1	CATGTTGGAGTTCTATGGCACGAATTC(G/T)AATAAG	35	
<i>Klebsiella pneumoniae</i>	fimH-f1	GGTGTGCTGTCGAGTTTTTCAGGCACCGTGAAATATAACG	40	261
	fimH-r1	GTTGTTGGCGTAGATGTTCCAGATGAAGTGAAGGAGTCG	40	
<i>Staphylococcus aureus</i>	ebpS-f1	CCAAATATCGCTAATGCACCGATAATTAGTACAGC	35	280
	ebpS-r1	ACTCGACTGAGGATAAAGCGTCTCAAGATAAGTCT	35	

^aВ скобках указаны вырожденные позиции, связанные с вариативностью данного участка генетической мишени у различных штаммов.

^bПолужирным шрифтом выделены значения длин продуктов, несущих флуоресцентную метку Cy5, — с целью лучшей дифференциации от продуктов близкой длины для других возбудителей (на фореграмме показано чередование продуктов “меченый—немеченый” — через один по увеличению длины).

наличия/отсутствия продукта амплификации проводили визуально по электрофореграмме.

Горизонтальный электрофорез для контроля продуктов RPA. Продукты RPA разделяли в 4%-ном агарозном геле Agarose LE (“Helicon”, Россия) в течение 20 мин при 10 В/см, для окрашивания использовали SYBR Green I (“Molecular Probes”, США). Визуализацию проводили с помощью трансиллюминатора TCP-20.MC 254/312 нм (“Vilber Lourmat”, Франция) при длине волны 260 нм, снабженного цифровым фотоаппаратом EOS 60D (“Canon”, Япония), — для получения цветного изображения; на системе гель-документирования ChemiScope 6200 Touch (“Clinx Science Instruments”, Китай) с помощью входящего в комплект УФ-трансиллюминатора — для общей детекции ДНК по окрашиванию SYBR Green I; а также с использованием встроенных LED-светодиодов и светофильтров “Red light excitation/emission”, соответствующих характеристикам флуоресцентного красителя Cy5, — для избирательной детекции продуктов удлинения меченых праймеров.

РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ

Для создания системы мультиплексной RPA сконструировали набор из 6 пар видоспецифичных праймеров для социально значимых бактериальных возбудителей пневмонии: *Staphylococcus aureus*, *S. pneumoniae*, *L. pneumophila*, *K. pneumoniae*, *H. influenzae* и *P. aeruginosa*. Для каждого бактериального возбудителя выбирали функционально значимые видоспецифичные участки генов. Ген

cpsB S. pneumoniae кодирует тирозинспецифичную протеинфосфатазу B; ген *ebpS Staphylococcus aureus* — эластинсвязывающий белок S; ген *fucK*, применяемый для видовой идентификации *H. influenzae*, относится к генам углеводного обмена; ген *oprL* кодирует пептидогликанассоциированный белок *P. aeruginosa*; для видовой идентификации *L. pneumophila* и *K. pneumoniae* использовали детерминанты вирулентности этих бактерий: *sidA* и *fimH* соответственно. Алгоритмы выбора генов и исследования по функциональной роли кодируемых ими белков подробно изложены в работе, посвященной разработке мультиплексной ОТ-ПЦР с иммобилизованными праймерами [17].

Основные подходы к конструированию праймеров, специфичных к участкам геномов бактериальных возбудителей пневмонии, описаны ранее [18] с учетом того факта, что для RPA в большинстве случаев требуются праймеры большей длины (30–35 н.), чем для ПЦР (Инструкция TwistAmp, официальные рекомендации по дизайну праймеров: <https://www.twistdx.co.uk/wp-content/uploads/2021/04/twistamp-assay-design-manual-v2-5.pdf>). Однако в процессе оптимизации этой системы длину праймеров корректировали в сторону увеличения по сравнению с рекомендованной, а именно 35–40 н., что резко повысило воспроизводимость результатов в мультиплексном режиме работы системы.

Последовательности праймеров указаны в табл. 1 в порядке увеличения длин образующихся продуктов.

Таблица 2. Штаммы возбудителей бактериальной пневмонии человека, использованные в работе

Возбудитель	№	Штамм	Источник
<i>Pseudomonas aeruginosa</i>	B-4807	10662 NCTC/ATCC 25668	Клинический штамм
	B-7565	ATCC 17650	
	B-8361	DE 177	
<i>Haemophilus influenzae</i>	B-7400	ATCC 49247	Клинический штамм
	B-7399	ATCC 9006 тип А	
	B-9029	Хак-1	
<i>Streptococcus pneumoniae</i>	B-7756	ATCC 49619	
	B-7397	ATCC 6303	
	B-7398	ATCC 6305	
<i>Legionella pneumophila</i>	B-5978	ATCC 33152	Природный штамм Природный штамм
	B-7783	7	
	B-8538	0042	
<i>Klebsiella pneumoniae</i>	B-4811	9633 NCTC/ATCC 13883	Клинический штамм Клинический штамм
	B-7877	KPB-594	
	B-7878	KPB-1802	
<i>Staphylococcus aureus</i>	B-5931	ATCC 25923	Клинический штамм
	B-7471	ATCC 43300	
	B-8377	Оренбург_2014 №67	

В исследовании использовали образцы деконтаминированной геномной ДНК, указанные в табл. 2.

При разработке системы использовали как коллекционные штаммы, так и клинические и природные изоляты. Основные параметры системы (аналитическая чувствительность, специфичность, воспроизводимость) оценивали с помощью амплификации в реальном времени (в изотермическом режиме) для разных концентраций образцов, с последующим электрофоретическим контролем. Для представленной в табл. 2 выборки специфичность составила 100%; причем перекрестной амплификации не наблюдали. Кроме того, геномную ДНК штамма *M. tuberculosis* H37_{rv} использовали в качестве контроля на специфичность, так как клинические проявления туберкулеза легких сходны с таковыми при пневмонии. Чувствительность метода составила от 10^2 до 10^3 копий ДНК в реакционной смеси, что несколько превышает чувствительность методов, основанных на мультиплексной ПЦР (10^3 – 10^4); при этом время выполнения RPA-анализа составляет менее часа, что гораздо меньше, чем для ПЦР (1.5–3 ч).

На рис. 1 представлен результат выполнения мультиплексной RPA в одном общем объеме с результирующим набором праймеров и при оптимизированных условиях амплификации.

Из данных, представленных на рис. 1, видно, что визуальная дискриминация вида возбудителя в ряде случаев затруднительна, так как выбранный интервал длин продуктов амплификации при конструировании системы RPA был ограничен официальными рекомендациями производителя – “TwistDX”. Кроме того, теоретически при наличии в образце двух и более генетических мишеней возможна конкурентная амплификация в пользу продуктов меньшей длины.

Учитывая эти факторы, для надежной дискриминации видовой принадлежности патогенов мы предложили двухволновую детекцию флуоресценции. Комплекс двухцепочечной ДНК с SYBR Green I имеет максимум поглощения при длине волны 497 нм ($\lambda_{\text{ex}} = 497$ нм) и максимумом излучения при 520 нм ($\lambda_{\text{em}} = 520$ нм). Окрашивание SYBR Green I применяли по завершении электрофореза для визуализации всех продуктов амплификации. Цианиновый краситель Cy5, ковалентно связанный с 5'-концом трех прямых праймеров (с возрастанием длин продуктов по схеме “через один”), применяли для избирательной визуализации продуктов меченых праймеров. Для этого красителя максимум поглощения соответствует длине волны 650 нм ($\lambda_{\text{ex}} = 650$ нм), а максимум излучения – 670 нм ($\lambda_{\text{em}} = 670$ нм). Такая двухволновая схема детекции не приводит к перекрестному

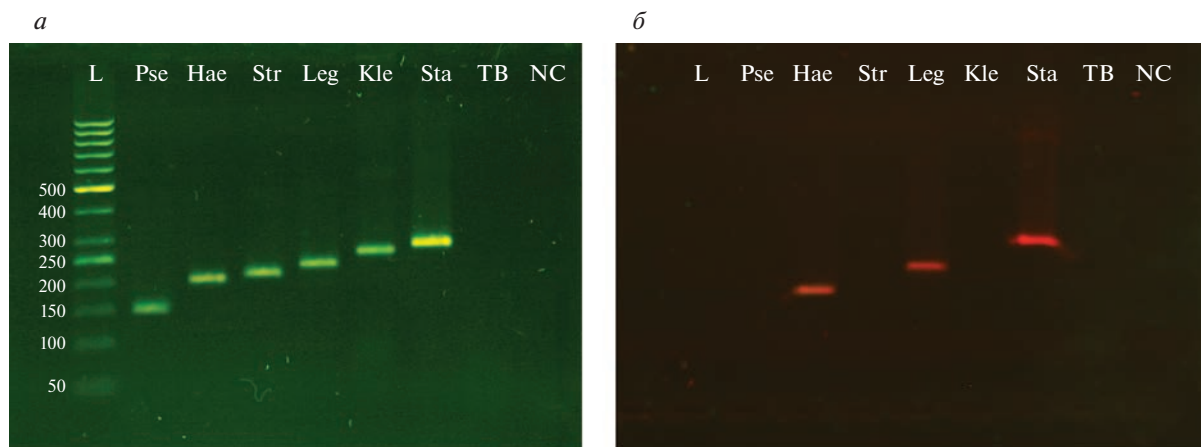


Рис. 1. Визуальная идентификация возбудителя пневмонии на электрофореграмме. *а* — Неизбирательная детекция ДНК с помощью красителя SYBR Green I; *б* — избирательная идентификация продуктов удлинения праймеров, содержащих цианиновый краситель Cy5. Окрашивание праймеров, специфичных к *H. influenzae*, *L. pneumophila* и *Staphylococcus aureus*, позволяет надежно дискриминировать соответствующие продукты амплификации от не несущих флуоресцентную метку: *P. aeruginosa*, *S. pneumoniae*, *K. pneumoniae*. Обозначения: L — маркер длин ДНК Ladder GeneRuler 50 bp (“Thermo Fisher Scientific”), Pse — *P. aeruginosa*, Hae — *H. influenzae*, Str — *S. pneumoniae*, Leg — *L. pneumophila*, Kle — *K. pneumoniae*, Sta — *Staphylococcus aureus*, TB — *M. tuberculosis* (контроль специфичности праймеров), NC — отрицательный контроль. Цифрами слева обозначена длина двухцепочечных ДНК (от 50 до 500 п.н.).

возбуждению флуорофоров и удобна для раздельной визуализации. Продукты амплификации, полученные с Cy5-мечеными праймерами, отличаются по цвету даже при использовании УФ-трансиллюминатора с применением оптического фильтра SYBR photographic filter S7569 (“Molecular Probes”, США), но для надежности интерпретации результатов мы использовали систему геледокументирования со специализированными источниками возбуждения и светофильтрами для каждого из красителей.

ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Рекомбиназная полимеразная амплификация обладает рядом преимуществ по сравнению с ПЦР: отсутствие необходимости в применении амплификатора (изотермические условия реакции), высокая скорость накопления продуктов и, в большинстве случаев, более высокая, по сравнению с ПЦР, чувствительность метода [15].

Потенциально технология RPA способна занять нишу ПЦР и других молекулярно-биологических методов, применяемых в профильных клинических лабораториях. Этот метод особенно востребован для экспресс-анализа инфекционных заболеваний с быстрой динамикой развития, когда требуется как можно быстрее получить результат и сформировать стратегию лечения пациента.

Разработка предлагаемой системы мультиплексного анализа, основанной на RPA, позволила сократить время анализа в два–три раза по сравнению с классической ПЦР. Как и в случае

мультиплексной ПЦР, предлагаемая система может быть расширена для анализа более широкого спектра патогенов. Как и в случае большинства диагностических систем, при использовании RPA возникают проблемы, которые обусловлены различиями в природе тестируемых образцов, присутствием ингибиторов, загрязнений и микроорганизмов-спутников. Присутствие в исследуемом образце нецелевых нуклеиновых кислот затрудняет анализ и интерпретацию результатов. Для преодоления этой проблемы необходим более внимательный подход к дизайну видоспецифичных праймеров.

