

ISSN 0026-8984

Том 58, Номер 4

Июль–Август 2024

# МОЛЕКУЛЯРНАЯ БИОЛОГИЯ



НАУКА

— 1727 —

# СОДЕРЖАНИЕ

Том 58, номер 4, 2024

## ОБЗОРЫ

- Современные знания о редактировании оснований и праймированном редактировании  
*О. А. Аверина, С. А. Кузнецова, О. А. Пермяков, П. В. Сергиев* 508
- Как сместить равновесие репарации разрывов ДНК в пользу гомологичной рекомбинации  
*О. А. Аверина, С. А. Кузнецова, О. А. Пермяков, П. В. Сергиев* 525

## ГЕНОМИКА. ТРАНСКРИПТОМИКА

- Паттерны аминокислотных замен в белках E6 и E7 вируса папилломы человека 16 типа:  
филогеография и эволюция  
*Е. Е. Зеленова, А. А. Карлсен, Д. В. Авдошина, К. К. Кюрегян, М. Г. Беликова, И. Д. Троценко* 549

## МОЛЕКУЛЯРНАЯ БИОЛОГИЯ КЛЕТКИ

- Способы повышения уровня нокина конструкции, кодирующей пептидный ингибитор слияния  
ВИЧ-1 MT-C34, в локус *CXCR4* в Т-клеточной линии СЕМ/R5  
*Д. С. Голубев, Д. С. Комков, М. В. Шепелев, Д. В. Мазуров, Н. А. Круглова* 575
- Повышение уровня нокина MT-C34-кодирующей конструкции в локус *CXCR4* с помощью  
модификации донорной ДНК сайтами-мишенями Cas9  
*М. В. Шепелев, Д. С. Комков, Д. С. Голубев, С. Е. Боровикова, Д. В. Мазуров, Н. А. Круглова* 590
- Чувствительность первичных клеточных линий глиобластомы человека  
к вакцинному штамму вируса паротита  
*Е. Ю. Николаева, Ю. Р. Желая, О. Ю. Сусова, А. А. Митрофанов, В. О. Варачев, Т. В. Наседкина,  
В. В. Зверев, О. А. Свитич, Ю. И. Амму* 601
- Белок CG9609 дрозофилы, содержащий домены цинковых пальцев, взаимодействует  
с деубиквитирующим (DUB) модулем комплекса SAGA и участвует в регуляции транскрипции  
*Ю. В. Николенко, М. М. Кушакова, Д. В. Копытова, Ю. А. Вдовина, Н. Е. Воробьева, А. Н. Краснов* 612
- Белки AEF1 и CG10543 дрозофилы, содержащие домены цинковых пальцев, колокализуются  
с комплексами SAGA, SWI/SNF и ORC на промоторах генов и участвуют в регуляции транскрипции  
*Ю. В. Николенко, М. М. Кушакова, Д. В. Копытова, Ю. А. Вдовина, Н. Е. Воробьева, А. Н. Краснов* 619
- Регуляция трансляции фактора eRF1 человека  
*А. В. Шувалов, А. А. Клишин, Н. С. Бизяев, Е. Ю. Шувалова, Е. З. Алкалаева* 627
- Метаболический профиль кишечной микробиоты и содержание трефоиловых факторов  
у взрослых с различными метаболическими фенотипами ожирения  
*И. М. Колесникова, Л. А. Ганенко, И. Ю. Васильев, Т. В. Григорьева,  
Н. И. Волкова, С. А. Румянцев, А. В. Шестопалов* 638
- Вирусоподобные частицы на основе E<sub>1v</sub> ВИЧ-1 с модулированным составом гликанов  
*Г. А. Каевицер, Е. И. Самохвалов, Д. В. Щебляков, А. Л. Гинцбург, А. Н. Взоров* 655

## СТРУКТУРНО-ФУНКЦИОНАЛЬНЫЙ АНАЛИЗ БИОПОЛИМЕРОВ И ИХ КОМПЛЕКСОВ

- Молекулярные блокаторы ионных каналов вирусов гриппа А и SARS-CoV-2  
*Ю. Н. Воробьев* 665

УДК 577.21

## СОВРЕМЕННЫЕ ЗНАНИЯ О РЕДАКТИРОВАНИИ ОСНОВАНИЙ И ПРАЙМИРОВАННОМ РЕДАКТИРОВАНИИ

© 2024 г. О. А. Аверина<sup>a, b, c</sup>, С. А. Кузнецова<sup>a, \*</sup>, О. А. Пермяков<sup>a, c</sup>, П. В. Сергиев<sup>a, b, c</sup><sup>a</sup>Институт функциональной геномики, Московский государственный университет им. М. В. Ломоносова, Москва, 119991 Россия<sup>b</sup>Научно-исследовательский институт физико-химической биологии им. А. Н. Белозерского, Московский государственный университет им. М. В. Ломоносова, Москва, 119991 Россия<sup>c</sup>Химический факультет, Московский государственный университет им. М. В. Ломоносова, Москва, 119991 Россия

\*e-mail: svetlana@belozersky.msu.ru

Поступила в редакцию 27.11.2023 г.

После доработки 05.01.2024 г.

Принята к публикации 09.01.2024 г.

Современные технологии генной инженерии, такие как редакторы оснований и праймированное редактирование, зарекомендовали себя как эффективные и надежные инструменты редактирования геномов, которые не требуют внесения двухцепочечных разрывов в молекулу ДНК и наличия донорских матриц. Появившись относительно недавно, эти технологии получили быстрое признание благодаря своей точности, простоте и возможностям мультиплексирования. В представленном обзоре суммированы новые данные об этих технологиях, их архитектуре и методах конструирования редакторов, специфичности, эффективности, универсальности. Обсуждаются преимущества, недостатки и перспективы использования редакторов в фундаментальных и прикладных исследованиях. Сведения, приведенные в обзоре, могут быть востребованы для планирования исследований по геномному редактированию и анализа их результатов в ходе решения различных задач фундаментальной биологии, биотехнологии, медицины и сельского хозяйства.

**Ключевые слова:** CRISPR/Cas-технология, редактирование оснований, нуклеозиддезаминазы, праймированное редактирование, обратная транскриптаза, системы репарации

DOI: 10.31857/S0026898424040013, EDN: INFPUZ

### ВВЕДЕНИЕ

Геномное редактирование CRISPR/Cas – одна из наиболее востребованных технологий направленного изменения генома и эпигенома млекопитающих. Благодаря своей простоте, доступности,

точности, высокой эффективности и огромному потенциалу, эта технология активно используется для решения различных задач фундаментальной биологии, биотехнологии, медицины и сельского хозяйства. С помощью геномного редактирования

Сокращения: ABE – редактор оснований A→G; BE – редактор оснований (Base Editor); BER – эксцизионная репарация оснований (Base Excision Repair); Cas – CRISPR-ассоциированная нуклеаза; CBE – редактор оснований C→T; CGBE – редактор оснований C→G; CRISPR – короткие палиндромные повторы, разделенные уникальными последовательностями-спейсерами; crPHK – направляющая РНК (crispR RNA); dCas – мутантная форма нуклеазы Cas, лишённая ферментативной активности в обоих каталитических доменах (dead или deactivated); DdCBE – редактор для замены пары C•G на T•A в дцДНК; EXO1 – экзонуклеаза I; FEN1 – флп-специфичная эндонуклеаза I; флп – отделенный от основной цепи ДНК “болтающийся/свисающий” олигонуклеотид; HDR – гомологичная рекомбинация (Homologous Recombination); HNH – каталитический домен нуклеазы Cas9, вносящий одноцепочечный разрыв в целевой участок ДНК и связывающийся со спейсером crPHK; MLH1dn – белок-репрессор MMR-репарации;

M–MLV RT – обратная транскриптаза вируса лейкоза мышей Молони; MMR – репарация неспаренных нуклеотидов (Mismatch Repair); nCas – мутантная форма нуклеазы Cas, способная разрезать только одну из цепей ДНК (nickase); NHEJ – негомологичное соединение концов ДНК (Nonhomologous DNA End Joining); NLS – последовательность ядерной локализации; PAM – смежный протоспейсерный мотив (Protospacer Adjacent Motif); PE – праймированное редактирование (Prime Editing); PFS – последовательность нуклеотидов, фланкирующая протоспейсер (Protospacer Flanking Sequence); RT – обратная транскриптаза; RuvC – каталитический домен нуклеазы Cas9, вносящий одноцепочечный разрыв в целевой участок ДНК (протоспейсер); transPHK – транс-активирующая РНК (trans-activating RNA); UGI – ингибитор урацил-ДНК-гликозилазы; UNG – урацил-ДНК-гликозилаза; rPHK – гидовая РНК (sgRNA – single-guide RNA); ДЦР – двухцепочечные разрывы.

можно изучать функции генов млекопитающих, создавать модели генетических заболеваний и разрабатывать подходы к их коррекции, улучшать значимые для сельского хозяйства виды. Вместе с тем, технология CRISPR/Cas не лишена таких недостатков, как относительно высокая вероятность образования нецелевых мутаций, зависимость эффективности редактирования целевых участков от геномного контекста и структуры хроматина, низкая эффективность гомологичной рекомбинации (HDR) и доставки компонентов CRISPR/Cas в клетки. Одним из ключевых этапов CRISPR/Cas является целенаправленная индукция сайт-специфических двухцепочечных разрывов (ДЦР) в ДНК, что приводит к серьезным повреждениям клетки. ДЦР могут быть причиной потери сегментов хромосом и хромосомных перестроек, вызывать гибель клеток или возникновение злокачественных новообразований. Для повышения эффективности и безопасности редактирования разрабатываются новые усовершенствованные версии CRISPR/Cas. Так, в последние годы появились альтернативные версии редактирования геномов, которые не требуют введения ДЦР. К ним относятся новые стратегии редактирования оснований и праймированного редактирования, получившие быстрое признание благодаря своей точности, простоте и возможностям мультиплексирования. Эти подходы основаны на применении химерных белков, состоящих из модифицированного белка Cas, не способного вносить ДЦР в ДНК, и ковалентно связанных с ним ферментов – нуклеозиддезаминазы (редактирование оснований) или обратной транскриптазы вируса лейкоза мышей Молони (M-MLV RT) (праймированное редактирование). В этих технологиях не используются синтетические экзогенные донорные молекулы ДНК и дополнительные экзогенные/эндогенные белковые факторы. В настоящем обзоре обобщены современные знания о редактировании оснований и праймированном редактировании: описаны общие принципы конструирования редакторов, суммированы сведения об их архитектуре, специфичности, эффективности, универсальности. Обсуждаются преимущества, недостатки и перспективы использования этих редакторов в фундаментальных и прикладных исследованиях.