Для разработанной системы мультиплексного анализа, основанной на методе RPA, показана высокая специфичность и чувствительность, для проведения анализа не требуется специализированное оборудование (используется изотермическая амплификация). Однако для корректной интерпретации результатов необходимо использовать двухволновую систему геле-документирования. Упрощению аппаратной части системы анализа может способствовать увеличение флуоресценции с избирательно окрашенных праймеров. Тогда сигнал красителя будет хорошо регистрироваться визуально. Это позволило бы использовать стандартный трансиллюминатор вместо двухволновой системы анализа. Достичь этого путем встраивания меченых дезоксиуридинтрифосфатов при настоящем дизайне системы невозможно, так как это приведет к неизбирательному окрашиванию продуктов амплификации. Возможна ковалентная пришивка нескольких молекул флуорофора в

один праймер, но при таком подходе возрастает размер и стерический эффект таких структур. Это может негативно влиять на ферментативную доступность комплексов праймер–целевая ДНК. Кроме того, возможны эффекты электронного переноса (EET, RET) и тушение флуорофора, что необходимо проверить экспериментально. Следует отметить, что дальнейшая оптимизация системы возможна при подборе подходящего индивидуального красителя для введения метки в праймеры. Это позволит проводить интерпретацию результатов с применением стандартного трансиллюминатора и упростить систему детекции сигнала флуоресцентных меток.

В результате проведенного исследования сконструированы высокоспецифичные RPA-праймеры для одновременной идентификации шести видов бактерий — возбудителей пневмонии человека. Разработан метод ускоренного дифференциального анализа бактериальных патогенов (время анализа, включая электрофоретическое разделение продуктов, — менее 1 ч). Мы стремились к упрощению как дизайна, так и эксплуатации разработанного подхода. Система ориентирована на применение в клинических лабораториях для ускоренного массового тестирования образцов с целью своевременного выбора или коррекции стратегии лечения пациентов.

Исследование поддержано грантом Российского научного фонда (№ 22-14-00257).

Настоящая статья не содержит описания каких-либо исследований с участием людей или животных в качестве объектов.

Авторы заявляют об отсутствии конфликта интересов.

СПИСОК ЛИТЕРАТУРЫ

1. Campigotto A., Mubareka S. (2015) Influenza-associated bacterial pneumonia; managing and controlling infection on two fronts. *Expert. Rev. Anti. Infect. Ther.* **13**, 55–68.
2. Noskin G.A., Glassroth J. (1996) Bacterial pneumonia associated with HIV-1 infection. *Clin. Chest. Med.* **17**, 713–723.
3. Henig O., Kaye K.S. (2017) Bacterial pneumonia in older adults. *Infect. Dis. Clin. North. Am.* **31**, 689–713.
4. Eshwara V.K., Mukhopadhyay C., Rello J. (2020) Community-acquired bacterial pneumonia in adults: an update. *Indian. J. Med. Res.* **151**, 287–302.
5. Harris M., Clark J., Coote N., Fletcher P., Harnden A., McKean M., Thomson A. (2011) British thoracic society standards of care committee. British thoracic society guidelines for the management of community acquired pneumonia in children: update 2011. *Thorax*. **66** (Suppl. 2), ii1–ii23.
6. Cunha B.A. (2006) The atypical pneumonias: clinical diagnosis and importance. *Clin. Microbiol. Infect.* **12**(Suppl. 3), 12–24.
7. Azoulay E., Russell L., Van de Louw A., Metaxa V., Bauer P., Pova P., Montero J.G., Loeches I.M., Meh-ta S., Puxty K., Schellongowski P., Rello J., Mokart D., Lemiale V., Mirouse A. (2020) Diagnosis of severe respiratory infections in immunocompromised patients. *Intensive Care Med.* **46**, 298–314.
8. Mabie M., Wunderink R.G. (2003) Use and limitations of clinical and radiologic diagnosis of pneumonia. *Semin. Respir. Infect.* **18**, 72–79.
9. Budayanti N.S., Suryawan K., Iswari I.S., Sukrama D.M. (2019) The quality of sputum specimens as a predictor of isolated bacteria from patients with lower respiratory tract infections at a tertiary referral hospital, Denpasar, Bali-Indonesia. *Front. Med. (Lausanne)*. **6**, 64.
10. Lee N., Rainer T.H., Ip M., Zee B., Ng M.H., Antonio G.E., Chan E., Lui G., Cockram C.S., Sung J.J., Hui D.S. (2006) Role of laboratory variables in differentiating SARS-coronavirus from other causes of community-acquired pneumonia within the first 72 h of hospitalization. *Eur. J. Clin. Microbiol. Infect. Dis.* **25**, 765–772.
11. Dorigo-Zetsma J.W., Zaat S.A., Wertheim-van Dillen P.M., Spanjaard L., Rijntjes J., van Waveren G., Jensen J.S., Angulo A.F., Dankert J. (1999) Comparison of PCR, culture, and serological tests for diagnosis of *Mycoplasma pneumoniae* respiratory tract infection in children. *J. Clin. Microbiol.* **37**, 14–17.
12. Лапа С.А., Ключихина Е.С., Мифтахов Р.А., Золотов А.М., Заседателев А.С., Чудинов А.В. (2020) Мультиплексная ПЦР для выявления бактериальных возбудителей инфекционной пневмонии. *Биоорг. химия*. **46**, 550–552.
13. Лапа С.А., Ключихина Е.С., Мифтахов Р.А., Заседателев А.С., Чудинов А.В. (2021) Development of multi-primer PCR system with an open architecture for rapid detection of infectious pneumonia causative agents. *AIP Conf. Proc.* **2388**, 030018.
14. Piepenburg O., Williams C.H., Stemple D.L., Armes N.A. (2006) DNA detection using recombination proteins. *PLoS Biol.* **4**, e204.
15. Lobato I.M., O'Sullivan C.K. (2018) Recombinase polymerase amplification: basics, applications and recent advances. *Trends Analyt. Chem.* **98**, 19–35.
16. Wilson K. (2001) Preparation of genomic DNA from bacteria. *Curr. Protoc. Mol. Biol.* **56**, 2.4.1–2.4.5.
17. Лапа С.А., Мифтахов Р.А., Ключихина Е.С., Ам-мур Ю.И., Благодатских С.А., Шершов В.Е., Заседателев А.С., Чудинов А.В. (2021) Разработка мультиплексной ОТ-ПЦР с иммобилизованными праймерами для идентификации возбудителей инфекционной пневмонии человека. *Молекуляр. биология*. **55**, 944–955.
18. Ключихина Е.С., Шершов В.Е., Кузнецова В.Е., Лапа С.А., Чудинов А.В. (2021) Особенности оптимизации мультипраймерной ПЦР для выявления возбудителей инфекционной пневмонии человека. *Тонкие хим. технологии*. **16**, 225–231.

Recombinase Polymerase Amplification for Rapid Detection of Human Bacterial Pneumonia Pathogens

S. A. Lapa^{1, *}, S. A. Surzhikov¹, S. A. Blagodatskikh², V. E. Shershov¹, and A. V. Chudinov¹

¹*Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Moscow, 119991 Russia*

²*Scientific Center of Applied Microbiology and Biotechnology, Obolensk, Moscow Region, 142279 Russia*

**e-mail: lapa@biochip.ru*

A diagnostic system based on recombinase polymerase amplification (RPA) has been developed to identify six bacterial pathogens of human pneumonia. Species-specific primers have been designed and optimized to conduct a multiplex reaction in one common volume. Labeled primers were used for reliable discrimination of amplification products close in size. Identification of the pathogen was carried out by visual analysis of the electrophoregram. The analytical sensitivity of the developed multiplex RPA was 10^2 – 10^3 copies of DNA. The specificity of the system was determined by the absence of cross-amplification of the studied DNA samples of pneumonia pathogens for each pair of primers, as well as for the DNA of *Mycobacterium tuberculosis* H37_{rv}, and amounted to 100%. The analysis execution time is less than an 1 h, including electrophoretic reaction control. The test system can be used in specialized clinical laboratories for rapid analysis of samples from patients with suspected pneumonia.

Keywords: recombinase polymerase amplification, RPA, bacterial pneumonia

УДК 577.29

НУКЛЕАЗЫ CRISPR-Cas II ТИПА: АЛГОРИТМ ПОИСКА И *in vitro* ХАРАКТЕРИСТИКА¹

© 2023 г. А. А. Васильева^{a, b, *}, С. А. Алюкас^c, П. А. Селькова^{a, b}, А. Н. Арсениев^{a, b}, В. Е. Чернова^b,
О. С. Мушарова^a, Е. И. Климук^a, М. А. Ходорковский^b, К. В. Северинов^{a, d, **}

^aИнститут молекулярной генетики Российской академии наук, Москва, 123182 Россия

^bСанкт-Петербургский политехнический университет Петра Великого, Санкт-Петербург, 195251 Россия

^cООО “Биотехнологический кампус”, Москва, 117437 Россия

^dWaksman Institute of Microbiology, Piscataway, NJ, 08854 USA

*e-mail: daucussativus7@gmail.com

**e-mail: severik@waksman.rutgers.edu

Поступила в редакцию 25.11.2022 г.

После доработки 18.12.2022 г.

Принята к публикации 23.12.2022 г.

Использование компонентов систем адаптивного иммунитета бактерий CRISPR-Cas для внесения направленных изменений в ДНК открыло широкие перспективы для программируемого редактирования геномов высших организмов. Наиболее широко используемые генные редакторы основаны на Cas9-эффекторах CRISPR-Cas систем II типа. Cas9 способен специфически вносить двуниевые разрывы в участки ДНК, комплементарные последовательностям направляющих РНК. Несмотря на широкий спектр охарактеризованных Cas9, поиск новых вариантов Cas9 остается актуальной задачей, так как доступные Cas9-редакторы имеют ряд ограничений. В данной работе представлен алгоритм поиска и последующей характеристики новых Cas9-нуклеаз, разработанный в нашей лаборатории. Приведены подробные протоколы, описывающие биоинформатический поиск, клонирование и выделение рекомбинантных белков Cas9, проверку наличия нуклеазной активности *in vitro* и определение PAM-последовательности, необходимой для узнавания ДНК-мишеней. Рассмотрены потенциальные сложности, которые могут возникнуть в процессе работы, а также способы их преодоления.

Ключевые слова: CRISPR-система, Cas9, нуклеаза, редактирование генома, биоинформатический поиск

DOI: 10.31857/S0026898423030163, **EDN:** CHWIKZ

ВВЕДЕНИЕ

CRISPR-Cas системы обеспечивают адаптивный иммунитет бактерий и архей. Полная CRISPR-Cas система состоит из CRISPR-кассеты и кластера генов, кодирующих Cas-белки [1]. CRISPR-кассета — это участок ДНК, в котором короткие повторы разделяются уникальными последовательностями — спейсерами. Несмотря на название CRISPR (Clustered Regularly Interspaced Short

Palindromic Repeats — кластеризованные регулярно разделенные короткие палиндромные повторы) в ряде случаев, в частности в системах II типа, рассматриваемых в данной работе, повторы не относятся к палиндромам. Некоторые CRISPR-спейсеры возникают за счет встраивания чужеродного генетического материала (фрагментов вирусных или плазмидных последовательностей) в CRISPR-кассету. В связи с этим многообразие спейсеров, накопленных в CRISPR-кассете конкретной бактерии, можно считать “памятью” об инфекциях, перенесенных ее предками. В результате транскрипции CRISPR-кассеты и процессинга образующейся некодирующей РНК формируются короткие CRISPR-РНК (CRISPR-RNA, crРНК), каждая из которых содержит последовательность одного из спейсеров и фланкирующую последовательность фрагмента повтора. Индивидуальные crРНК связываются с Cas-нуклеазами,

¹ Дополнительная информация для этой статьи доступна по doi 10.31857/S0026898423030163 для авторизованных пользователей.

Сокращения: CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) — кластеризованные регулярно разделенные короткие палиндромные повторы; crРНК (CRISPR-RNA) — CRISPR-РНК; DR (Direct Repeat) — прямой повтор; ORF (Open Reading Frame) — открытая рамка считывания; PAM (Protospacer Adjacent Motif) — примыкающий к протоспейсеру мотив; tracrРНК (trans-activating CRISPR-RNA) — транс-активирующая crРНК.

образуя эффекторные комплексы. При появлении в клетке ДНК (а для некоторых систем РНК), содержащей участки, соответствующие одному из спейсеров CRISPR-кассеты (например, при повторной инфекции вирусом), эффекторный комплекс узнает ее за счет комплементарных взаимодействий с *cr*РНК. Активация нуклеазной активности белкового компонента эффекторного комплекса приводит к расщеплению чужеродной нуклеиновой кислоты в месте узнавания и последующей ее деградации [2–4].

Описанный общий механизм защитного действия систем CRISPR-Cas позволяет использовать эффекторные комплексы в качестве инструмента по направленному изменению последовательностей ДНК в клетках самых разных организмов, включая человека, что актуально для решения многих задач, стоящих перед биотехнологией и биомедициной. В этом случае используют эффекторный белок, запрограммированный не природной *cr*РНК, а специально подобранной направляющей РНК (RNA guide, РНК-гид). Участок направляющей РНК, соответствующий спейсеру *cr*РНК, комплементарен последовательности генома, в которую необходимо внести изменение.

Cas-белки можно разделить на две группы, исходя из их функций. Белки адаптационного модуля обеспечивают процесс адаптации – встраивания новых спейсеров в CRISPR-кассету. Белки эффекторного модуля отвечают за интерференцию – расщепление чужеродных нуклеиновых кислот. Исходя из строения эффекторов, CRISPR-Cas системы подразделяют на два класса [5]. В системах класса 1 эффекторы представляют собой комплексы, состоящие из нескольких Cas-белков. В системах класса 2 эффекторы – это односубъединичные белки. По наличию генов, кодирующих характерные дополнительные белки, системы класса 1 подразделяются на типы I, III и IV, а системы класса 2 – на типы II, V и VI. В свою очередь, каждый тип включает разнообразные подтипы [5].

Наличие лишь одного белка, достаточного для программируемого расщепления нуклеиновой кислоты, делает системы класса 2 наиболее удобными для практического применения. Эффекторы данного класса носят названия Cas9 (тип II), Cas12 (тип V) и Cas13 (тип VI). К наиболее изученным эффекторам класса 2 относятся несколько представителей типа II. Так, нуклеазу Cas9 из бактерии *Streptococcus pyogenes* (SpCas9) активно используют для редактирования геномов эукариот [6–8]. Однако геномные редакторы, основанные на охарактеризованных на сегодня Cas9-белках, не лишены недостатков. Во-первых, некоторые из этих нуклеаз имеют слишком большой размер, что затрудняет их доставку в клетки. Во-вторых, в

ряде случаев эффективность и специфичность редактирования оказывается невысокой. Наконец, мишени, доступные для конкретного Cas9-эффектора, требуют, кроме комплементарности к направляющей РНК, наличия специфической последовательности РАМ (см. “Организация CRISPR-Cas систем II типа”), что ограничивает число мишеней, которые могут быть подвергнуты редактированию. Cas9-нуклеазы с новыми свойствами позволяют расширить спектр существующих инструментов геномного редактирования и подобрать оптимальные эффекторы для решения конкретных задач. Таким образом, поиск новых Cas9-нуклеаз и их характеристика – задача, актуальная как для фундаментальной, так и прикладной науки.

ОРГАНИЗАЦИЯ CRISPR-Cas СИСТЕМ II ТИПА

В системах CRISPR-Cas типа II роль эффектора выполняет нуклеаза Cas9. Состав адаптационного модуля зависит от конкретного подтипа (II-A, II-B или II-C) и может включать в себя белки Cas1, Cas2, Csn2, Cas4 (рис. 1). Подавляющее большинство CRISPR-кассет систем II типа имеет длину прямого повтора (Direct Repeat, DR) от 34 до 37 п.н. и длину спейсера – от 29 до 32 п.н. Число спейсеров обычно не превышает 30 (табл. S1, см. Дополнительные материалы).

Зрелые индивидуальные *cr*РНК содержат частичную последовательность спейсера (например, в случае хорошо изученной системы II-A бактерии *S. pyogenes* длина спейсеров в кассете составляет 30 п.н., а участок, соответствующий спейсеру в *cr*РНК, составляет лишь 20 нуклеотидов). Кроме того, *cr*РНК содержит фрагмент DR, примыкающий к 3'-концевой части спейсерного участка (рис. 2). Для процессинга первичного транскрипта CRISPR-кассеты и функционирования эффектора необходима дополнительная *транс*-активирующая *cr*РНК (*trans*-activating CRISPR-RNA, *tracr*РНК). Последовательность *tracr*РНК имеет участок, комплементарный фрагменту последовательности DR (антиповтор) (рис. 2). Молекулы *tracr*РНК образуют дуплексы с комплементарными участками DR в первичном транскрипте CRISPR-кассеты. Расщепление дуплексов клеточными РНКазам приводит к образованию индивидуальных *cr*РНК; *tracr*РНК остается связанной с молекулами *cr*РНК и становится частью эффекторного комплекса [9]. Как правило, ген *tracr*РНК находится поблизости от остальных компонентов локуса CRISPR-Cas (рис. 1).

Для узнавания и расщепления комплементарных *cr*РНК мишеней эффекторным комплексом Cas9 необходим мотив, примыкающий к прото-спейсеру (Protospacer Adjacent Motif, РАМ) [10].

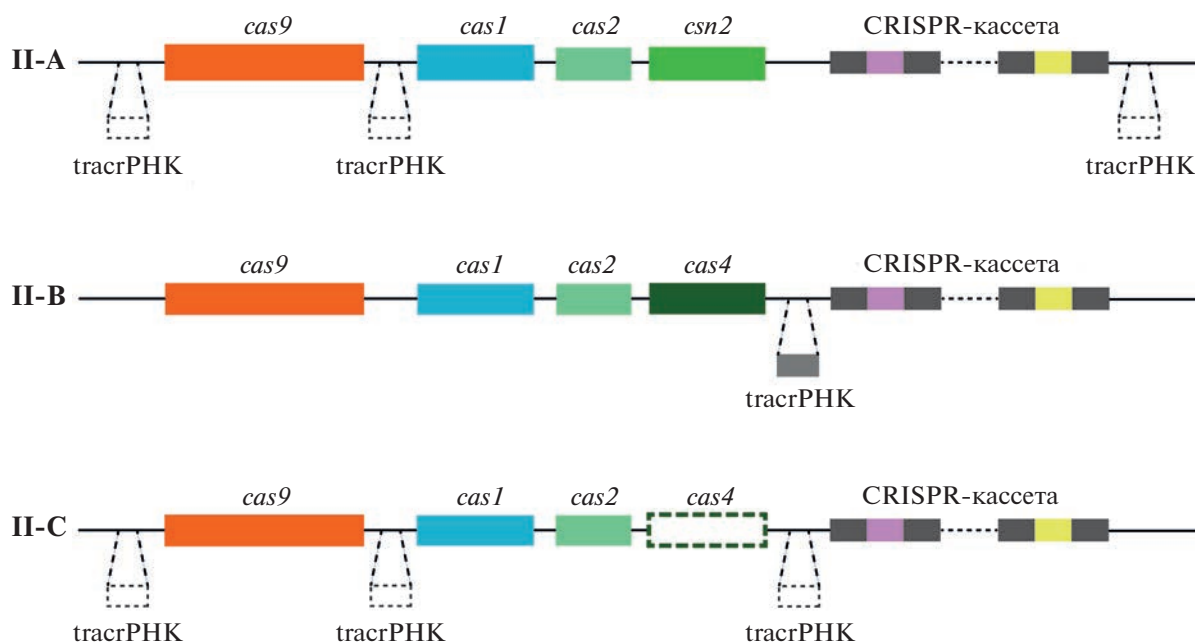


Рис. 1. Схема устройства локусов CRISPR-Cas систем подтипов II-A, II-B и II-C. Каждый локус содержит один ген транс-активирующей crPHK (trans-activating CRISPR-RNA, tracrPHK); в системах, в которых местоположение этого гена варьирует, оно отмечено прямоугольниками, нарисованными пунктирными линиями. Для системы II-C ген белка Cas4 отмечен пунктиром, так как его присутствие опционально. Направление транскрипции отдельных элементов локуса не указано, поскольку оно может отличаться у разных представителей одного подтипа.

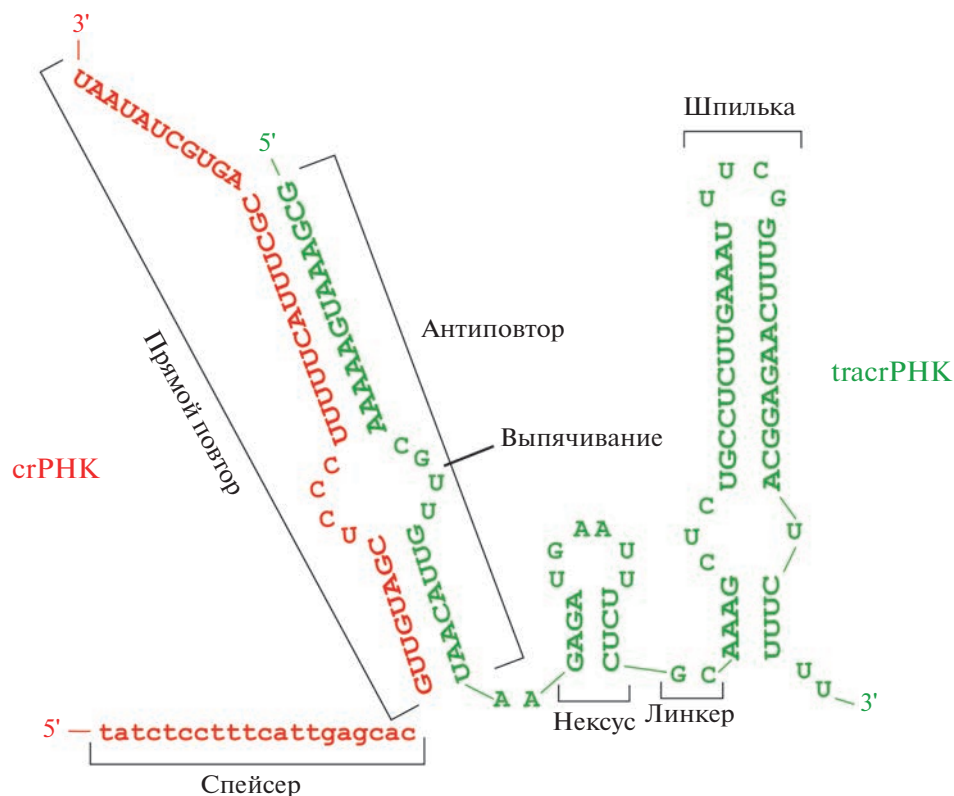


Рис. 2. Предсказанная *in silico* вторичная структура комплекса crPHK-tracrPHK системы CRISPR-Cas II типа бактерии *Pasteurella pneumotropica* ATCC 35149. Последовательность crPHK окрашена в красный цвет, tracrPHK — в зеленый.