## РЕДАКТИРОВАНИЕ ОСНОВАНИЙ

Редактирование оснований – это способ прямого изменения последовательности нуклеотидов целевого участка генома. Редактирование проводится путем дезаминирования гетероциклического основания в ДНК без введения ДЦР, HDR, в отсутствие донорской матрицы и дополнительных экзогенных/эндогенных белковых факторов. Основными составляющими единицами редакторов оснований (BE) являются нуклеозиддезаминаза,

каталитически неактивная Cas-нуклеаза и РНК, направляющая комплекс к целевому локусу для преобразования пар оснований. Точечные мутации могут быть внесены как в ДНК, так и в транскрипты РНК. В последнем случае происходит прямое редактирование патогенных транскриптов без изменения генетической информации. Поскольку редактирование РНК носит временный характер, в геномной ДНК не возникнут побочные продукты или нежелательные мутации, которые могут быть переданы дочерним клеткам во время митоза. Такой подход перспективен для лечения заболеваний, не имеющих генетического происхождения, и имитации генетических вариантов, которые обеспечивают преимущества для организма [1].

## ТРАНСВЕРСИЯ ЦИТИДИНА В ТИМИДИН

Первыми редакторами оснований стали BE, обеспечивающие направленное введение замен С→Т, поскольку природные цитидиндезаминазы, входящие в суперсемейство APOBEC/AID, способны на уровне РНК или ДНК дезаминировать С до U, имеющего такие же комплементарные свойства, как и Т. Первыми охарактеризованными членами этого суперсемейства стали белки APOBEC1 и AID. Белок APOBEC1 является каталитической субъединицей ферментного комплекса, редактирующего мРНК аполипопротеина В. Фермент AID играет ключевую роль в соматическом гипермутационном рекомбинации генов иммуноглобулинов в В-клетках при переключении классов антител. Также в суперсемейство входит группа цитидиндезаминаз APOBEC3 *Homo sapiens* (hA3), которые инактивируют геномы патогенных ретровирусов, таких как вирус иммунодефицита человека, с помощью случайных мутаций [2–4]. Цитидиндезаминазы суперсемейства APOBEC/AID сильно различаются по активности дезаминирования и предпочтениям в отношении нуклеотидного окружения мишени. Например, APOBEC1 *Rattus norvegicus* (rAPO1) выбирает ближайших соседей в следующем порядке: TC ≥ CC ≥ AC > GC [5]. Все семь дезаминаз группы hA3, несмотря на их структурное сходство, сильно различаются активностью дезаминирования и предпочтениями в отношении ДНК-субстрата [6–11]. На базе цитидиндезаминаз суперсемейства APOBEC/AID почти одновременно были разработаны три разных BE, которые параллельно прошли схожие пути совершенствования. Первые BE были сконструированы на основе природной rAPO1 и названы редакторами цитидиновых оснований (CBE) [5]. На основе ферментов, дезаминирующих ДНК, разработаны редакторы Target-AID и TAM. В Target-AID задействован ортолог AID – цитидиндезаминаза *Petromyzon marinus* CDA1 (PmCDA1), объединенная с N-концом каталитической субъединицы dCas или nCas (d/nCas), ингибитором урацил-ДНК-гликозилазы (UGI)

и гРНК [12]. Редактор TAM сконструирован с использованием цитидиндезаминазы AID, гибризированной с d/nCas и гРНК, а также при участии UGI [13]. Все три BE обладают не только высокой эффективностью, но и такими формами индивидуальной специфичности, как температурная чувствительность, размер окна активности, предпочтение нуклеотидной последовательности протоспейсера. Эти различия предполагают, что разные редакторы должны дополнять друг друга и обеспечивать возможность выбора [14].

Рассмотрим механизм функционирования первого редактора – CBE [5]. CBE первого поколения получен путем гибридизации цитидиндезаминазы гAPO1 с N-концом неактивной нуклеазы dCas9, содержащей мутации, препятствующие образованию ДЦР в ДНК [5]. CBE1 дезаминирует С до U, который затем заменяется на Т (рис. 1). Однако в технологии CBE1 есть нюанс, снижающий

эффективность редактирования – это работа урацил-ДНК-гликозилазы (UNG). В нормальных условиях эксцизионная репарация оснований (BER) является первичным ответом клетки на несоответствие пары G • U, который инициируется вырезанием U с помощью UNG для предотвращения накопления мутаций в результате спонтанного окислительного дезаминирования С [15]. В процессе редактирования оснований промежуточная пара G • U может репарироваться клеточной системой BER с образованием исходной пары G • С. Поэтому, чтобы защитить урацил, входящий в некомплементарную пару G • U, от удаления с помощью UNG, CBE1 модифицировали, присоединив 83-аминокислотный UGI к С-концу комплекса гAPO1-dCas9 [5] (рис. 1). UGI включены также в Target-AID и TAM [12, 13]. В этой новой версии CBE2 урацил не удаляется путем BER с участием UNG, что существенно повышает эффективность

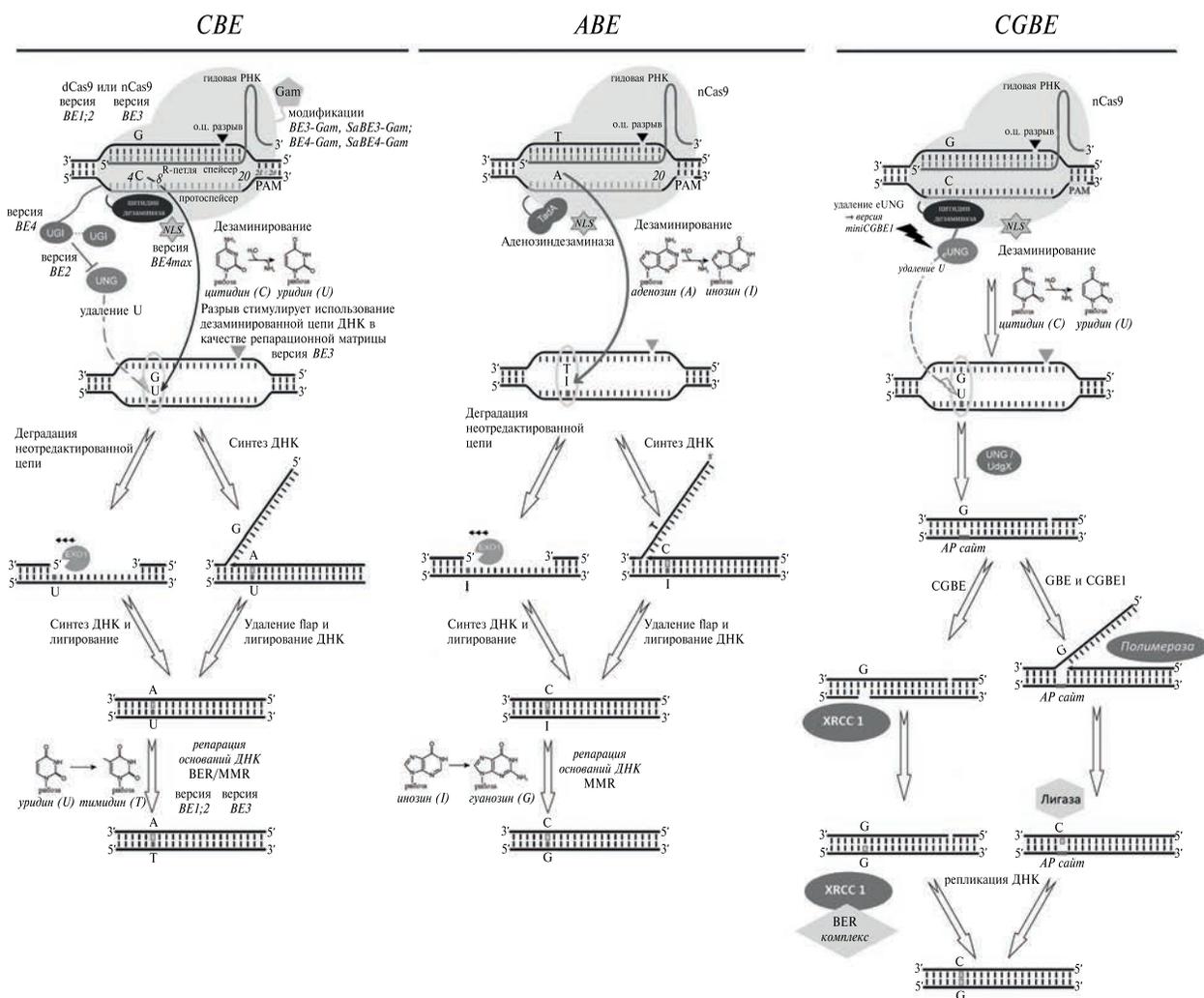


Рис. 1. Схематичное изображение функционирования CBE, осуществляющих трансверсию цитидина в тимидин (C→T); ABE, осуществляющих трансверсию аденозина в гуанозин (A→G); CGBE, осуществляющих трансверсию цитидина в гуанозин (C→G).

редактирования С→Т. Для дальнейшего повышения эффективности редактирования разработана версия СВЕ3, в которой вместо неактивной нуклеазы dCas9 используется nCas9 (альтернативное название D10A) [5]. В этом случае разрыв происходит в паре G•U недезаминированной (G-содержащей) цепи ДНК. Разрыв стимулирует использование дезаминированной (U-содержащей) цепи в качестве матрицы для репарации с переходом в промежуточное состояние с парой A•U, которая в итоге преобразуется в A•T [16]. Эта модификация, а также присоединение второго домена UGI (модификация СВЕ4) и белка Gam бактериофага Mu, который, как известно, защищает концы ДЦР от деградации, дополнительно повысило эффективность редактирования С (рис. 1) [17]. Дальнейшее усовершенствование ВЕ осуществлено в версии СВЕ4max путем модификации последовательности ядерной локализации (NLS) и оптимизации кодонов [18].

Доступный для дезаминазы участок оцДНК-мишени в пределах протоспейсера R-петли, называемый “окном редактирования/активности”, остается критичным звеном целевого редактирования оснований. Классическое окно составляет примерно 5 из 20 п.н. всего протоспейсера. Это означает, что для дезаминирования доступен только С, локализованный между позициями 4 и 8, при этом РАМ соответствует положениям с 21 по 23 (рис. 1) [5]. Расширение окна активности считается одним из направлений повышения эффективности редактирования. При этом, если в окне активности или в непосредственном окружении присутствуют несколько нуклеотидов-мишеней, то они могут подвергнуться сопутствующей нежелательной трансверсии. В связи с этим необходимо учитывать оба направления корректировки размера окна активности и выбирать оптимальный в зависимости от последовательности протоспейсера в каждом эксперименте. На окно активности может влиять ряд факторов, например, рядом с целевой последовательностью должен находиться подходящий РАМ. Так, ВЕ, в которых задействованы d/nCas9 *Streptococcus pyogenes* (d/nSpCas9), ограничены редактированием геномных локусов, содержащих NGG РАМ. Для расширения возможностей редактирования разработаны ВЕ, содержащие ортологи d/nCas9, распознающие РАМ, отличные от NGG, окно активности которых смещается или расширяется до целевых оснований. В противном случае эти основания становятся недоступными из-за отсутствия оптимального РАМ. Варианты ВЕ, созданные на основе разных ортологов SpCas9, подробно рассмотрены [14, 19], сопоставлены их активности и размеры окон редактирования. Важно отметить, что у редактора оснований *Staphylococcus aureus* окно редактирования шире, чем у *S. pyogenes*, а именно позиции 3–12, что позволяет редактировать основания, расположенные

ближе к РАМ [20–27]. Интересный вариант Cas9 с циклической перестановкой (Cas9-CP) изменяет локализацию домена нуклеозиддезаминазы по отношению к R-петле и таким образом может обеспечить более эффективное преобразование оснований в более широком окне редактирования [28]. ВЕ с участием Cas12a (Cpf1) распознает TTTV РАМ для трансверсии С в расширенном до 6 п.н. окне редактирования [29]. Напротив, ВЕ, включающий миниатюрный вариант Cas12f (CasMINI), имеет очень узкое окно, всего из 2 п.н. [30]. А использование оптимизированной версии nCas12a *Acidaminococcus sp.* демонстрирует улучшенную трансверсию С→Т [31].

Другим подходом к уменьшению зависимости ВЕ от структуры РАМ стало создание нуклеазы SpRY – структурно модифицированного варианта SpCas9. С помощью ВЕ, несущих в своей структуре SpRY, стало возможным воздействовать на гены, ранее недоступные для редактирования [32]. Однако такая “релаксация” РАМ снижает специфичность и способствует нецелевому редактированию оснований ДНК. В таких случаях можно предусмотреть возникновение “опечаток” редактирования и принять ряд мер по нивелированию проблемы. Рассмотрены два типа нецелевых эффектов ВЕ: независимые и зависимые от гРНК [33]. Зависимое от гРНК нецелевое редактирование происходит, когда комплекс ВЕ направляется гРНК к последовательностям ДНК, схожим с протоспейсером. Хотя зависимое от гРНК нецелевое редактирование оснований происходит относительно редко, идентификация этих “опечаток” крайне важна. Для оценки потенциальных нецелевых эффектов ВЕ используют полногеномное секвенирование и методы GUIDE-seq [34], Digenome-seq [35] и EndoV-seq [36]. Нивелировать эту проблему можно с помощью тщательного дизайна гРНК с использованием высокоточных вариантов Cas и РНП-комплексов для доставки ВЕ в клетки. Возникновение случайных нецелевых мутаций также не зависит от гРНК, особенно при высоком уровне экспрессии комплекса редактирования, и наблюдается существенно реже при использовании дезаминаз AmAPOBEC1, PpAPOBEC1, RrA3F, SsAPOBEC3B [37] и Ye1 [38]. Сильное влияние на уровень нецелевого редактирования оказывает специфика гибридизации дезаминазы с белком Cas [39]. Так, при соединении с Cas через N-конец опечатки редактирования возникают чаще, чем при встраивании редактирующих ферментов в центральную зону (середины) nCas9 (такой вариант конструкции называют CE “Cas Embedding”) [40]. Логично, что ВЕ с более узкими окнами активности демонстрируют более низкий уровень нецелевого редактирования всего генома, потому что для редактирования как в целевых, так и в нецелевых сайтах доступно меньше нуклеотидов [38], хотя такой подход часто приводит к компромиссу между

эффективностью и точностью редактирования. Как и в случае гРНК-зависимых опечаток ВЕ, снизить независимое от гРНК нецелевое редактирование можно, используя для доставки в клетку РНП или мРНК, что ограничивает экспрессию ВЕ по времени, уменьшая нецелевое редактирование [37].

На расположение и ширину окна активности, точность и специфичность редактирования могут влиять различные варианты дезаминазы [11]. Например, классическая дезаминаза гАРО1 в составе СВЕ3 имеет окно активности из 5 п.н. (позиции 4–8 протоспейсера) [5], но изменения аминокислотных остатков гАРО1, участвующих в катализе и связывании ДНК, могут привести к сужению окна редактирования. Так, внесение мутаций W90Y или W90F в гАРО1 может снижать гидрофобность ее активного центра и, соответственно, внутреннее сродство к ДНК. Такие замены уменьшают неспецифическое дезаминирование и сокращают окно редактирования до 3 п.н., сохраняя при этом высокий уровень целевого дезаминирования. Оба варианта мутаций (R126E и R132E) в гАРО1 влияют на доступность оцДНК, сужают окно активности до 3 п.н., но снижают эффективность ВЕ. Показано, что комбинация мутаций W90Y и R126E в гАРО1 (эта мутантная гАРО1 названа YE1, или YE), W90Y и R132E (YE2), R126E и R132E (EE) приводит к сужению окна активности до 2 п.н. при изменении эффективности редактирования в ряду дезаминаз в следующем порядке: гАРО1 ≈ YE1 > YE2 ≈ EE. При этом, если пойти дальше и ввести в гАРО1 тройную мутацию, объединив W90Y или W90F с R132E и R126E (мутант YEE), окно активности не только сократится до 2 п.н., но область редактирования сконцентрируется на С-6 протоспейсера, если в положении 5 также будет С. В связи с этим упоминается, что у ВЕ, в состав которого входит YEE, окно сужается до 1 п.н. В этой ситуации точность ВЕ сильно повышается, но эффективность при этом резко падает [20]. Анализ ВЕ по соотношению “размер окна активности/эффективность редактирования” показал, что именно мутация R132E в дезаминазе ответственна за снижение эффективности, поэтому исследовали и альтернативные варианты мутаций в гАРО1. Так, гАРО1 с мутацией Y120F, расположенной вблизи активного центра гАРО1 и имеющей решающее значение для контакта с С-мишенью, показала высокую эффективность, что подтолкнуло к созданию тройного мутанта гАРО1, аналогичного YEE. В итоге эффективность ВЕ, в состав которого вошла гАРО1 с мутациями W90Y, Y120F и R126E, оказалась сопоставимой с ВЕ4 при сужении окна активности до 3 п.н. [41].

Напротив, использование цитидиндезаминазы hA3A расширяет, по разным данным, окно редактирования от 9 (позиции 3–11 протоспейсера) [42] до 12 п.н. (позиции 2–13 протоспейсера)

и улучшает превращение С→Т в метилированных областях генома [9]. Включение оцДНК-связывающего домена белка RAD51 между nCas9 и hA3A расширяет окно редактирования до 13 п.н. (позиции 3–15 (~17) протоспейсера) [42]. Одновременное включение нескольких доменов дезаминазы в ВЕ увеличивает доступность нуклеотидов ДНК-мишени в R-петле, расширяя тем самым окно редактирования. Представителем этой технологии является система “ВЕ-PLUS”, которая включает комплекс nCas9, SunTag, scFv-АРОВЕС-UGI-GB1, гРНК, где SunTag – полипептидный мотив, содержащий 10 копий эпитопа GCN4. С SunTag специфически взаимодействует одноцепочечный вариативный фрагмент scFv. Таким образом, действующий ВЕ собирается при помощи указанного взаимодействия и содержит гРНК, nCas9 и 10 молекул гибридного белка scFv-АРОВЕС-UGI-GB1. При этом для предотвращения агрегации этой белковой конструкции используется домен GB1 G-белка *Streptococcus* [43]. В свою очередь, аминокислотная замена N57G, привела к повышению точности редактирования hA3A. Цитидиндезаминаза eA3A вошла в состав ВЕ третьего поколения (ВЕ3). eA3A-ВЕ3 показывает специфичность высокоточного даже по сравнению с описанными ВЕ суженного окна редактирования С в мотиве TCR (R = А или G) [10]. Вместе с тем проблемой остается высокоточное редактирование С в других мотивах, в частности, в целевых сайтах, содержащих несколько остатков С. Решением этой проблемы стала модернизация цитидиндезаминазы *H. sapiens* АРОВЕС3G (hA3G) и создание на базе ВЕ4max редактора eA3G-ВЕ с контекстной специфичностью СС. Цитидиндезаминаза hA3G содержит С-концевой каталитический домен и второй псевдокаталитический домен на N-конце, который сохраняет ту же пространственную укладку, но не обладает каталитической активностью. В улучшенную версию hA3G встроены С-концевой каталитический домен (CTD, C-terminal domain), что приводит к образованию hA3G-CTD (eA3G), который преимущественно дезаминирует С в соответствии с иерархией ССС>ССС>СС. Эффективность и точность eA3G-ВЕ показана на клетках человека и эмбрионах кролика, что указывает на возможность использования этого редактора в терапии генетических заболеваний и создании животных моделей с измененным геномом для фундаментальных и доклинических исследований [44].

Сузить окно редактирования можно также с использованием цитидиндезаминазы PmCDA1. Target-AID позволяет сузить окно активности до 4 п.н. (позиции 16–19 протоспейсера) [12]. Кроме того, можно сузить окно редактирования и путем укорачивания дезаминазы PmCDA1 с С-конца, что приводит к возможности выборочного редактирования С в положениях 15 или 16 протоспейсера [45]. Укорачивание PmCDA1 не только

с С-конца, но и с N-конца способствует снижению ее сродства к ДНК, сводя тем самым к минимуму неспецифическое дезаминирование С. Одновременное усечение N- и С-концов приводит к уменьшению количества участков с “открытыми” гидрофобными остатками, что делает поверхность дезаминазы более “гладкой”, а внесение мутаций W122E и W139Q в эти гидрофобные остатки значительно повышает активность модернизированной цитидиндезаминазы (tCDA1EQ). В tCDA1EQ был удален UGI, в присутствии которого увеличивалась скорость мутаций. Поскольку tCDA1EQ имеет низкое сродство к ДНК и менее стабильна, чем исходная PmCDA1, архитектура ее присоединения к nCas9 может влиять на редактирующую активность. Для минимизации данной проблемы tCDA1EQ можно не только соединить с N-концом nCas9, но и встроить в середину этого белка. Действительно, tCDA1EQ, присоединенная к N-концу nCas9, показала нестабильную эффективность редактирования, тогда как версия, встроенная внутрь nCas9, имела стабильно высокую эффективность, сравнимую с Target-AID. Структура “tCDA1EQ-N-конец nCas9” была названа AID-2S, где 2S означает “Small & Specific” (маленький и специфичный), а встроенная в nCas9 – AID-3S, где 3S это “Small & Specific & Superior” (маленький, специфичный и улучшенный). Показано, что AID-2S имеет более узкую область редактирования, тогда как у AID-3S окно активности сдвинуто в сторону последовательности PAM по сравнению с исходным Target-AID. Эффективность AID-2S и AID-3S находилась на уровне Target-AID, YE1 и YE2, при этом нецелые эффекты были значительно снижены [46].

Использование улучшенной цитидиндезаминазы AID в составе BE четвертого поколения, eAID-BE4max, обеспечивает окно редактирования из 11 п.н. и редактирование С в контексте GC [47], а использование нескольких eгРНК (последовательности сгРНК, связанные линкерной петлей с каркасной последовательностью tracrРНК), позволяет еще больше расширить окно редактирования [13].

Окно редактирования может быть изменено также путем изменения ориентации доменов в составе BE и/или свойств связывающего их линкера [11]:

- В целом, дезаминазы, присоединенные к N-концу d/nCas9, имеют более широкие окна редактирования, чем присоединенные к С-концу, поскольку эта ориентация обеспечивает, вероятно, более легкий доступ к оцДНК внутри R-петли [48, 49]. В структурах, где дезаминаза встроена в середину nCas, размер окна редактирования будет варьировать в зависимости от положения дезаминазы в белке [28, 39]. Предполагается, что уменьшение расстояния между дезаминазой и белком Cas9 приводит к более точному позиционированию дезаминазы на последовательности-мишени [45, 48].