РАМ фланкирует мишень с 3'-конца и узнается специальным доменом (РАМ Interacting Domain, PID) Cas9. Экспериментально охарактеризованные белки Cas9 узнают РАМ разной последовательности и длины (табл. S1, см. Дополнительные материалы). На данный момент четкой взаимосвязи между последовательностями PID различных Cas9-нуклеаз и конкретными последовательностями РАМ не просматривается. Существующие данные говорят лишь о наиболее предпочтительных вариантах взаимодействия определенных аминокислотных остатков PID и нуклеотидных оснований РАМ. Так, например, известно, что остатки аргинина в PID предпочтительно взаимодействуют с гуанинами, а остатки глутамина — с аденинами РАМ [11]. Подобного рода анализ помог предсказать наличие А-богатого РАМ у нуклеазы SmacCas9, что затем было подтверждено и экспериментально (SmacCas9: 5'-NAAN-3' РАМ) [12]. Однако вывести точную последовательность РАМ только лишь на основе анализа последовательности PID на данный момент не удастся. Более того, на примере белка Nme1-Cas9 и его ортологов было показано, что даже близкие по последовательности Cas9-нуклеазы могут значительно различаться по узнаваемым последовательностями РАМ (у Nme1Cas9 это 5'-NNNNGATT-3', а у близкого ортолога Nsp2Cas9 — 5'-NNNNCC-3') [13]. Следовательно, экспериментальное определение РАМ-последовательности — абсолютно необходимый шаг для характеристики не исследованных ранее белков Cas9 и принятии решения о перспективности их использования для геномного редактирования.

Таким образом, поиск и первичный (*in vitro*) анализ новых Cas9-нуклеаз включают следующие основные этапы:

1. Биоинформатический поиск генов, кодирующих Cas9-нуклеазы и соответствующие tracrRNA, и предсказание последовательности зрелой crRNA.
2. Гетерологическая экспрессия гена *cas9* в модельном штамме и получение рекомбинантного белка.
3. Нарботка crRNA и tracrRNA.
4. Демонстрация программируемого расщепления мишени рекомбинантной нуклеазой Cas9 и идентификация РАМ.

Полученная в результате выполнения этих этапов информация используется для принятия решения о перспективности дальнейших исследований, в частности оценки работы нового редактора в эукариотических клетках. Ниже приведено подробное описание алгоритма исследования для каждого из этапов первичного анализа на примере нуклеазы PpCas9, охарактеризованной в работе “PpCas9 from *Pasteurella pneumotropica* — a compact Type II-C Cas9 ortholog active in human cells” Я. Федоровой (I. Fedorova) и соавт. [14].

БИОИНФОРМАТИЧЕСКИЙ ПОИСК НОВЫХ СИСТЕМ CRISPR-Cas II ТИПА

Ниже мы приводим пошаговое описание такого поиска.

1. Выбрать нуклеотидную последовательность, в которой будет проведен поиск системы CRISPR-Cas II типа.

Поиск можно вести как с полногеномными последовательностями, так и с метагеномными данными. Мы использовали как данные из общедоступных баз, так и результаты секвенирования, полученные в нашей лаборатории (использование собственных данных облегчает последующее патентование редакторов). В случае PpCas9 поиск проводили в находящейся в общем доступе последовательности генома бактерии *Rodentibacter pneumotropicus* ATCC 35149 ([*Pasteurella*] *pneumotropica*; GenBank Acc. No BBIX01000009.1).

2. Провести поиск CRISPR-кассеты.

Использовать программу CRT CRISPR Recognition Tool [15]. Минимальное число повторов в кассете (—minNR) — 2. Остальные параметры — по умолчанию.

Альтернативные программы для поиска: CRISPRCasFinder [16], PILER-CR [17], CRISPR-Detect [18], CRISPRidentify [19].

3. В обнаруженной в результате поиска нуклеотидной последовательности выбрать участок, окружающий CRISPR-кассету и содержащий 30000 п.н. с каждой ее стороны.

4. В выбранном участке определить все открытые рамки считывания (ORFs) длиной более 900 кодонов.

Использовать модель для генетического кода 11 (Бактерии, Археи и Пластиды растений). В качестве старт-кодонов использовать любой смысловой кодон.

5. Подготовить профиль HMM (Hidden Markov Model) для белков Cas9.

Для построения профиля использовать команду **hmmbuild** пакета HMMER [20] из множественного выравнивания последовательностей Cas9-нуклеаз. Выравнивание можно подготовить самостоятельно или использовать доступное в сторонних базах данных (Conserved Domain Database (CDD) или общий fasta-файл типа II в zip-архиве S1, см. Дополнительные материалы).

6. Использовать HMM-профиль Cas9 из п. 5 для поиска среди ORFs, предсказанных в п. 4, при помощи команды **hmmsearch** пакета HMMER.

7. Для более корректного определения подтипа гомологов Cas9, обнаруженных в п. 6, анализ повторяют, используя профили нуклеаз подтипов II и V (fasta-файлы в zip-архиве S1, см. Дополнительные материалы).

Дополнительная проверка с использованием профилей близкородственной группы (для типа II это системы типа V) рекомендуется для отсеивания ложноположительных совпадений с искомым профилем. Для дальнейшей работы выбирают те ORFs, для которых получен максимальный score с профилями интереса (II-A, II-B, или II-C).

8. Проверить уникальность отобранных ORFs.

Для этого построить множественное выравнивание найденных ORFs и ранее описанных Cas9 (табл. S1, см. Дополнительные материалы). Вычислить попарные дистанции, используя матрицу BLOSUM62. Отобрать те ORFs, для которых не нашлось пары с дистанцией менее 0.5.

9. Идентифицировать ген *tracrPHK*.

Существует ряд программ, предназначенных для поиска последовательностей *tracrPHK*. Например, алгоритм S. Dooley и др. [21] и CRISPR-*tracrRNA* [22]. Рекомендуется провести первичный поиск с использованием одной из этих программ с менее строгими, чем с заданными по умолчанию, параметрами (для CRISPR*tracrRNA*: `-anti_repeat_similarity_threshold 0.3 -anti_repeat_coverage_threshold 0.3`). Отсеять предсказанные последовательности, которые находятся на расстоянии более чем 4000 п.н. от начала или конца ORF изучаемой Cas9-нуклеазы. Стоит отметить, что программы зачастую некорректно предсказывают начало и конец гена *tracrPHK*. В связи с этим рекомендуется использовать их только для определения приблизительного местоположения искомого гена, а затем вручную индивидуально оценивать каждый из предсказанных кандидатов, ориентируясь на следующие известные характеристики *tracrPHK*:

- Местоположение *tracrPHK* в локусе (зависит от подтипа системы, рис. 1).
- Последовательность *tracrPHK* не должна превышать длину в 150 п.н., длина антиповтора должна равняться приблизительно 24 п.н.
- “Хвостовая” часть *tracrPHK*, идущая за антиповтором, должна складываться в определенные вторичные структуры — так называемый нексус и следующие за ним шпильки, число которых может различаться (рис. 2). Терминальная шпилька *tracrPHK* вместе с примыкающей последовательностью из нескольких (чаще всего пяти) урацилов образует Rho-независимый терминатор транскрипции [23].

Существующие варианты вторичных структур *tracrPHK* подробно представлены в работе S. Dooley и соавт. [21]. Для предсказания и визуализации вторичной структуры РНК можно использовать программу RNAfold [24].

Рекомендуется отбирать кандидатов, для которых биоинформатически предсказываемая *tracrPHK* соответствует всем перечисленным ха-

рактеристикам. Такие кандидаты могут быть использованы для наработки РНК и подтверждения функции в экспериментах *in vitro*.

Если предсказать *tracrPHK* биоинформатически не удастся, можно использовать подход RNAseq (секвенирование РНК): ввести генетическую конструкцию с клонированным локусом изучаемой CRISPR-Cas системы в клетки *Escherichia coli*, а затем выделить из них РНК и провести высокопроизводительное секвенирование. Данный подход позволяет не только обнаружить *tracrPHK* (и *crPHK*) [25], но и корректно определить их границы [26].

10. Для дополнительной оценки обнаруженной CRISPR-Cas системы II типа проверить наличие остальных компонентов, характерных для конкретного подтипа.

Для этого повторить этапы 4, 5 и 6 с HMM-профилями Cas1, Cas2, Csn2, Cas4 (fasta-файлы в zip-архиве S1, см. Дополнительные материалы). Минимальное число кодонов при определении ORF (п. 4): 150 для Cas1, 50 для Cas2, 150 для Csn2 и 80 для Cas4. Этот шаг не считается обязательным, поскольку полное или частичное отсутствие данных компонентов не означает, что предсказанный Cas9-кандидат не обладает нуклеазной активностью.

11. Предсказать структуру зрелой *crPHK*.

Для этого необходимо определить направление транскрипции CRISPR-кассеты. Это можно сделать визуально, поскольку DR уже охарактеризованных CRISPR-Cas систем II типа похожи между собой: например, DR систем II типа бактерий *Neisseria meningitidis* 8013 и *Otariodibacter oris* DSM 23800 различаются всего тремя позициями (табл. S1, см. Дополнительные материалы). Также определению правильного направления DR может помочь предсказанная *tracrPHK* — последовательность DR должна быть способна образовывать дуплекс с участком антиповтора (рис. 2). В случае затруднений при предсказании направления транскрипции CRISPR-кассеты рекомендуется протестировать оба направления DR в экспериментах *in vitro* или воспользоваться подходом RNAseq (см. выше).

В последовательность *crPHK*, как правило, входит лишь часть повтора, комплементарная антиповтору *tracrPHK*. При подборе *crPHK* для *in vitro* проверки можно начать с варианта с полноразмерным повтором, а затем проанализировать несколько укороченных вариантов.

Помимо DR зрелая *crPHK* содержит участок спейсера, комплементарного целевой мишени. Известно, что изменение длины данного участка значительно влияет на эффективность интерференции. Так, например, для нуклеаз SpCas9, SaCas9, Cje1Cas9 и PrCas9 наибольшая эффективность в эукариотических клетках наблюдается

при длинах спейсера 20, 21, 22, 24 п.н. соответственно [14, 27, 28]. В связи с этим рекомендуется тестировать sgРНК с несколькими вариантами длин спейсера.

Программные коды, реализующие отдельные этапы биоинформатического поиска, можно скачать из репозитория GitHub по ссылке https://github.com/CRISPRNanoBio/Cas9_Pipeline.

По результатам биоинформатического поиска (рис. S1, см. Дополнительные материалы) для дальнейшей экспериментальной работы отбирают те системы-кандидаты, которые соответствуют следующим критериям:

- ORF Cas9 уникальна, то есть кодируемая ею аминокислотная последовательность непохожа на ранее охарактеризованные нуклеазы.
- Система содержит в себе все обязательные компоненты: Cas9, CRISPR-кассета, tracrРНК. Наличие других компонентов системы желательно, но не относится к обязательным условиям.

Организация локуса CRISPR-Cas системы II типа бактерии *Pasteurella pneumotropica* ATCC 35149, определенная по результатам описанного выше биоинформатического анализа, представлена в табл. S2 Дополнительных материалов.

КЛОНИРОВАНИЕ ГЕНА НУКЛЕАЗЫ-КАНДИДАТА

На следующем этапе работы необходимо получить генетическую конструкцию, несущую ген нуклеазы-кандидата, для дальнейшего получения ее рекомбинантной версии.

1. Получить нуклеотидную последовательность гена нуклеазы-кандидата.

В случае недоступности образцов ДНК с выбранной для изучения системой, например, если система была обнаружена в открытых базах метагеномных данных, необходимо заказать синтетическую последовательность гена нуклеазы-кандидата.

2. Выбрать вектор для клонирования.

Для получения рекомбинантных белков Cas9 рекомендуется использовать экспрессионные векторы, позволяющие проводить дальнейшую хроматографическую очистку целевого белка аффинным методом. Комбинация аффинной хроматографии с последующей очисткой гель-фильтрацией позволяет получить препарат нуклеазы достаточной чистоты для проведения экспериментов *in vitro*.

Для этих целей хорошо подходит вектор pET21a (рис. S2, см. Дополнительные материалы), так как к полилинкеру, используемому для клонирования, примыкает последовательность, кодирующая шесть гистидинов, что позволяет проводить очистку рекомбинантных белков на металл-хелат-

ном аффинном сорбенте (Ni-NTA-агарозе). Транскрипция целевого гена происходит с промотора РНК-полимеразы фага Т7. Транскрипции гена РНК-полимеразы и целевого гена регулируется репрессором лактозного оперона.