- Модификация линкера между дезаминазой и белком Cas путем укорачивания линкера и/или изменения его свойств, например увеличения жесткости, приводит к сужению окна активности при сохранении высокой эффективности редактирования [11, 48]. Однако укороченный линкер не влияет существенно на окно редактирования [10, 20].

С целью внесения случайных мутаций в заранее выбранные участки генома разработана стратегия, использующая CRISPR-направляемые ДНК-полимеразы. ДНК-полимеразы способны создавать все 12 замен, а также делеции. Так, ДНК-полимеразы, транслирующие одноцепочечный разрыв ДНК, способны инициировать синтез на одноцепочечном разрыве дцДНК, замещая расположенные по ходу репликации нуклеотиды, т.е. в сторону 3'-конца, а их flap-эндонуклеазный домен впоследствии разрушает смещенные нуклеотиды, оставляя лигируемый одноцепочечный разрыв. ДНК-полимеразы, транслирующие одноцепочечный разрыв ДНК, участвуют в конструировании новых BE. Показано, что введение в nCas9 трех мутаций (K848A, K1003A и R1060A), способствующих диссоциации комплекса nCas9 с ДНК после разрыва целевого локуса, увеличивает частоту целевых мутаций. Усовершенствованная версия nCas9 названа enCas9. В итоге гибридизация enCas9 с N-концом ДНК-полимеразы I *Escherichia coli* (PolII), несущей мутации D424A, I709N и A759R (PolI3M), привела к созданию enCas9-PolI3M, названной EvolvR. По сравнению с использованием направленного мутагена на основе dCas9-AIDx, где окно редактирования заключено в пределах 5 п.н., EvolvR способна непрерывно диверсифицировать все нуклеотиды в пределах настраиваемой длины окна, до 350 п.н., в определяемых локусах. Следовательно, nCas9-PolI3M генерирует более широкий спектр замещающих мутаций, чем dCas9-AIDx, что применимо для крупномасштабных генетических скрининговых исследований [50, 51].

Эффективность направленного мутагенеза дополнительно повышается при использовании CRISPR-X. В CRISPR-X неактивная нуклеаза dCas9 и цитидиндезаминаза AID взаимодействуют посредством MS2-шпилек и MS2-связывающих белков. Две шпильки MS2, входящие в состав eгРНК, рекрутируют два MS2-связывающих белка, которые соединены с AID, лишенной С-конца ( $\Delta$ AID). Использование MS2- $\Delta$ AID ограничивает локализацию AID в ядре и улучшает эффективность мутагенеза по сравнению с полноразмерным AID. Более того, гиперактивный вариант AID с усиленной активностью соматической гипермутации увеличивал скорость мутагенеза до  $\sim 1/1000$  оснований. Экспериментально показана возможность использования CRISPR-X в направленной эволюции GFP в усиленный

GFP с более высокой интенсивностью излучения (eGFP), а также для идентификации вариантов генов, устойчивых к бортезомибу [52]. Кроме того, CRISPR-X использовали для получения моноклональных антител человека с повышенной аффинностью к антигену в клетках НЕК293 [53].

В качестве альтернативного подхода к увеличению точности редактирования можно рассматривать систему нековалентного привлечения цитидиндезаминазы к конкретной геномной С-мишени. Этот подход предполагает отход от классического ковалентного взаимодействия деаминазы с Cas9, которое неразрывно связывает события деаминации С как на целевых участках генома, так и на близлежащих нежелательных мишенях. Напротив, нековалентное связывание должно способствовать разделению событий редактирования на целевом сайте и рядом с ним, тем самым повышая вероятность точных мутаций с заменой одного основания. Ключевыми игроками этой системы становятся белки, которые связывают деаминазу, не блокируя активный сайт от взаимодействия с С-мишенью. Эти белки, связанные с комплексом Cas9п/гРНК, используются для привлечения коэкспрессируемого фермента АРОВЕС для редактирования. По ассоциации с физическим процессом магнетизма эта технология была названа “магнитным редактором” (MagnEdit) [54].

### ТРАНСВЕРСИЯ АДЕНОЗИНА В ГУАНОЗИН

Альтернативным подходом в разработке ВЕ стало использование деаминаз А. Редактор основания А (АВЕ) включает nCas, аденозиндезаминазу, гРНК и функционирует подобно СВЕ3. Комплекс АВЕ связывается с ДНК-мишенью, после чего аденозиндезаминаза катализирует превращение А→I, создавая из Т•А пару Т•I, которая путем репарации неспаренных оснований (MMR) последовательно превращается в С•I и затем в С•G. Отсутствие природной ДНК-аденозиндезаминазы для нацеливания и перехода А→G привело к созданию из гРНК-аденозиндезаминазы *E. coli* (TadA<sup>wt</sup>), мутантного варианта TadA\*, способного эффективно деаминировать около 53% остатков А в ДНК (рис. 1). Создано несколько поколений АВЕ, которые различаются вводимыми в гомодимер TadA мутациями [55]. Среди них выделяют наиболее эффективные АВЕ6.3, АВЕ7.8, АВЕ7.9 с окнами редактирования 3 п.н. (позиции 8–10 протоспейсера) и АВЕ7.10 с окном редактирования из 4 п.н. (позиции 4–7 протоспейсера). АВЕ7.10 состоит из гетеродимера TadA\*-TadA<sup>wt</sup>, она обеспечивает высокоэффективное редактирование А в линиях клеток мышей [56] и человека [55]. Активность АВЕ7.10 была дополнительно повышена путем модификации NLS и оптимизации кодонов до версии АВЕmax [18]. Стратегии изменения окна

активности АВЕ развивались таким же путем, как и в СВЕ. Окно редактирования в зависимости от поставленной экспериментальной задачи можно сузить, вводя, например, мутацию F148A в домен TadA [57], или расширить, используя Cas9-CP [28], а также SpRY [32]. Подход с заменой классического SpCas9 ортологами позволил осуществлять трансверсию А→С в геномных сайтах, содержащих *не-NGG* PAM [22, 58]. Например, для создания АВЕ, действующих на сайты, содержащие NNNRRT и NGA PAM, соответственно, nSpCas9 был заменен на nSaKKH или nSpCas9-VQR в АВЕ7.10 и АВЕmax. При этом Sa-АВЕmax и SaKKH-АВЕmax имеют большое окно редактирования (позиции 4–14 протоспейсера), хотя эффективность редактирования невысока [28, 59]. Использование xCas9 (вариант SpCas9), созданного для работы с широким спектром последовательностей PAM [26], в АВЕ7.10 и АВЕmax позволило распознавать широкий спектр последовательностей *не-NGG* PAM, включая NR, NG, GAA и GAT [28]. Однако в отличие от усовершенствованных версий СВЕ, АВЕ, содержащие ортологи SpCas9, имели более низкую эффективность редактирования оснований, чем канонический АВЕ, потому что белок TadA был изначально усовершенствован и представлен в виде гибрида с SpCas9. В следующей версии АВЕ8 гетеродимер TadA\*-TadA<sup>wt</sup> не содержит TadA<sup>wt</sup>, а TadA\* дополнен восемью новыми мутациями, что позволило улучшить кинетику деаминации. Удаление TadA<sup>wt</sup> не влияет на активность редактирования АВЕ8 и указывает на то, что новые аденозиндезаминазы TadA-8e и TadA-8s могут эффективно работать как мономеры [34, 60]. Так, эффективность АВЕ8 с TadA-8e была выше, чем у АВЕ7.10 при сочетании деаминазы не только с SpCas9, но и с SpnCas9-NG, SaCas9, SaCas9-KKH, LbCas12a, enAsCas12a, а окно активности было шире (позиции 3–10 протоспейсера) [60]. АВЕ8 с TadA-8s также показал высокую эффективность редактирования в сочетании с SaCas9 и SpCas9-NG [34]. Менее эффективной по сравнению с TadA-8, но не менее интересной является новая версия аденозиндезаминазы, в которую добавили мутации V82G и A56G, а также удалили TadA<sup>wt</sup> из гетеродимера. Смысл этой двойной мутации заключается в том, что замена V82G (версия miniАВЕ) нивелирует неспецифическое сродство к РНК-субстратам [61], представляющее проблему как для АВЕ, так и СВЕ, которые могут индуцировать независимое от гРНК редактирование оснований РНК в масштабах всего транскриптома [62]. Однако мутация V82G также снижает сродство TadA к ДНК-субстратам, что в результате отрицательно влияет на эффективность редактирования. Поэтому для увеличения сродства к ДНК-мишени в miniАВЕ добавили вторую мутацию A56G. Полученная TadA-A56G&V82G эффективно редактировала ДНК-мишени и характеризовалась сниженной нецелевой активностью

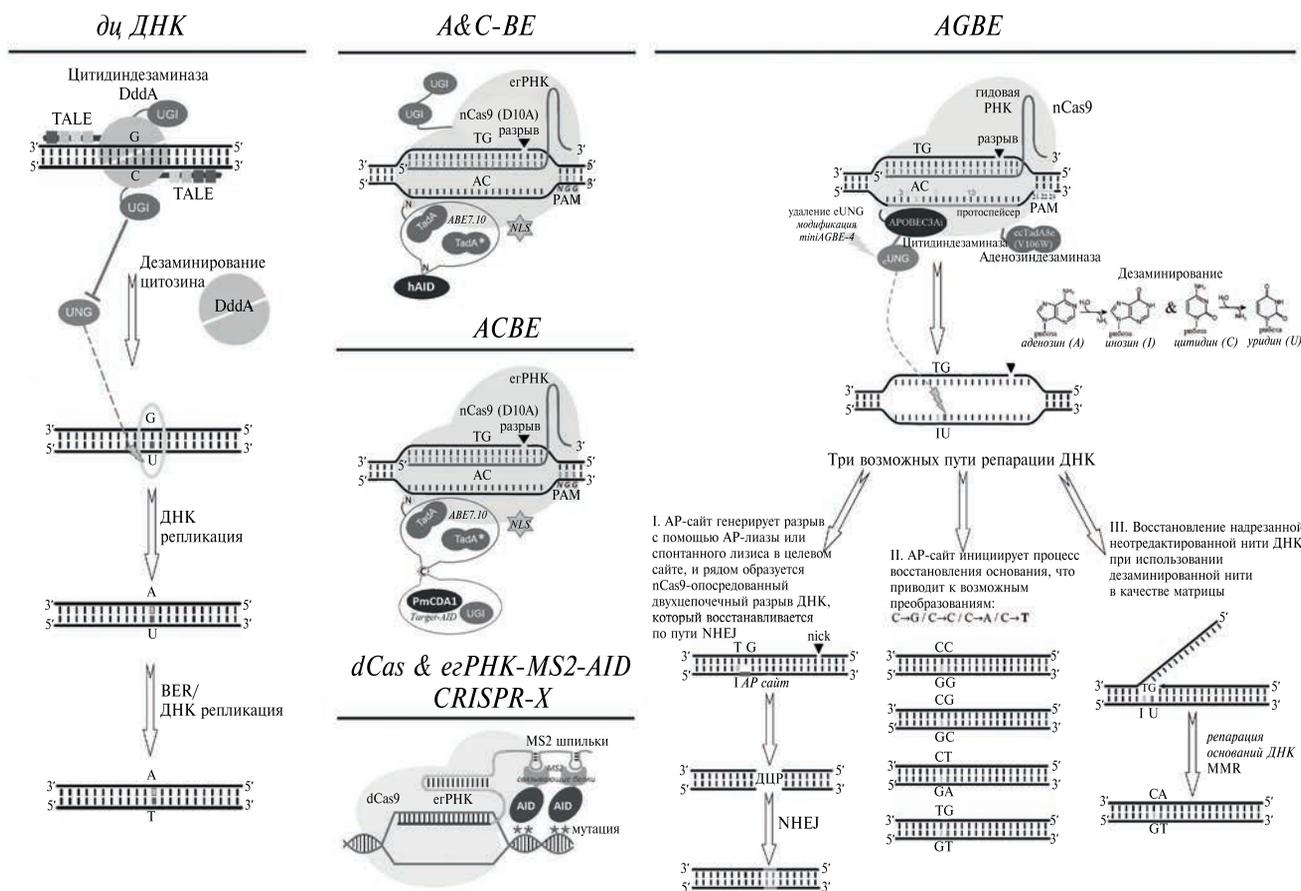
дезаминирования РНК [63]. Дополнительно с модификациями АВЕ можно ознакомиться в работе Jeong Y. K. и соавт. [14].

**ТРАНСВЕРСИЯ ЦИТИДИНА В ГУАНОЗИН**

Следующей ступенью в развитии технологии ВЕ стала трансверсия С→G. Сначала был получен GBE, состоящий из nCas9, цитидиндезаминазы и UNG. После дезаминирования и образования остатка дезоксиуридина UNG вырезает урацил, что приводит к появлению AP-сайта. В то же самое время nCas9 вносит одноцепочечный разрыв, создавая предпосылки для удлинения 3'-конца разрыва с использованием в качестве матрицы цепи ДНК, содержащей AP-сайт. В трансверсии С→G в клетках млекопитающих участвует гAPO1. Система гAPOBES-nCas9-UNG показала высокую специфичность редактирования в шестой позиции протоспейсера, эффективность которой составила 5.3–53.0% [64]. Параллельно на основе платформы ВЕ4max был разработан ВЕ С→G [65]. На первом этапе из ВЕ4max удалили две UGI и получили ВЕ4maxΔUGI, что привело к увеличению

количества трансверсий С→G; при этом также наблюдалось наиболее эффективное редактирование в шестом положении протоспейсера [65]. В качестве цитидиндезаминазы была выбрана гAPO1 с мутацией R33A и сниженной нецелевой активностью редактирования РНК и ДНК [62]. Добавление к N-концу ВЕ4maxΔUGI ортолога UNG из *E. coli* (eUNG) привело к формированию ВЕ4max(R33A)ΔUGI, названной CGBE1 (рис. 1). Однако ВЕ CGBE1, несмотря на высокую эффективность целевого редактирования С→G, оказался довольно громоздким, поэтому пришлось его значительно уменьшить, удалив домен eUNG, что привело к умеренному снижению эффективности и получению нового редактора miniCGBE1 [65].

Кроме UNG-опосредованной трансверсии С→G [64, 65], где количество переходов С•G→G•С прямо зависит от генерации AP-сайтов, разработан новый редактор CGBE, основанный на манипулировании репарацией ДНК после создания AP-сайта. Так, при создании AP-сайта в Cas9-индуцированной R-петле клеточный фермент UNG замещается эндонуклеазой APE1, после чего белок XRCC1 (X-ray repair cross-complementing protein 1),



**Рис. 2.** Схематичное изображение редактирования оснований в дцДНК; синхронной трансверсии А→G и С→T в одном целевом сайте (A&C-BE, ACBE); одновременного введения четырех типов трансверсий: С→G, С→T, С→A, А→G, а также InDel на одной и той же цепи ДНК в сайтах-мишенях (AGBE).

участвующий в эксцизионной репарации ДНК, рекрутирует такие компоненты BER, как ДНК-полимераза  $\beta$  (Pol  $\beta$ ) и ДНК-лигазы, для заполнения бреши G. Последующая репарация ДНК преобразует пару G•G в G•C (рис. 1) [66]. Каждый редактор описывается как высокоэффективный и обладающий широким потенциалом применения в научных и терапевтических целях. Для предсказания эффективности той или иной модификации BE, частоты редактирования C•G→G•C и определения его оптимальных стратегий разработаны модели машинного обучения CGBE-Hive [67].

Недавно стало возможным редактирование оснований в контексте дцДНК, ранее недоступное из-за специфичной к оцДНК активности цитидин- и аденозиндезаминаз. Показано, что DddA – токсин, секретируемый бактериями *Burkholderia cenocepacia* для индукции репликативного стресса ДНК [68] и нестабильности генома у других бактерий, проявляет цитидиндезаминазную активность в отношении дцДНК [69]. Для замены C•G → T•A в дцДНК как клеточных линий, так и мышей, сконструирован редактор DdCBE: токсичный домен DddA был разделен на две нетоксичные части, которые соединили с UGI и ДНК-связывающими активаторами транскрипции TALE, что позволило восстановить функциональный домен цитидиндезаминазы (рис. 2) [70, 71]. Этот BE может работать не только с яДНК, но и с мтДНК, что предоставляет большие возможности для изучения митохондриальных заболеваний, поскольку более 90% патогенных мутаций мтДНК являются точечными [72]. Следует отметить, что известные BE, управляемые РНК, не подходят для редактирования мтДНК, поскольку гРНК не могут быть доставлены в митохондрии [73]. Для расширения возможностей подобных систем разработан BE TALEd, превращающий A в G в мтДНК. Отличие редактора TALEd от DdCBE заключается в использовании нетоксичного полноразмерного варианта DddA с мутацией E1347A в активном сайте дезоксиаденозиндезаминазы, полученной из Tada и слитой с TALE [74].

В принципе, используя описанные режимы BE, можно осуществить многоэтапное редактирование. Например, C в паре C•G может быть первоначально превращена в U с помощью CBE (U•G), а U•G, в свою очередь, в T•A с помощью механизмов репликации и репарации клеточной ДНК. В качестве альтернативы C•G можно превратить в G•C с помощью CGBE с последующим опосредованным CBE переходом G•C → A•T. Затем с использованием этой стратегии целевой C можно превратить в T (CBE), G (CGBE) или A (CGBE + CBE). Точно так же A можно преобразовать в G (ABE), C (ABE + CGBE) или T (ABE + CGBE + CBE). Тем не менее, многоэтапное редактирование трудно провести с высокой эффективностью и специфичностью

из-за ряда ограничений, таких как окно высокой дезаминазной активности, расположенное в определенном положении относительно PAM. Кроме того, на результат и эффективность редактирования может влиять контекст последовательности целевого локуса, а наличие нескольких редактируемых азотистых оснований в окне активности, а иногда и за его пределами, может быть причиной нежелательных мутаций [64, 75–77]. Несмотря на широкий спектр BE с различными характеристиками и свойствами, в том числе обладающих высокой эффективностью и специфичностью, совместная трансфекция редакторов разных оснований дает низкий коэффициент совместной трансверсии A→G и C→T. Скорее всего, это связано и с конкуренцией разных BE за целевой сайт [78].

### СИНХРОННЫЕ ТРАНСВЕРСИИ НЕСКОЛЬКИХ ОСНОВАНИЙ

В 2020 году опубликован ряд BE для синхронного введения замен A→G и C→T в одном и том же целевом сайте. Путем гибридизации цитидиндезаминазы и адениндезаминазы вместе с nCas9 и гРНК были сконструированы двойные BE как для растений, так и для млекопитающих:

1. STEME (saturated targeted endogenous mutagenesis editors) в протопластах риса одновременно редактирует C и A в одном и том же сайте-мишени с эффективностью 15.10% [79].

2. SPACE создан на платформах miniABEmax-V82G (A→G) и Target-AID (C→T) с добавлением двух UGI [80].

3. Target-ACEmax объединяет цитидиндезаминазу PmCDA1 и аденозиндезаминазу Tada для эффективного синхронного редактирования A→G и C→T в клетках HEK293T [81].

4. A&C-BEmax сконструирован из цитидиндезаминазы hAID, соединенной с N-концом ABE7.10. Комплекс ABE7.10-N-hAIDmax содержит NLS, что увеличивает активность двухкомпонентного редактирования в большей степени за счет повышения активности CBE, чем ABE. Версия A&C-BEmax дополнена двумя копиями UGI (рис. 2). Коэффициент одновременной трансверсии A и C при совместной трансфекции CBE и ABE7.10 с шестью егРНК был в большинстве случаев существенно ниже, чем при использовании A&C-BEmax. A&C-BEmax даже рассматривают в качестве возможного терапевтического агента [78].

5. ACBE объединяет возможности редакторов Target-AID и ABE7.10 (рис. 2). В ACBE цитидиндезаминаза PmCDA1 и UGI в составе Target-AID соединены через C-конец с ABE7.10, который присоединен к nCas9 через N-конец, при этом NLS находится в N- и C-концевых участках. Особенностью ACBE считают одновременное введение точечных

мутаций A→G и C→T в соматические клетки млекопитающих для получения с трансгенных животных и коррекции генетических заболеваний человека [82].

6. AGBE сконструирован на платформах CGBE и ABE для одновременного введения четырех типов трансверсий C→G, C→T, C→A, A→G, а также InDel в сайтах-мишенях на одной и той же цепи ДНК (рис. 2). В максимально эффективной и компактной версии miniAGBE-4 eUNG удалена, CGBE содержит мутантную цитидиндезаминазу APOBEC3Ai, а ABE – высокоэффективную мономерную аденозиндезаминазу eTadA8e(V106W) [83].

Интересным направлением, расширяющим диапазон применения BE, стало введение преждевременного стоп-кодона в середину гена-мишени с целью дестабилизации его экспрессии. Два независимых подхода (CRISPR-STOP и iSTOP) показывают, что BE3 может превращать кодоны CAA, CAG, CGA и/или TGG в стоп-кодоны TAA, TAG и TGA в целевых генах [84, 85]. Вносить мутации в стартовый кодон ATG может также редактор ABEmax. Эта стратегия, названная i-Silence, включает трансверсию ATG в GTG/ACG, подавляя тем самым экспрессию целевого гена. В отличие от CRISPR-STOP и iSTOP дестабилизация генов с помощью i-Silence не приводит к образованию укороченных вариантов белка, поскольку блокируется начало трансляции [86]. Еще одна система, CRISPR-pass, превращает преждевременный стоп-кодон в кодоны глутамин (CAA или CAG) или аргинин (CGA), обеспечивая продолжение трансляции [87]. BE способны также вызывать программируемый пропуск экзонов путем мутации оснований ДНК-мишени в акцепторных сайтах сплайсинга. Система CRISPR-SKIP в акцепторных сайтах сплайсинга осуществляет трансформацию G, которым заканчивается большинство интронов, в A через редактирование C в комплементарной цепи целевого сайта. В результате соответствующие экзоны не включаются в зрелые транскрипты, тогда как другие экзоны экспрессируются нормально [88].