3. Проанализировать последовательности гена нуклеазы-кандидата и вектора pET21a и выбрать оптимальные для клонирования сайты рестрикции.

Так, например, для гена *ppcas9* были выбраны сайты эндонуклеаз рестрикции XhoI и NdeI.

4. Провести клонирование целевого гена.

Альтернативой классическому подходу с использованием рестрикции и лигирования может быть клонирование по Гибсону, при котором у вектора и вставки есть частично перекрывающиеся последовательности на концах. Использовать NEBuilder HiFi DNA Assembly Master Mix (NEB) согласно протоколу производителя.

5. Верифицировать вставку целевого гена нуклеазы-кандидата в вектор.

Для этого использовать метод рестрикционно-го анализа: контрольный вектор pET21a и полученную на его основе рекомбинантную плазмиду обработать парой эндонуклеаз рестрикции.

В случае обработки вектора pET21a эндонуклеазами рестрикции HindIII и XbaI образуются фрагменты длиной 5300 и 103 п.н.; обработка вектора pET21a с клонированным геном *ppcas9* (pET21a_PpCas9) этими эндонуклеазами рестрикции приводит к образованию трех фрагментов длиной 5400, 2500 и 630 п.н. (рис. S3, см. Дополнительные материалы).

Для дополнительной верификации провести секвенирование клонированного гена по Сэнгеру.

НАРАБОТКА РЕКОМБИНАНТНОЙ НУКЛЕАЗЫ-КАНДИДАТА

Получение рекомбинантной версии нуклеазы-кандидата включает в себя следующие основные этапы:

1. Экспрессия гена нуклеазы-кандидата в клетках продуцентах
2. Проверка продукции нуклеазы
3. Проверка растворимости нуклеазы
4. Лизис клеток-продуцентов
5. Аффинная хроматография
6. Гель-фильтрация

Ниже приведены подробные протоколы каждого из этапов.

Экспрессия гена нуклеазы-кандидата в клетках-продуцентах

1. Трансформировать клетки *E. coli* Rosetta (DE3) плазмидой pET21a_PpCas9 стандартными методами и высеять на чашки Петри, содержащие ампициллин и хлорамфеникол.

2. Перенести единичную колонию в 5 мл среды LB, содержащей ампициллин и хлорамфеникол, и растить в течение ночи при 37°C.

3. Развести 5 мл ночной культуры в 500 мл среды LB, содержащей ампициллин.

4. Инкубировать клеточные культуры при температуре 37°C и равномерной аэрации на вращающейся качалке до достижения оптической плотности 0.6 при длине волны 600 нм.

5. Индуцировать транскрипцию клонированного целевого гена добавлением изопропил- β -D-1-тиогалактопиранозид (ИПТГ) до концентрации 1 мМ.

6. Продолжить рост клеточных культур при температуре 20–24°C в течение 4–6 ч.

7. Собрать клетки центрифугированием в настольной центрифуге на скорости 4000–5000 об/мин (~3500 g) в течение 30 мин.

Полученные осадки можно сразу использовать для выделения и очистки белка либо заморозить и хранить при температуре –20°C.

Проверка продукции нуклеазы-кандидата

1. Отобрать по 1 мл клеточных культур до внесения индуктора и после инкубации с индуктором (ИПТГ) (п. 4 и п. 6 “*Экспрессия гена нуклеазы-кандидата в клетках-продуцентах*”).

Для корректного сравнения образцов и детекции продукта экспрессии гена важно, чтобы образцы до и после индукции содержали одинаковое число клеток.

2. Собрать клетки центрифугированием в настольной мини-центрифуге на скорости 13000 об/мин (~12000 g) в течение 10 мин.

3. Удалить надосадочную жидкость.

4. Ресуспендировать осадки в небольшом объеме (~100 мкл) фосфатно-солевого буфера (PBS).

5. Отобрать из обеих проб по 15 мкл и проанализировать наличие индуцированного белка при помощи электрофореза в денатурирующем 10%-ном ПААГ по методике Лэммли [29].

Результаты проверки продукции нуклеазы PpCas9 представлены на рис. S4 Дополнительных материалов. Продукт, соответствующий ожидаемой молекулярной массе целевого белка (121 кДа), обнаружен только в образце “После индукции”.

Проверка растворимости нуклеазы-кандидата

Для хроматографической очистки белка Cas9 необходимо убедиться в его растворимости.

1. Подготовить лизисный буфер.

Состав: 50 мМ Трис, 500 мМ NaCl, 10 мМ имидазол, 1 мМ β -меркаптоэтанол, pH 7.5 (при комнатной температуре).

2. Отобрать 1 мл клеточной культуры после инкубации с индуктором (ИПТГ) (п. 6 “*Экспрессия гена нуклеазы-кандидата в клетках-продуцентах*”).

6. Собрать клетки центрифугированием в настольной мини-центрифуге на скорости 13000 об/мин (~12000 g) в течение 10 мин.

7. Удалить надосадочную жидкость.

3. Ресуспендировать осадок в 500 мкл лизисного буфера с добавлением лизоцима (концентрация 1 мг/мл).

4. Инкубировать полученную клеточную суспензию на льду в течение 15 мин.

5. Разрушить клеточные оболочки методом ультразвуковой дезинтеграции на ледяной бане.

Ультразвуковую дезинтеграцию проводили на приборе LABSONIC (“Sartorius AG”, Германия): амплитуда ультразвука 30%, интервал времени активности 0.2 с.

Продолжительность одного цикла обработки – 20 с, с последующим охлаждением в течение 40 с, общее время обработки – 10 мин.

6. Центрифугировать в настольной центрифуге с охлаждением на скорости 13000 об/мин (~16000 g) в течение 20 мин при 4°C.

7. Отобрать надосадочную жидкость.

8. К осадку добавить 500 мкл лизисного буфера, ресуспендировать.

9. Отобрать из ресуспендированного осадка и из надосадочной жидкости по 15 мкл и проанализировать наличие индуцированного белка методом электрофореза в денатурирующем 10%-ном ПААГ по методике Лэммли [29].

Результаты проверки для нуклеазы PpCas9 представлены в Дополнительных материалах, рис. S5. В пробе “Надосадочная жидкость” присутствует продукт, соответствующий ожидаемой молекулярной массе целевого белка, что свидетельствует о том, что полученный белок PpCas9 растворим.

Лизис клеток-продуцентов

1. Осадок, полученный из 500 мл жидкой клеточной культуры, содержащей индуцированный белок, разморозить на льду в течение 30 мин.

2. Ресуспендировать в 15 мл лизисного буфера с добавлением лизоцима (концентрация 1 мг/мл).

3. Инкубировать на льду в течение 30 мин.

4. Разрушить клеточные оболочки методом ультразвуковой дезинтеграции на ледяной бане.

Параметры прибора аналогичны таковым на этапе “Проверка растворимости нуклеазы-кандидата” п. 5. Общее время обработки увеличено и составляет 40 мин.

5. Центрифугировать в настольной центрифуге с охлаждением на скорости 11000 об/мин (~16000 g) в течение 60 мин при 4°C.

6. Надосадочную жидкость профильтровать через шприцевой фильтр с целлюлозно-нитратной мембраной (размер пор 0,4 мкм).

Для выделения функционально активного белка проводят двухстадийную очистку лизата методами аффинной хроматографии и гель-фильтрации. Очистку нуклеазы PrCas9 проводили на приборе АКТА (“GE Healthcare”, США). Можно использовать оборудование других производителей. Для аффинной хроматографии также возможно использование гравитационных колонок.

Аффинная хроматография

1. Полученный после фильтрации лизат нанести на колонку HisTrap HP 1 mL (“GE Healthcare”), содержащую иммобилизованные ионы никеля, на скорости 0,5 мл/мин.

2. Промыть колонку от неспецифически связанных белков лизисным буфером (5–10 объемов колонки).

3. Промыть колонку лизисным буфером, содержащим 30 мМ имидазол (15 объемов колонки).

4. Элюировать целевой белок лизисным буфером, содержащим 300 мМ имидазол (5 объемов колонки).

Гель-фильтрация

1. Выбрать фракции, содержащие наибольшее количество белка, и объединить их.

В рассматриваемом примере нуклеаза PrCas9 выходила во фракциях 3 и 4 (рис. S6, см. Дополнительные материалы).

2. При помощи центрифужных концентраторов Amicon (“Merck Millipore”, США) с фильтром на 100 кДа сконцентрировать объединенные фракции до объема 600 мкл.

3. Нанести образец на гель-фильтрационную колонку Superdex 200 Increase 10/300 GL (“Sigma-Aldrich”, США), уравновешенную буфером: 50 мМ Трис-HCl (pH 7,5 при комнатной температуре), 500 мМ NaCl, 1 мМ DTT.

4. Ориентируясь на показания интенсивности поглощения при длине волны 280 нм и известные калибровочные данные используемой хроматографической колонки, собрать фракции элюата, содержащие белок.

5. Проверить наличие и чистоту целевого белка в собранных фракциях методом электрофореза в денатурирующем 10%-ном ПААГ по методике Лэммли [29].

В рассматриваемом примере нуклеазы PrCas9 для электрофореза были взяты фракции 4–6 (рис. S7, см. Дополнительные материалы).

6. Объединить фракции и сконцентрировать при помощи центрифужных концентраторов с фильтром на 100 кДа до конечной концентрации целевого белка 1–3 мг/мл (для используемых объемов жидкой клеточной культуры, указанных выше, конечный объем на данном этапе, как правило, составляет 250–300 мкл).

7. К концентрату добавить глицерин до конечной концентрации 10%.

8. Полученный образец белка разделить на аликвоты объемом 5–10 мкл, распределить их по микробиркам и заморозить в жидком азоте.

9. Очищенный белок хранить при температуре –80°C. Остатки от использованных образцов повторно не замораживать.

Описанные протоколы получения рекомбинантного белка применимы не только для ортологов Cas9-нуклеаз, но и для представителей других типов: например, Cas12a и Cas12e [30].

Возможные проблемы. Самая часто встречающаяся проблема связана с нерастворимостью целевой нуклеазы в лизисном буфере (раздел “Проверка растворимости нуклеазы-кандидата”). В некоторых случаях для повышения растворимости целевого белка рекомендуется проводить индукцию при более низких температурах (раздел “Экспрессия гена нуклеазы-кандидата в клетках-продуцентах”, п. 6). Также решению данной проблемы может способствовать наработка белкового продукта гена нуклеазы-кандидата, слитого с белком, связывающим мальтозу (Maltose-Binding Protein; MBP) [31].

ДЕМОНСТРАЦИЯ ПРОГРАММИРУЕМОГО РАСЩЕПЛЕНИЯ МИШЕНИ НУКЛЕАЗОЙ-КАНДИДАТОМ И ИДЕНТИФИКАЦИЯ РАМ

In vitro расщепление РАМ-библиотеки нуклеазой-кандидатом

Как было сказано ранее, для узнавания и расщепления мишени нуклеазой Cas9 последняя должна быть фланкирована определенной последовательностью РАМ (см. “Организация CRISPR-Cas систем II типа”). Понятно, что для нуклеазы кандидата она неизвестна. Поэтому для проверки наличия нуклеазной активности используется специальный набор последовательностей, в котором участок, соответствующий РАМ, заменен случайными нуклеотидами. Такой набор после-

довательностей далее обозначен как “РАМ-библиотека”. Часть последовательностей РАМ-библиотеки будет содержать участок случайных нуклеотидов, совпадающий с РАМ. В результате взаимодействия с эффекторным комплексом нуклеазы такие последовательности подвергнутся расщеплению.

Мы использовали *in vitro* РАМ-библиотеку, содержащую 7 случайных нуклеотидов (7N), так как РАМ-последовательность известных Cas9 чаще всего короче 7 н. (табл. S1, см. Дополнительные материалы).