### ПРАЙМИРОВАННОЕ РЕДАКТИРОВАНИЕ

На основе технологии редактирования оснований и CRISPR/Cas9 разработана молекулярная платформа праймированного редактирования генома (PE), применение которой позволяет выполнять все 12 возможных преобразований, вставок и делеций от основания к основанию. В отличие от HDR, PE не зависит от ДЦР ДНК или синтетических экзогенных донорных молекул ДНК для замены целевых геномных последовательностей. Более того, PE расширяет возможности редактирования генома, позволяя, теоретически, исправить до 89% мутаций более чем в 75 000 генетических заболеваний человека, о которых сообщается

в ClinVar [89, 90]. Метод PE более универсален, чем HDR, так как способен вносить точечные мутации не только в пределах 10 п.н. от места разрыва, но и на больших (> 30 п.н.) расстояниях [91]. Хотя эффективность редактирования PE и BE еще не сравнима, но этот подход непрерывно модернизируется и при определенных условиях его эффективность может превышать 50% [91–93]. Технология PE может применяться в фундаментальных и прикладных исследованиях не только на клеточных культурах [89] и мелких лабораторных животных [94], но и на моделях органоидных культур [95], *Danio rerio* [96], *Drosophila melanogaster* [97], кроликах [98], а также растениях [93, 99, 100].

Редакторы PE сконструированы путем слияния M–MLV RT с C-концом nCas9. В этой технологии используется специфическая 3'-удлиненная egРНК, содержащая PBS и RT-матрицу (pegРНК). Общий принцип работы PE заключается в том, что система, направляемая pegРНК, связывается с ДНК-мишенью, nCas9 создает разрыв в цепи ДНК, содержащей PAM. Созданный таким образом одноцепочечный участок ДНК со свободным 3'-концом связывается с PBS на pegРНК. Далее RT связывается с 3'-концом экспонированной цепи ДНК-мишени и выполняет обратную транскрипцию участка pegРНК, после чего нередатируемый выступающий 5'-флар на цепи, содержащей PAM, разрушается клеточными эндонуклеазами. Наконец, происходит лигирование и репарация ДНК с образованием нужной последовательности (рис. 3) [89]. Более высокая специфичность редакторов PE по сравнению с классическим методом CRISPR/Cas9 может быть обусловлена тремя последовательными гибридизациями: между ДНК-мишенью и спейсером в молекуле pegРНК (гРНК, содержащей PBS и RT-матрицу), между ДНК-мишенью и PBS в молекуле pegРНК и между ДНК-мишенью и редактируемым участком флар-ДНК [92].

Разработаны три основные версии PE. Редактор первого поколения (PE1) получен путем слияния C-конца nCas9 (H840A) с M–MLV RT, при участии pegРНК его эффективность составила 0.7–5.5% при использовании точечных трансверсий [89]. Во второй версии (PE2) в M–MLV RT для увеличения активности при более высоких температурах были внесены три мутации (D200N, L603W и T330P), а также две дополнительные мутации (T306K и W313F) для усиления связывания с комплексом pegРНК. В итоге вариант PE2 оказался в 1.6–5.1 раза эффективнее PE1 [92]. PE3 разработан для устранения неопределенности в редактировании 5'-флар, который спаривается с немодифицированной цепью ДНК. С этой целью применяются дополнительную гРНК, которая с помощью nCas9 вносит разрыв в нередатируемую цепь ДНК, используя эндогенный путь восстановления несоответствия для сохранения информации об

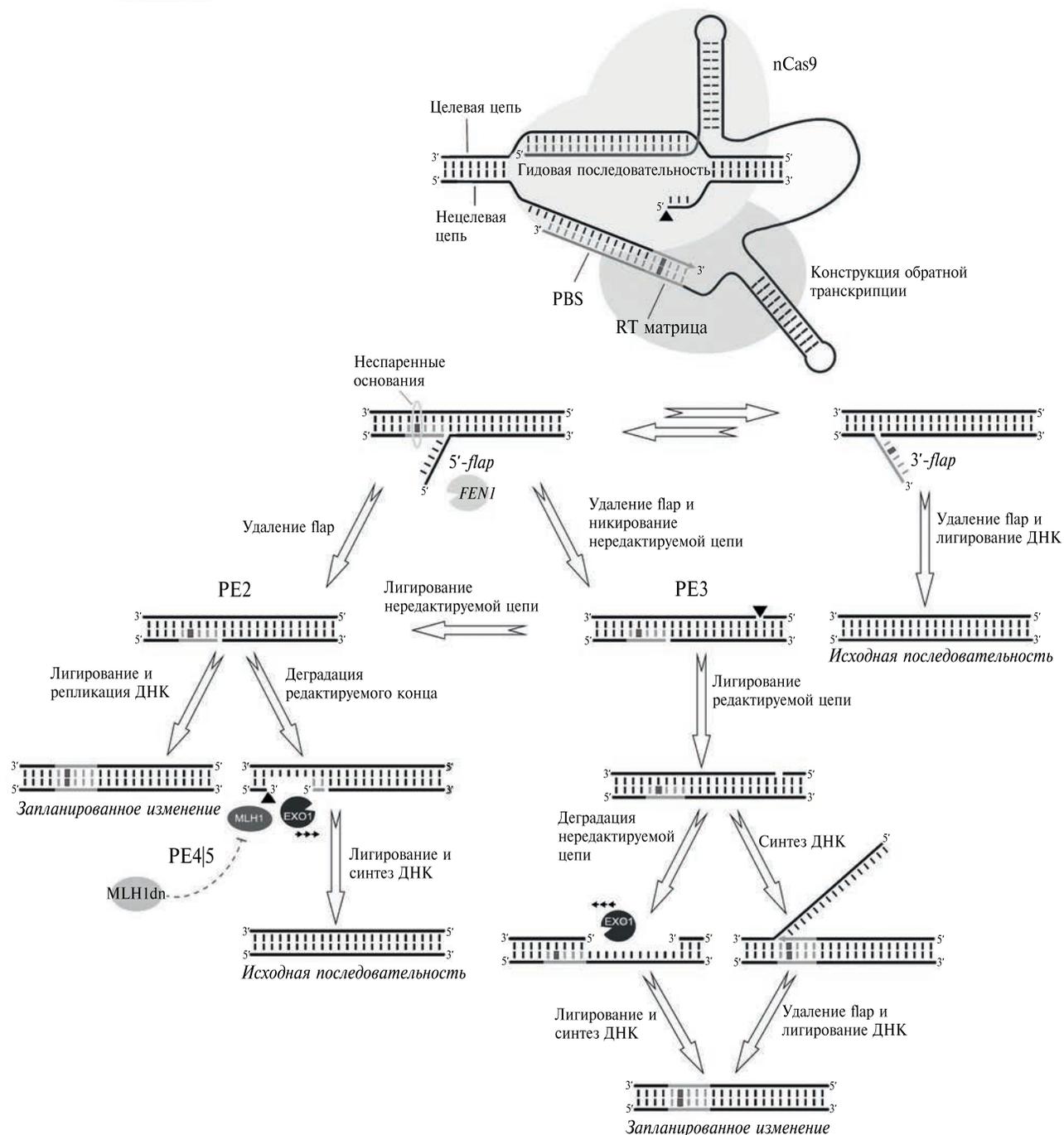


Рис. 3. Схематичное изображение функционирования основных модификаций PE.

отредактированной цепи. Благодаря этому подходу можно с эффективностью 33% получить все 12 комбинаций оснований. Количество нецелевых эффектов и нежелательных мутаций при этом становится ниже, чем при использовании CRISPR/Cas9 (рис. 3) [16]. Развитие PE следующих поколений основано на подавлении MMR-пути [101]. Предложены более эффективные системы, где редактирование осуществляется на фоне временной

экспрессии белка-репрессора MMR (MLH1dn). На основе PE2 и PE3 сконструированы редакторы с подавленной активностью MMR, названные PE4 и PE5 соответственно. Кроме того, редакторы второго–пятого поколений были модернизированы до версии PE<sub>max</sub> за счет использования кодон-оптимизированной RT человека, линкера из 34 аминокислотных остатков, двухкомпонентной последовательности NLS SV40, которая увеличивает

соотношение ядерной и цитоплазматической концентраций белка, дополнительной С-концевой последовательности (с-Мус NLS), включающей остатки Pro320-Asp328, что позволяет осуществлять транспорт в ядро, а также за счет включения в SpCas9 мутаций R221K и N394K, улучшающих ее активность. В клетках с MMR эта новая система превзошла по эффективности редактирования классический PE2 в 2.8 раза, а в клетках с дефицитом MMR в 1.2 раза [101, 102].

В оценке эффективности технологии PE существуют некоторые расхождения [101, 103–106]. В настоящее время основными ограничениями PE считаются его крупный размер, зависимость от целевых сайтов и типов клеток, более сложный дизайн гРНК. В стремлении усовершенствовать систему разработаны различные модификации PE, в которых в той или иной степени были устранены многие из ограничений. Например, эффективность PE2 повысили, добавив на N-конец пептиды IGFp1 и NFATC2IPp1, связанные с репарацией ДНК (модификация IN-PE2) [107], или ДНК-связывающий домен Rad51 (модификация huPE2) [108].

Расширение диапазона редактирования PE зависит в основном от белка Cas, который распознает последовательность PAM. PE первых трех поколений сконструированы на основе SpCas9, распознающей NGG-последовательности PAM, которые встречаются в среднем 1 раз в каждых 16 случайно выбранных геномных локусах, что сильно ограничивает диапазон мишеней для редактирования [89]. Использование таких вариантов SpCas9, как NG, SpG, SpRY, VQR, VRQR и VRER, после введения мутации H840A в PE второго поколения привело к следующим результатам [109]:

- модификации PE2-NG, PE2-SpG и PE2-SpRY позволяют редактировать сайты рядом с PAM NGN;
- оба варианта PE2-VQR и PE2-VRQR активны в пяти из семи сайтов NGA PAM;
- PE2-VRER предпочитает NGCG другим последовательностям PAM;
- PE2-SpG имеет самую высокую активность в сайтах NGA/C/T;
- PE2-SpRY функционирует независимо от вариаций последовательности PAM (активен в 43 из 58 сайтов), хотя и со сниженной активностью.