Анализ реакции расщепления РАМ-библиотеки проводят электрофорезом в агарозном геле. При наличии нуклеазной активности Cas9-кандидата на агарозном геле будут видны фрагменты нескольких размеров. Одни соответствуют последовательностям, которые не подверглись расщеплению, вторые — расщепленным последовательностям. Используемая 7N РАМ-библиотека содержит последовательности длиной 374 п.н. Выбранный нами участок спейсера *sgRNA* позволяет расщеплять последовательности такой библиотеки на фрагменты длиной приблизительно 49 и 325 п.н. (рис. S8a, см. Дополнительные материалы).

Далее приводится подробный протокол проверки активности нуклеазы-кандидата с использованием 7N РАМ-библиотеки.

1. Подготовить 7N *in vitro* РАМ-библиотеку.

В представленной здесь работе за основу РАМ-библиотеки был взят фрагмент гена *grin2b* человека. Аналогичным образом можно использовать любые нуклеотидные последовательности. Для более удобной визуализации результата на агарозном геле рекомендуется использовать последовательности длиной более 300 п.н.

Семь случайных нуклеотидов вводят в нужную последовательность при помощи амплификации с использованием прямого и обратного праймеров для 7N-библиотеки. Прямой праймер при этом включает участок из 7 случайных нуклеотидов (табл. S2, см. Дополнительные материалы).

Провести очистку целевого фрагмента из агарозного геля.

2. Синтезировать *sgRNA* и *tracrRNA* (табл. S3, см. Дополнительные материалы), используя набор для проведения *in vitro* транскрипции HiScribe T7 High Yield RNA Synthesis Kit (“NEB”, США) согласно протоколу производителя.

Последовательность DR *sgRNA* и последовательность *tracrRNA* были определены ранее (см. “Биоинформатический поиск новых систем CRISPR-Cas II типа” п. 9 и п. 11). В качестве спейсера *sgRNA* выбирают область, примыкающую к 7N-участку *in vitro* РАМ-библиотеки с 5'-конца.

Матрицы для синтеза *sgRNA* и *tracrRNA* должны содержать последовательность промотора T7 РНК-полимеразы (табл. S2, см. Дополнительные материалы). Матрица может быть либо полностью двухцепочечная, либо одноцепочечная с двухцепочечной областью промотора T7 РНК-полимеразы.

3. Провести очистку полученных в результате *in vitro* транскрипции *sgRNA* и *tracrRNA* при помощи набора для очистки РНК — RNA Clean & Concentrator-25 (“Zymo Research”, США) согласно протоколу производителя.

4. Спектрофотометрически измерить концентрации синтезированных *sgRNA*, *tracrRNA* и 7N *in vitro* РАМ-библиотеки.

5. Провести *in vitro* реакцию расщепления 7N РАМ-библиотеки изучаемой нуклеазой в объеме 40 мкл.

Для этого объединить: вода (до 40 мкл), 1× CutSmart-буфер (“NEB”), 0.5 мМ DTT, 2 мкМ *sgRNA*, 2 мкМ *tracrRNA*, 400 нМ рекомбинантный белок Cas9, 100 нМ 7N *in vitro* РАМ-библиотека.

6. Провести аналогичную реакцию, не добавляя *sgRNA* (контроль).

7. Обе реакционные смеси инкубировать при 37°C в течение 30 мин.

8. Добавить по 1 мкл протеиназы К (“Thermo Fisher Scientific”, США) для деградации белков, которые могут препятствовать движению ДНК в агарозном геле. Инкубировать смесь при температуре 37°C в течение 20 мин.

9. Контроль и образец проанализировать методом геле-электрофореза в 2%-ном агарозном геле.

Возможные проблемы. Мы рекомендуем использовать 7N *in vitro* РАМ-библиотеку. Однако в ряде случаев (например, нуклеаза NmeI Cas9 и ее ближайшие ортологи, а также CjCas9) участок РАМ оказывается длиннее 7 п.н. (5'-NNNN-GATT-3' [32] и 5'-NNNNRYAC-3' [27] соответственно). В таком случае существует вероятность, что РАМ-библиотека не будет содержать последовательностей, соответствующих РАМ исследуемой нуклеазы. Отрицательный результат расщепления РАМ-библиотеки при этом ошибочно будет интерпретирован как отсутствие нуклеазной активности белка-кандидата. В таком случае можно использовать библиотеки с более длинным участком случайных нуклеотидов (8N, 9N, 10N и т.д.). Однако на следующем этапе работы (“Идентификация РАМ”) при использовании таких библиотек потребуется увеличить покрытие высокоэффективного секвенирования.

Эффективность нуклеазной активности белка Cas9 сильно зависит от условий проведения реакции. Мы рекомендуем проводить реакцию расщепления в присутствии ионов магния и калия (входят в состав буфера CutSmart) при температу-

ре 37°C. При отсутствии расщепления РАМ-библиотеки рекомендуется проверить эффективность гидролиза при различных температурах и в присутствии различных ионов (например, марганца или меди). Так, для нуклеазы PpCas9 наиболее эффективное расщепление РАМ-библиотеки зарегистрировано при 42°C (см. [14]).

Также решению проблемы отсутствия расщепления последовательностей библиотеки может способствовать изменение концентраций основных компонентов: белка, crРНК и/или tracrРНК.

Идентификация РАМ

Определить консенсус РАМ-последовательности можно по результатам анализа продуктов расщепления РАМ-библиотеки. К сожалению, на практике не удастся воссоздать идеальную ситуацию, в которой все последовательности, содержащие РАМ нуклеазы, расщепляются на 100% и могут быть полностью отделены от последовательностей, не подвергшихся расщеплению. В связи с этим используют подход, включающий анализ двух РАМ-библиотек: исходной (контроль) и библиотеки, часть последовательностей которой подверглась расщеплению нуклеазой (образец). Обе библиотеки проходят одинаковую пробоподготовку, а полученные на них результаты – сравнительный анализ. По результатам этого сравнения отбирают те последовательности, которые вероятнее всего соответствуют РАМ.

Данный подход можно реализовать двумя способами. В первом случае сначала нужно отобрать из образца последовательности, которые подверглись расщеплению. Полученный набор последовательностей далее сравнивают с контролем и отбирают те, доля которых от всех последовательностей в образце достоверно увеличилась по сравнению с их долей в контроле. Такой подход получил название “тест обогащения”. Во втором случае из образца, наоборот, отбирают те последовательности, которые не подверглись расщеплению, а при последующем сравнении с контролем ищут последовательности, доля которых достоверно уменьшилась в образце. Этот подход получил название “деплеционный тест” [33]. В представленной работе мы реализовали один из вариантов *in vitro* деплеционного теста (рис. S86, см. Дополнительные материалы).

7N *in vitro* деплеционный тест

Этот анализ проводят с использованием образца и контроля, полученных на этапе “*In vitro* расщепление РАМ-библиотеки нуклеазой-кандидатом”, п. 9.

1. Часть образца, которая не подверглась расщеплению, и контроль экстрагировать из геля. Длина экстрагируемых последовательностей 374 п.н.

2. Очищенные последовательности подготовить для анализа высокоэффективным секвенированием с помощью двухстадийной ПЦР (праймеры см. в табл. S2 Дополнительных материалов). Праймеры первого этапа содержат последовательности, которые отжигаются к участку РАМ-библиотеки длиной 92 п.н., включающему 7 случайных нуклеотидов и 20 нуклеотидов мишени, комплементарной crРНК. Дополнительно в праймеры включены фрагменты адаптеров Illumina (“Illumina”, США). Полученные на этом этапе ампликоны используют в качестве матрицы для второй ПЦР. На втором этапе адаптеры полностью достраивают и вносят последовательности, содержащие 6 или 8 рандомизированных нуклеотидов (баркоды).

3. Последовательности проанализировать на платформе Illumina одиночными или парными прочтениями с минимальной длиной прочтения 75 п.н. с каждой стороны и минимальным покрытием 10 млн прочтений на образец.

Парные прочтения в данном случае будут практически полностью перекрывать друг друга, что позволит исправить возможные ошибки секвенирования. Для дальнейшей работы пары прочтений необходимо объединить и построить соответствующие консенсусы.

4. Из полученных прочтений (либо консенсусов прямых и обратных прочтений) вырезать участки, соответствующие положению 7 случайных нуклеотидов.

5. Сравнить наборы последовательностей для контроля и образца и визуализировать полученный результат.

Визуализация результата *in vitro* деплеционного теста с помощью сервиса WebLogo [34]

В случае деплеционного теста РАМ-лого представляет собой консенсус последовательностей, частота встречаемости которых уменьшилась в образце по сравнению с контролем (деплецированные последовательности). Те позиции на полученном лого, которые характеризуются наибольшей консервативностью, выраженной в битах, рассматривают как возможные значимые позиции РАМ.

1. При помощи критерия χ^2 сравнить наборы последовательностей, соответствующих 7N-участку РАМ-библиотеки, из контроля и образца и отобрать те, которые достоверно деплецировались.

Мы рекомендуем использовать пороговое значение $p\text{-value} \leq 10^{-12}$. Однако получаемые в этом тесте данные содержат большое число как ложноположительных, так и ложноотрицательных результатов. Именно поэтому возможно увеличение или наоборот уменьшение порогового значения $p\text{-value}$ для получения лого с высококонсервативными отдельными позициями РАМ.

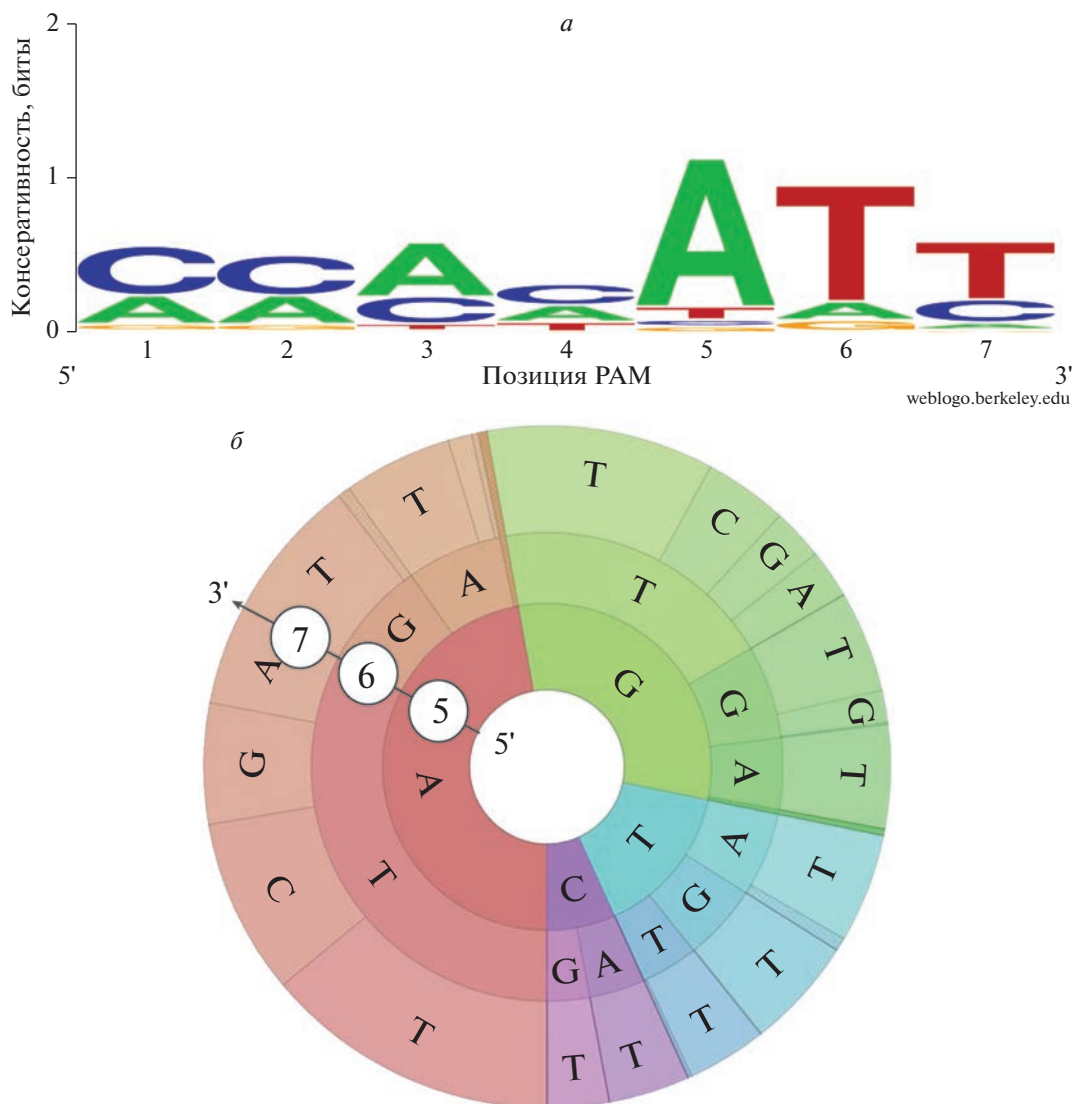


Рис. 3. Результаты 7N *in vitro* деpleционного теста для нуклеазы PpCas9. Визуализация результатов в виде PAM-лого для позиций 1–7 PAM (а) и в виде PAM-колеса для позиций 5–7 PAM (б). Адаптировано из работы [14, Creative Commons Attribution License].

2. Построить PAM-лого из деpleцированных последовательностей с использованием сервиса WebLogo (<https://weblogo.berkeley.edu/logo.cgi>).

По лого, полученном по результатам 7N *in vitro* деpleционного теста для нуклеазы PpCas9, можно предположить значимость позиций 5, 6, 7 от конца мишени (рис. 3а).

Аналогично для нуклеазы SauriCas9 по лого была выявлена консервативность позиций 3 и 4 PAM и преобладание доли G в них ([35]).

*Визуализация результата *in vitro* деpleционного теста в виде PAM-колеса [36]*

PAM-колесо (PAM wheel) строят с использованием всех вариантов последовательностей, соответствующих 7N-участку PAM-библиотеки, и ин-

формации по частоте их встречаемости в образце и контроле. Это отличает PAM-колесо от PAM-лого, для построения которого берут только деpleцированные последовательности.

Однако, у PAM-колеса есть ограничение – результаты удобно визуализировать только для небольшого числа позиций PAM (до четырех). Таким образом, этот инструмент можно считать дополнением к уже построенному PAM-лого.

1. Из имеющихся наборов последовательностей, соответствующих 7N-участку PAM-библиотеки, необходимо отобрать только интересные позиции (до четырех), которые могут иметь значимость согласно PAM-лого. После этого необходимо уточнить, сколько раз встречается каждый новый вариант последовательности.

Так, для нуклеазы PrCas9 отбирают позиции 5, 6 и 7. Например, если в изначальном наборе присутствует одна последовательность **CCCCATT** и одна последовательность **GGGGATT**, в новый набор необходимо внести две последовательности **ATT**.

2. Для каждого нового варианта последовательности, соответствующей участку случайных нуклеотидов, рассчитать коэффициент деплеции (D_{coef}):

$$3. D_{\text{coef}} = \log_2 \frac{N_{\text{sample}}}{N_{\text{control}}} \frac{N_{\text{seq_var.control}}}{N_{\text{seq_var.sample}}},$$

где N_{sample} — общее число последовательностей в образце; N_{control} — общее число последовательностей в контроле; $N_{\text{seq_var.sample}}$ — число последовательностей определенного варианта в образце; $N_{\text{seq_var.control}}$ — число последовательностей определенного варианта в контроле.

В исходном методе используют “тест обогащения” и рассчитывают соответствующий коэффициент обогащения. Таким образом, D_{coef} рассчитывается как коэффициент, обратный рассчитанному Leenay и др. [36].

4. Полученные значения коэффициентов записать в таблицу специального формата в файл KronaExcelTemplate (табл. S4, см. Дополнительные материалы), в котором проводят построение РАМ-колеса.

Дополнительная проверка результатов с построением РАМ-колеса для нуклеазы PrCas9 подтверждает вероятную значимость позиций 5, 6 и 7. Также можно предположить, что присутствие нуклеотидов А либо Г равновероятно допустимо в позиции 5 РАМ (рис. 3б). По результатам построения РАМ-лого и РАМ-колеса для нуклеазы PrCas9 можно сделать вывод о том, что ее РАМ соответствует консенсусу 5'-NNNNR~~TT~~-3'.

Для нуклеазы FrCas9, с РАМ-последовательностью 5'-NNTA-3', при визуализации РАМ-колеса для позиций 2, 3 и 4 РАМ четко видно преобладание Т в третьем положении и А в четвертом ([37]).

Экспериментальное подтверждение предсказанной РАМ-последовательности

Проверку значимости отдельных нуклеотидных позиций проводят по представленной ниже методике.

1. Среди наиболее деплецированных последовательностей, отобранных для построения РАМ-лого, выбрать одну.

Для нуклеазы PrCas9 была выбрана последовательность **СААСАТТ** (рис. 4).

Далее необходимо проверить, что выбранная последовательность распознается нуклеазой как РАМ.

2. Синтезировать ДНК-мишень по аналогии с получением 7N-библиотеки для *in vitro* деплеционного теста, заменив в прямом праймере рандомизированный участок случайных нуклеотидов на последовательность, обратнoкомплементарную выбранному варианту РАМ.

3. Провести *in vitro* реакцию расщепления полученного фрагмента в объеме 20 мкл.

Для этого объединить: вода (до 20 мкл), 1× CutSmart-буфер (“NEB”), 0.5 мМ DTT, 2 мкМ crPНК, 2 мкМ tracrPНК, 400 нМ рекомбинантный белок Cas9, 40 нМ фрагмент ДНК с вариантом РАМ из п. 1.

4. Провести контрольную реакцию без добавления crPНК.

5. Обе реакционные смеси инкубировать при 37°C в течение 30 мин (либо при той температуре, которая обеспечивает наиболее эффективное расщепление мишеней нуклеазой).

6. Добавить по 1 мкл протеиназы К (“Thermo Fisher Scientific”). Инкубировать при 37°C в течение 20 мин.

7. Продукты реакций проанализировать методом гель-электрофореза в 2%-ном агарозном геле.

Дальнейшие этапы проводить, если фрагмент ДНК, содержащий выбранный вариант последовательности РАМ, подвергся расщеплению. В случае отсутствия расщепления необходимо вернуться к п. 1 и выбрать другую последовательность.

8. Составить набор последовательностей, которые могут быть модификациями выбранного варианта.

Каждая последовательность такого набора включает замену одной нуклеотидной позиции; при этом пуриновые основания заменить на пиримидиновые и наоборот.

Для PrCas9 был создан следующий набор из 7 последовательностей:

СААСАТТ (исходный вариант)
АААСАТТ
СТАСАТТ
САТСАТТ
СААГАТТ
СААСТТТ
СААСААТ
СААСАТГ

9. Для набора полученных последовательностей повторить п. 2–7.

По отсутствию расщепления нуклеазой последовательности из данного набора можно судить о значимости соответствующей позиции РАМ.

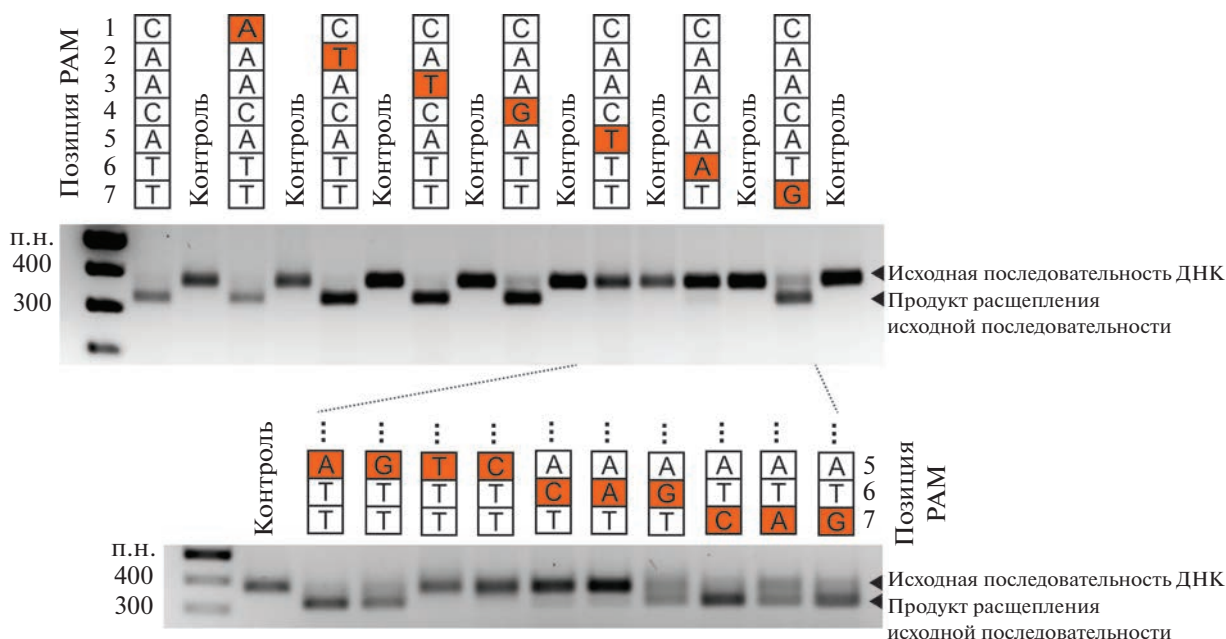


Рис. 4. Проверка значимости отдельных нуклеотидных позиций в РАМ-последовательности РрCas9. Для позиций 5, 6 и 7 проведена дополнительная проверка влияния замен на все четыре варианта нуклеотидов (нижняя часть рисунка). Замены в позиции 5 на Т или С и любые замены в позиции 6 блокируют или значительно снижают расщепление соответствующей последовательности нуклеазой. Замены в позиции 7 снижают эффективность расщепления. Адаптировано из работы [14, Creative Commons Attribution License].

Дополнительно для наиболее значимых позиций РАМ-последовательности (или для всех позиций) можно проверить, как замены на все возможные варианты нуклеотидов влияют на эффективность распознавания и расщепления нуклеазой мишени.

Для нуклеазы РрCas9 вышеописанная проверка показала, что значимы нуклеотиды в позициях 5, 6 и 7 РАМ. Также показано, что замена А на Г в позиции 5 не влияет на эффективность расщепления. Таким образом, удалось подтвердить, что РАМ нуклеазы РрCas9 соответствует консенсусу 5'-NNNNRTT-3', предсказанному по результатам деплеционного теста (рис. 4).

Для дополнительной проверки соответствия РАМ-последовательности нуклеазы предсказанному консенсусу также рекомендуется проверить эффективность расщепления различных мишеней, которые фланкированы этим консенсусом ([14]). Реакцию проводят с любым фрагментом ДНК, содержащим участки, совпадающие с консенсусом РАМ (условия реакции аналогичны тем, которые используются при проверке значимости отдельных нуклеотидных позиций).

ЗАКЛЮЧЕНИЕ

Мы многократно проверили приведенные протоколы на практике и доказали их эффективность. Используя их, мы успешно охарактеризо-

вали ряд новых CRISPR-Cas редакторов: нуклеазы РрCas9 из бактерии *Pasteurella pneumotropica* [14], CcCas9 из бактерии *Clostridium cellulolyticum* H10 [38], DfCas9 из бактерии *Deftluviimonas* sp. 20V17 [14] и CoCas9 из бактерии *Capnocytophaga ochracea* (в печати).

Работа выполнена при финансовой поддержке Министерства науки и высшего образования Российской Федерации (грант № 075-15-2021-1062). Исследования А.А.В., П.А.С. и А.Н.А. также поддержаны грантом Российского научного фонда (грант № 21-14-00122).

Настоящая статья не содержит описания каких-либо исследований с участием людей или животных в качестве объектов.

Авторы заявляют об отсутствии конфликта интересов.