Модификация regРНК и ее компонентов тоже способствует повышению эффективности PE. Так показано, что увеличение длины PBS позволяет повысить активность regРНК с 3.29% для 9-нуклеотидного PBS до 16.6% для 13-нуклеотидного [109]. Добавление в RT-матрицу синонимичных мутаций, т.е. мутаций, при которых изменение нуклеотидной последовательности гена не приводит к изменению аминокислотной последовательности белка,

формирует версию spregРНК, которая увеличивает эффективность PE в разы [110]. Классическая regРНК циклизуется во время отжига, поскольку PBS на ее 3'-конце комплементарен части спейсера на 5'-конце, а это препятствует работе редакторов PE. Предотвратить этот феномен можно, защитив 3'-конец шпилечной структурой, полученной путем присоединения сайта узнавания рибонуклеазы Csy4 из 20 п.н. Гибридизация Csy4-T2A с nCas9 и комплексом RT на базе PE третьего поколения привела к созданию более эффективной версии ePE [111]. Также 3'-конец regРНК нуждается в защите от экзонуклеазной деградации, которая может привести к потере PBS и, как следствие, снизить эффективность редактирования PE. Для решения этой проблемы решили добавить к 3'-концу псевдоузел, представляющий собой модифицированный аптамер evorgeQ1. Этот аптамер, состоящий из 42 п.н., относится к классу природных РНК минимального размера с четко определенной третичной структурой. regРНК с псевдоузлом на 3'-конце (eregРНК) дополнительно усовершенствовали, вставив 8-нуклеотидный линкер между 3'-концом PBS и evorgeQ1, чтобы уменьшить влияние аптамера на функцию regРНК [106]. Защитить regРНК от 3'-деградации можно и присоединив к 3'-концу мотив, устойчивый к вирусной экзорибонуклеазе (версия хгРНК) [112]. Сходным образом работает и технология добавления на 3'-конец G-квадруплекса hTR (теломеразная РНК человека) [113]. Другой подход к увеличению стабильности regРНК с целью повышения эффективности PE состоит во включении С•G или замены других пар оснований на С•G в шпильке гРНК. Этот вариант гРНК назван aregРНК [110]. Расширить спектр применений PE можно с использованием сразу двух regРНК. В сочетании с сайт-специфичной рекомбиназой PE с двумя regРНК (twinPE) способна реализовать целевую интеграцию ДНК-плазмид размером более 5000 п.н. и инверсию последовательностей из 40 т.п.н. в клетках человека [114].

Таким образом, следует отметить, что технологии PE и BE произвели революцию в области редактирования генома. Вместе с тем, нерешенными остаются такие проблемы, как повышение эффективности и специфичности редактирования, оптимизация методов доставки и устранение нецелевых эффектов. Ранее мы показали, что несмотря на значительный прогресс PE2, оптимальной технологией целевого редактирования генома мышей остается egРНК-направленное расщепление ДНК-мишени Cas9 с последующей гомологичной рекомбинацией с использованием оцДНК-матрицы [104]. Однако попытки модернизации технологий PE и BE не стоят на месте и ведутся параллельно. Предложена новая стратегия DAP “Drive-And-Process” мультиплексного редактирования оснований до 31 локуса и мультиплексного праймированного редактирования до трех локусов.

DAR существенно снизила нецелевое редактирование и была адаптирована для доставки с помощью векторов на основе аденоассоциированного вируса и лентивируса для одновременного редактирования нескольких локусов [115]. Технологии PE и VE широко обсуждаются в сфере генной терапии и, несмотря на обеспокоенность по поводу возможной генотоксичности [116], результаты доклинических испытаний формируют мнение, что однажды PE и VE будут использоваться для лечения генетических заболеваний [117].

Работа поддержана грантом Российского научного фонда (№ 21-64-00006).

Настоящая статья не содержит каких-либо исследований с использованием животных в качестве объектов.

Авторы заявляют об отсутствии конфликта интересов.

### СПИСОК ЛИТЕРАТУРЫ

- Reardon S. (2020) Step aside CRISPR, RNA editing is taking off. *Nature*. **578**, 24–27.
- Лада А.Г., Фрам-Крик К., Козьмин С.Г., Майоров В.И., Карпова Т.С., Рогозин И.Б., Павлов Ю.И. (2011) Мутаторные эффекты и мутационная специфичность редактирующих дезаминаз. *Биохимия*. **76**(1), 157–175.
- Kohli R.M., Maul R.W., Guminski A.F., McClure R.L., Gajula K.S., Saribasak H., McMahon M.A., Siliciano R.F., Gearhart P.J., Stivers J.T. (2010) Local sequence targeting in the AID/APOBEC family differentially impacts retroviral restriction and antibody diversification. *J. Biol. Chem.* **285**, 40956–40964.
- Coticello S.G. (2008) The AID/APOBEC family of nucleic acid mutators. *Genome Biol.* **9**, 229.
- Komor A.C., Kim Y.B., Packer M.S., Zuris J.A., Liu D.R. (2016) Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature*. **533**, 420–424.
- Hou S., Lee J.M., Myint W., Matsuo H., Kurt Yilmaz N., Schiffer C.A. (2021) Structural basis of substrate specificity in human cytidine deaminase family APOBEC3s. *J. Biol. Chem.* **297**, 100909.
- Ziegler S.J., Liu C., Landau M., Buzovetsky O., Desimie B.A., Zhao Q., Sasaki T., Burdick R.C., Pathak V.K., Anderson K.S., Xiong Y. (2018) Insights into DNA substrate selection by APOBEC3G from structural, biochemical, and functional studies. *PLoS One*. **13**, e0195048.
- Lee S., Ding N., Sun Y., Yuan T., Li J., Yuan Q., Liu L., Yang J., Wang Q., Kolomeisky A.B., Hilton I.B., Zuo E., Gao X. (2020) Single C-to-T substitution using engineered APOBEC3G-nCas9 base editors with minimum genome- and transcriptome-wide off-target effects. *Sci. Adv.* **6**, eab1773.
- Wang X., Li J., Wang Y., Yang B., Wei J., Wu J., Wang R., Huang X., Chen J., Yang L. (2018) Efficient base editing in methylated regions with a human APOBEC3A-Cas9 fusion. *Nat. Biotechnol.* **36**, 946–949.
- Gehrke J.M., Cervantes O., Clement M.K., Wu Y., Zeng J., Bauer D.E., Pinello L., Joung J.K. (2018) An APOBEC3A-Cas9 base editor with minimized bystander and off-target activities. *Nat. Biotechnol.* **36**, 977–982.
- Tan J., Forner J., Karcher D., Bock R. (2022) DNA base editing in nuclear and organellar genomes. *Trends Genet.* **38**, 1147–1169.
- Nishida K., Arazoe T., Yachie N., Banno S., Kakimoto M., Tabata M., Mochizuki M., Miyabe A., Araki M., Hara K.Y., Shimatani Z., Kondo A. (2016) Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems. *Science*. **353**, aaf8729.
- Ma Y., Zhang J., Yin W., Zhang Z., Song Y., Chang X. (2016) Targeted AID-mediated mutagenesis (TAM) enables efficient genomic diversification in mammalian cells. *Nat. Methods*. **13**, 1029–1035.
- Jeong Y.K., Song B., Bae S. (2020) Current status and challenges of DNA base editing tools. *Mol. Ther.* **28**, 1938–1952.
- Pearl L.H. (2000) Structure and function in the uracil-DNA glycosylase superfamily. *Mut. Res.* **460**, 165–181.
- Nambiar T.S., Baudrier L., Billon P., Ciccio A. (2022) CRISPR-based genome editing through the lens of DNA repair. *Mol. Cell*. **82**, 348–388.
- Komor A.C., Zhao K.T., Packer M.S., Gaudelli N.M., Waterbury A.L., Koblan L.W., Kim Y.B., Badran A.H., Liu D.R. (2017) Improved base excision repair inhibition and bacteriophage Mu Gam protein yields C: G-to-T: A base editors with higher efficiency and product purity. *Sci. Adv.* **3**, eaao4774.
- Koblan L.W., Doman J.L., Wilson C., Levy J.M., Tay T., Newby G.A., Maianti J.P., Raguram A., Liu D.R. (2018) Improving cytidine and adenine base editors by expression optimization and ancestral reconstruction. *Nat. Biotechnol.* **36**, 843–846.
- Yu S.Y., Birkenshaw A., Thomson T., Carlaw T., Zhang L.H., Ross C.J.D. (2022) Increasing the targeting scope of CRISPR base editing system beyond NGG. *CRISPR J.* **5**, 187–202.
- Kim Y.B., Komor A.C., Levy J.M., Packer M.S., Zhao K.T., Liu D.R. (2017) Increasing the genome-targeting scope and precision of base editing with engineered Cas9-cytidine deaminase fusions. *Nat. Biotechnol.* **35**, 371–376.
- Nishimasu H., Shi X., Ishiguro S., Gao L., Hirano S., Okazaki S., Noda T., Abudayyeh O.O., Gootenberg J.S., Mori H., Oura S., Holmes B., Tanaka M., Seki M., Hirano H., Aburatani H., Ishitani R., Ikawa M., Yachie N., Zhang F., Nureki O. (2018) Engineered CRISPR-Cas9 nuclease with expanded targeting space. *Science*. **361**, 1259–1262.
- Miller S.M., Wang T., Randolph P.B., Arbab M., Shen M.W., Huang T.P., Matuszek Z., Newby G.A., Rees H.A., Liu D.R. (2020) Continuous evolution of Sp-Cas9 variants compatible with non-G PAMs. *Nat. Biotechnol.* **38**, 471–481.
- Kleinstiver B.P., Prew M.S., Tsai S.Q., Topkar V.V., Nguyen N.T., Zheng Z., Gonzales A.P., Li Z., Peterson R.T., Yeh J.R., Aryee M.J., Joung J.K. (2015) Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature*. **523**, 481–485.
- 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.