СПИСОК ЛИТЕРАТУРЫ

1. Jansen R., van Embden J.D.A., Gastra W., Schouls L.M. (2002) Identification of genes that are associated with DNA repeats in prokaryotes. *Mol. Microbiol.* **43**, 1565–1575.
2. Barrangou R., Fremaux C., Deveau H., Richards M., Boyaval P., Moineau S., Romero D.A., Horvath P. (2007) CRISPR provides acquired resistance against viruses in prokaryotes. *Science*. **315**, 1709–1712.
3. Marraffini L.A., Sontheimer E.J. (2008) CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science*. **322**, 1843–1845.

4. Garneau J.E., Dupuis M.-È., Villion M., Romero D.A., Barrangou R., Boyaval P., Fremaux C., Horvath P., Magadán A.H., Moineau S. (2010) The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature*. **468**, 67–71.
5. Makarova K.S., Wolf Y.I., Iranzo J., Shmakov S.A., Alkhnbashi O.S., Brouns S.J.J., Charpentier E., Cheng D., Haft D.H., Horvath P., Moineau S., Mojica F.J.M., Scott D., Shah S.A., Siksnys V., Terns M.P., Venclovas Č., White M.F., Yakunin A.F., Yan W., Zhang F., Garrett R.A., Backofen R., van der Oost J., Barrangou R., Koonin E.V. (2020) Evolutionary classification of CRISPR-Cas systems: a burst of class 2 and derived variants. *Nat. Rev. Microbiol.* **18**, 67–83.
6. Cong L., Ran F.A., Cox D., Lin S., Barretto R., Habib N., Hsu P.D., Wu X., Jiang W., Marraffini L.A., Zhang F. (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science*. **339**, 819–823.
7. Jinek M., East A., Cheng A., Lin S., Ma E., Doudna J. (2013) RNA-programmed genome editing in human cells. *Elife*. **2**, e00471.
8. Mali P., Yang L., Esvelt K.M., Aach J., Guell M., Di-Carlo J.E., Norville J.E., Church G.M. (2013) RNA-guided human genome engineering via Cas9. *Science*. **339**, 823–826.
9. Deltcheva E., Chylinski K., Sharma C.M., Gonzales K., Chao Y., Pirzada Z.A., Eckert M.R., Vogel J., Charpentier E. (2011) CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature*. **471**, 602–607.
10. Shah S.A., Erdmann S., Mojica F.J.M., Garrett R.A. (2013) Protospacer recognition motifs: mixed identities and functional diversity. *RNA Biol.* **10**, 891–899.
11. Luscombe N.M., Laskowski R.A., Thornton J.M. (2001) Amino acid–base interactions: a three-dimensional analysis of protein–DNA interactions at an atomic level. *Nucleic Acids Res.* **29**, 2860–2874.
12. Chatterjee P., Lee J., Nip L., Koseki S.R.T., Tysinger E., Sontheimer E.J., Jacobson J.M., Jakimo N. (2020) A Cas9 with PAM recognition for adenine dinucleotide. *Nat. Commun.* **11**, 2474.
13. Wei J., Hou L., Liu J., Wang Z., Gao S., Qi T., Gao S., Sun S., Wang Y. (2022) Closely related type II-C Cas9 orthologs recognize diverse PAMs. *Elife*. **11**, e77825.
14. Fedorova I., Vasileva A., Selkova P., Abramova M., Arseniev A., Pobegalov G., Kazalov M., Musharova O., Goryanin I., Artamonova D., Zyubko T., Shmakov S., Artamonova T., Khodorkovskii M., Severinov K. (2020) PpCas9 from *Pasteurella pneumotropica* – a compact type II-C Cas9 ortholog active in human cells. *Nucleic Acids Res.* **48**, 12297–12309.
15. Bland C., Ramsey T.L., Sabree F., Lowe M., Brown K., Kyrpides N.C., Hugenholtz P. (2007) CRISPR Recognition Tool (CRT): a tool for automatic detection of clustered regularly interspaced palindromic repeats. *BMC Bioinformatics*. **8**, 209.
16. Couvin D., Bernheim A., Toffano-Nioche C., Touchon M., Michalik J., Néron B., Rocha E.P.C., Vergnaud G., Gautheret D., Pourcel C. (2018) CRISPRCasFinder, an update of CRISPRFinder, includes a portable version, enhanced performance and integrates search for Cas proteins. *Nucleic Acids Res.* **46**, W246–W251.
17. Edgar R.C. (2007) PILER-CR: fast and accurate identification of CRISPR repeats. *BMC Bioinformatics*. **8**, 18.
18. Biswas A., Staals R.H.J., Morales S.E., Fineran P.C., Brown C.M. (2016) CRISPRDetect: a flexible algorithm to define CRISPR arrays. *BMC Genomics*. **17**, i356–i356.
19. Mitrofanov A., Alkhnbashi O.S., Shmakov S.A., Makarova K.S., Koonin E.V., Backofen R. (2021) CRISPRidentify: identification of CRISPR arrays using machine learning approach. *Nucleic Acids Res.* **49**, e20.
20. Mistry J., Finn R.D., Eddy S.R., Bateman A., Punta M. (2013) Challenges in homology search: HMMER3 and convergent evolution of coiled-coil regions. *Nucleic Acids Res.* **41**, e121.
21. Dooley S.K., Baken E.K., Moss W.N., Howe A., Young J.K. (2021) Identification and evolution of Cas9 tracrRNAs. *CRISPR J.* **4**, 438–447.
22. Mitrofanov A., Ziemann M., Alkhnbashi O.S., Hess W.R., Backofen R. (2021) CRISPRtracrRNA: robust approach for CRISPR tracrRNA detection. *Bioinformatics*. **38**, ii42–ii48.
23. Livny J., Waldor M.K. (2007) Identification of small RNAs in diverse bacterial species. *Curr. Opin. Microbiol.* **10**, 96–101.
24. Lorenz R., Bernhart S.H., zu Siederdissen C.H., Tafer H., Flamm C., Stadler P.F., Hofacker I.L. (2011) ViennaRNA Package 2.0. *Algorithms Mol. Biol.* **6**, 26.
25. Harrington L.B., Burstein D., Chen J.S., Paez-Espino D., Ma E., Witte I.P., Cofsky J.C., Kyrpides N.C., Banfield J.F., Doudna J.A. (2019) Programmed DNA destruction by miniature CRISPR-Cas14 enzymes. *Science*. **362**, 839–842.
26. Zhang Y., Heidrich N., Ampattu B.J., Gunderson C.W., Seifert H.S., Schoen C., Vogel J., Sontheimer E.J. (2013) Processing-independent CRISPR RNAs limit natural transformation in *Neisseria meningitidis*. *Mol. Cell*. **50**, 488–503.
27. Ran F.A., Cong L., Yan W.X., Scott D.A., Gootenberg J.S., Kriz A.J., Zetsche B., Shalem O., Wu X., Makarova K.S., Koonin E.V., Sharp P.A., Zhang F. (2015) *In vivo* genome editing using *Staphylococcus aureus* Cas9. *Nature*. **520**, 186–191.
28. Kim E., Koo T., Park S.W., Kim D., Kim K., Cho H.-Y., Song D.W., Lee K.J., Jung M.H., Kim S., Kim J.H., Kim J.H., Kim J.-S. (2017) *In vivo* genome editing with a small Cas9 orthologue derived from *Campylobacter jejuni*. *Nat. Commun.* **8**, 14500.
29. Laemmli U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*. **227**, 680–685.
30. Selkova P., Vasileva A., Pobegalov G., Musharova O., Arseniev A., Kazalov M., Zyubko T., Shcheglova N., Artamonova T., Khodorkovskii M., Severinov K., Fedorova I. (2020) Position of *Deltaproteobacteria* Cas12e nuclease cleavage sites depends on spacer length of guide RNA. *RNA Biol.* **17**, 1472–1479.
31. Sun P., Tropea J.E., Waugh D.S. (2011) Enhancing the solubility of recombinant proteins in *Escherichia coli* by

- using hexahistidine-tagged maltose-binding protein as a fusion partner. *Methods Mol. Biol.* **705**, 259–274.
32. Esvelt K.M., Mali P., Braff J.L., Moosburner M., Yaung S.J., Church G.M. (2013) Orthogonal Cas9 proteins for RNA-guided gene regulation and editing. *Nat. Methods.* **10**, 1116–1121.
 33. Zetsche B., Gootenberg J.S., Abudayyeh O.O., Slaymaker I.M., Makarova K.S., Essletzbichler P., Volz S.E., Joung J., van der Oost J., Regev A., Koonin E.V., Zhang F. (2015) Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell.* **163**, 759–771.
 34. Crooks G.E., Hon G., Chandonia J.-M., Brenner S.E. (2004) WebLogo: a sequence logo generator. *Genome Res.* **14**, 1188–1190.
 35. Hu Z., Wang S., Zhang C., Gao N., Li M., Wang D., Wang D., Liu D., Liu H., Ong S.-G., Wang H., Wang Y. (2020) A compact Cas9 ortholog from *Staphylococcus auricularis* (SauriCas9) expands the DNA targeting scope. *PLoS Biol.* **18**, e3000686.
 36. Leenay R.T., Maksimchuk K.R., Slotkowski R.A., Agrawal R.N., Gomaa A.A., Briner A.E., Barrangou R., Beisel C.L. (2016) Identifying and visualizing functional PAM diversity across CRISPR-Cas systems. *Mol. Cell.* **62**, 137–47.
 37. Cui Z., Tian R., Huang Z., Jin Z., Li L., Liu J., Huang Z., Xie H., Liu D., Mo H., Zhou R., Lang B., Meng B., Weng H., Hu Z. (2022) FrCas9 is a CRISPR/Cas9 system with high editing efficiency and fidelity. *Nat. Commun.* **13**, 1425.
 38. Fedorova I., Arseniev A., Selkova P., Pobegalov G., Goryanin I., Vasileva A., Musharova O., Abramova M., Kazalov M., Zyubko T., Artamonova T., Artamonova D., Shmakov S., Khodorkovskii M., Severinov K. (2020) DNA targeting by *Clostridium cellulolyticum* CRISPR-Cas9 type II-C system. *Nucleic Acids Res.* **48**, 2026–2034.

Type II CRISPR-Cas System Nucleases: a Pipeline for Prediction and *in vitro* Characterization

A. A. Vasileva^{1, 2, *}, S. A. Aliukas³, P. A. Selkova^{1, 2}, A. N. Arseniev^{1, 2}, V. E. Chernova²,
O. S. Musharova¹, E. I. Klimuk¹, M. A. Khodorkovskii², and K. V. Severinov^{1, 4, **}

¹*Institute of Molecular Genetics, Russian Academy of Sciences, Moscow, 123182 Russia*

²*Peter the Great St. Petersburg Polytechnic University, St. Petersburg, 195251 Russia*

³*LLC “Biotech campus”, Moscow, 117437 Russia*

⁴*Waksman Institute of Microbiology, Piscataway, NJ, 08854 USA*

*e-mail: daucussativus7@gmail.com

**e-mail: severik@waksman.rutgers.edu

The use of CRISPR-Cas bacterial adaptive immunity systems components for targeted DNA changing has opened broad prospects for programmable genome editing of higher organisms. The most widely used gene editors are based on the Cas9 effectors of the type II CRISPR-Cas systems. In complex with guide RNAs, Cas9 proteins are able to directionally introduce double-strand breaks into DNA regions complementary to guide RNA sequences. Despite the wide range of characterized Cas9s, the search for new Cas9 variants remains an actual task, since the available Cas9 editors have several limitations. This paper presents a workflow for the search and subsequent characterization of new Cas9 nucleases developed in our laboratory. Detailed protocols describing the bioinformatical search, cloning and isolation of recombinant Cas9 proteins, testing for nuclease activity *in vitro*, and determining the PAM sequence required for recognition of DNA targets, are presented. Potential difficulties that may arise, as well as ways to overcome them, are considered.

Keywords: CRISPR, Cas9, nuclease, genome editing, bioinformatics search