25. Kleinstiver B.P., Prew M.S., Tsai S.Q., Nguyen N.T., Topkar V.V., Zheng Z., Joung J.K. (2015) Broadening the targeting range of *Staphylococcus aureus* CRISPR-Cas9 by modifying PAM recognition. *Nat. Biotechnol.* **33**, 1293–1298.
26. Hu J.H., Miller S.M., Geurts M.H., Tang W., Chen L., Sun N., Zeina C.M., Gao X., Rees H.A., Lin Z., Liu D.R. (2018) Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. *Nature.* **556**, 57–63.
27. Liu Z., Shan H., Chen S., Chen M., Song Y., Lai L., Li Z. (2019) Efficient base editing with expanded targeting scope using an engineered Spy-mac Cas9 variant. *Cell Discov.* **5**, 58.
28. Huang T.P., Zhao K.T., Miller S.M., Gaudelli N.M., Oakes B.L., Fellmann C., Savage D.F., Liu D.R. (2019) Circularly permuted and PAM-modified Cas9 variants broaden the targeting scope of base editors. *Nat. Biotechnol.* **37**, 626–631.
29. Li X., Wang Y., Liu Y., Yang B., Wang X., Wei J., Lu Z., Zhang Y., Wu J., Huang X., Yang L., Chen J. (2018) Base editing with a Cpf1-cytidine deaminase fusion. *Nat. Biotechnol.* **36**, 324–327.
30. Xu X., Chemparathy A., Zeng L., Kempton H.R., Shang S., Nakamura M., Qi L.S. (2021) Engineered miniature CRISPR-Cas system for mammalian genome regulation and editing. *Mol. Cell.* **81**, 4333–4345.e4.
31. Kleinstiver B.P., Sousa A.A., Walton R.T., Tak Y.E., Hsu J.Y., Clement K., Welch M.M., Horng J.E., Malagon-Lopez J., Scarfò I., Maus M.V., Pinello L., Aryee M.J., Joung J.K. (2019) Engineered CRISPR-Cas12a variants with increased activities and improved targeting ranges for gene, epigenetic and base editing. *Nat. Biotechnol.* **37**, 276–282.
32. Walton R.T., Christie K.A., Whittaker M.N., Kleinstiver B.P. (2020) Unconstrained genome targeting with near-PAMless engineered CRISPR-Cas9 variants. *Science.* **368**, 290–296.
33. Rees H.A., Wilson C., Doman J.L., Liu D.R. (2019) Analysis and minimization of cellular RNA editing by DNA adenine base editors. *Sci. Adv.* **5**, eaax5717.
34. Gaudelli N.M., Lam D.K., Rees H.A., Solá-Esteves N.M., Barrera L.A., Born D.A., Edwards A., Gehrke J.M., Lee S.J., Liguori A.J., Murray R., Packer M.S., Rinaldi C., Slaymaker I.M., Yen J., Young L.E., Ciaramella G. (2020) Directed evolution of adenine base editors with increased activity and therapeutic application. *Nat. Biotechnol.* **38**, 892–900.
35. Kim D., Lim K., Kim S.T., Yoon S.H., Kim K., Ryu S.M., Kim J.S. (2017) Genome-wide target specificities of CRISPR RNA-guided programmable deaminases. *Nat. Biotechnol.* **35**, 475–480.
36. Liang P., Xie X., Zhi S., Sun H., Zhang X., Chen Y., Chen Y., Xiong Y., Ma W., Liu D., Huang J., Songyang, Z. (2019) Genome-wide profiling of adenine base editor specificity by EndoV-seq. *Nat. Commun.* **10**, 67.
37. Yu Y., Leete T.C., Born D.A., Young L., Barrera L.A., Lee S.J., Rees H.A., Ciaramella G., Gaudelli N.M. (2020) Cytosine base editors with minimized unguided DNA and RNA off-target events and high on-target activity. *Nat. Commun.* **11**, 2052.
38. Doman J.L., Raguram A., Newby G.A., Liu D.R. (2020) Evaluation and minimization of Cas9-independent off-target DNA editing by cytosine base editors. *Nat. Biotechnol.* **38**, 620–628.
39. Chu S.H., Packer M., Rees H., Lam D., Yu Y., Marshall J., Cheng L.I., Lam D., Olins J., Ran F.A., Liguori A., Gantzer B., Decker J., Born D., Barrera L., Hartigan A., Gaudelli N., Ciaramella G., Slaymaker I.M. (2021) Rationally designed base editors for precise editing of the sickle cell disease mutation. *CRISPR J.* **4**, 169–177.
40. Liu Y., Zhou C., Huang S., Dang L., Wei Y., He J., Zhou Y., Mao S., Tao W., Zhang Y., Yang H., Huang X., Chi T. (2020) A Cas-embedding strategy for minimizing off-target effects of DNA base editors. *Nat. Commun.* **11**, 6073.
41. Liu Z., Chen S., Shan H., Jia Y., Chen M., Song Y., Lai L., Li Z. (2020) Efficient base editing with high precision in rabbits using YFE-BE4max. *Cell Death Dis.* **11**, 36.
42. Zhang X., Chen L., Zhu B., Wang L., Chen C., Hong M., Huang Y., Li H., Han H., Cai B., Yu W., Yin S., Yang L., Yang Z., Liu M., Zhang Y., Mao Z., Wu Y., Liu M., Li D. (2020) Increasing the efficiency and targeting range of cytidine base editors through fusion of a single-stranded DNA-binding protein domain. *Nat. Cell Biol.* **22**, 740–750.
43. Jiang W., Feng S., Huang S., Yu W., Li G., Yang G., Liu Y., Zhang Y., Zhang L., Hou Y., Chen J., Chen J., Huang X. (2018) BE-PLUS: a new base editing tool with broadened editing window and enhanced fidelity. *Cell Res.* **28**, 855–861.
44. Liu Z., Chen S., Shan H., Jia Y., Chen M., Song Y., Lai L., Li Z. (2020) Precise base editing with CC context-specificity using engineered human APOBEC3G-nCas9 fusions. *BMC Biol.* **18**, 111.
45. Tan J., Zhang F., Karcher D., Bock R. (2020) Expanding the genome-targeting scope and the site selectivity of high-precision base editors. *Nat. Commun.* **11**, 629.
46. Li A., Mitsunobu H., Yoshioka S., Suzuki T., Kondo A., Nishida K. (2022) Cytosine base editing systems with minimized off-target effect and molecular size. *Nat. Commun.* **13**, 4531.
47. Liu Z., Shan H., Chen S., Chen M., Zhang Q., Lai L., Li Z. (2019) Improved base editor for efficient editing in GC contexts in rabbits with an optimized AID-Cas9 fusion. *FASEB J.* **33**, 9210–9219.
48. Tan J., Zhang F., Karcher D., Bock R. (2019) Engineering of high-precision base editors for site-specific single nucleotide replacement. *Nat. Commun.* **10**, 439.
49. Cheng T.L., Li S., Yuan B., Wang X., Zhou W., Qiu Z. (2019) Expanding C-T base editing toolkit with diversified cytidine deaminases. *Nat. Commun.* **10**, 3612.
50. Halperin S.O., Tou C.J., Wong E.B., Modavi C., Schaffer D.V., Dueber J.E. (2018) CRISPR-guided DNA polymerases enable diversification of all nucleotides in a tunable window. *Nature.* **560**, 248–252.
51. Tou C.J., Schaffer D.V., Dueber J.E. (2020) Targeted diversification in the *S. cerevisiae* genome with CRISPR-guided DNA polymerase I. *ACS Synth. Biol.* **9**, 1911–1916.
52. Hess G.T., Frésard L., Han K., Lee C.H., Li A., Cimprich K.A., Montgomery S.B., Bassik M.C. (2016) Directed evolution using dCas9-targeted somatic hypermutation in mammalian cells. *Nat. Methods.* **13**, 1036–1042.
53. Devilder M.C., Moyon M., Gautreau-Rolland L., Navet B., Perroteau J., Delbos F., Gesnel M.C., Breathnach R., Saulquin X. (2019) *Ex vivo* evolution of human antibodies by CRISPR-X: from a naive B cell repertoire to affinity matured antibodies. *BMC Biotechnol.* **19**, 14.
54. McCann J.L., Salamango D.J., Law E.K., Brown W.L., Harris R.S. (2020) MagnEdit-interacting factors that recruit

- DNA-editing enzymes to single base targets. *Life Sci. Alliance*. **3**, e201900606.
55. Gaudelli N.M., Komor A.C., Rees H.A., Packer M.S., Badran A.H., Bryson D.I., Liu D.R. (2017) Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. *Nature*. **551**, 464–471.
  56. Ryu S.M., Koo T., Kim K., Lim K., Baek G., Kim S.T., Kim H.S., Kim D.E., Lee H., Chung E., Kim J.S. (2018) Adenine base editing in mouse embryos and an adult mouse model of Duchenne muscular dystrophy. *Nat. Biotechnol.* **36**, 536–539.
  57. Zhou C., Sun Y., Yan R., Liu Y., Zuo E., Gu C., Han L., Wei Y., Hu X., Zeng R., Li Y., Zhou H., Guo F., Yang H. (2019) Off-target RNA mutation induced by DNA base editing and its elimination by mutagenesis. *Nature*. **571**, 275–278.
  58. Chatterjee P., Jakimo N., Jacobson J.M. (2018) Minimal PAM specificity of a highly similar SpCas9 ortholog. *Sci. Adv.* **4**, eaau0766.
  59. Yang L., Zhang X., Wang L., Yin S., Zhu B., Xie L., Duan Q., Hu H., Zheng R., Wei Y., Peng L., Han H., Zhang J., Qiu W., Geng H., Siwko S., Zhang X., Liu M., Li D. (2018) Increasing targeting scope of adenosine base editors in mouse and rat embryos through fusion of Tada deaminase with Cas9 variants. *Protein Cell*. **9**, 814–819.
  60. Richter M.F., Zhao K.T., Eton E., Lapinaite A., Newby G.A., Thuronyi B.W., Wilson C., Koblan L.W., Zeng J., Bauer D.E., Doudna J.A., Liu D.R. (2020) Phage-assisted evolution of an adenine base editor with improved Cas domain compatibility and activity. *Nat. Biotechnol.* **38**, 883–891.
  61. Grünewald J., Zhou R., Iyer S., Lareau C.A., Garcia S.P., Aryee M.J., Joung J.K. (2019) CRISPR DNA base editors with reduced RNA off-target and self-editing activities. *Nat. Biotechnol.* **37**, 1041–1048.
  62. Grünewald J., Zhou R., Garcia S.P., Iyer S., Lareau C.A., Aryee M.J., Joung J.K. (2019) Transcriptome-wide off-target RNA editing induced by CRISPR-guided DNA base editors. *Nature*. **569**, 433–437.
  63. Xu L., Zhang C., Li H., Wang P., Gao Y., Mokadam N.A., Ma J., Arnold W.D., Han R. (2021) Efficient precise *in vivo* base editing in adult dystrophic mice. *Nat. Commun.* **12**, 3719.
  64. Zhao D., Li J., Li S., Xin X., Hu M., Price M.A., Rosser S.J., Bi C., Zhang X. (2021) Glycosylase base editors enable C-to-A and C-to-G base changes. *Nat. Biotechnol.* **39**, 35–40.
  65. Kurt I.C., Zhou R., Iyer S., Garcia S.P., Miller B.R., Langner L.M., Grünewald J., Joung J.K. (2021) CRISPR C-to-G base editors for inducing targeted DNA transversions in human cells. *Nat. Biotechnol.* **39**, 41–46.
  66. Chen L., Park J.E., Paa P., Rajakumar P.D., Prekop H.T., Chew Y.T., Manivannan S.N., Chew W.L. (2021) Programmable C: G to G: C genome editing with CRISPR-Cas9-directed base excision repair proteins. *Nat. Commun.* **12**, 1384.
  67. Koblan L.W., Arbab M., Shen M.W., Hussmann J.A., Anzalone A.V., Doman J.L., Newby G.A., Yang D., Mok B., Rempel J.M., Xu A., Sisley T.A., Weissman J.S., Adamson B., Liu D.R. (2021) Efficient C•G-to-G•C base editors developed using CRISPRi screens, target-library analysis, and machine learning. *Nat. Biotechnol.* **39**, 1414–1425.
  68. Zeman M.K., Cimprich K.A. (2014) Causes and consequences of replication stress. *Nat. Cell Biol.* **16**, 2–9.
  69. de Moraes M.H., Hsu F., Huang D., Bosch D.E., Zeng J., Radey M.C., Simon N., Ledvina H.E., Frick J.P., Wiggins P.A., Peterson S.B., Mougous J.D. (2021) An interbacterial DNA deaminase toxin directly mutagenizes surviving target populations. *eLife*. **10**, e62967.
  70. Mok B.Y., de Moraes M.H., Zeng J., Bosch D.E., Kotrys A.V., Raguram A., Hsu F., Radey M.C., Peterson S.B., Mootha V.K., Mougous J.D., Liu D.R. (2020) A bacterial cytidine deaminase toxin enables CRISPR-free mitochondrial base editing. *Nature*. **583**, 631–637.
  71. Guo J., Chen X., Liu Z., Sun H., Zhou Y., Dai Y., Ma Y., He L., Qian X., Wang J., Zhang J., Zhu Y., Zhang J., Shen B., Zhou F. (2021) DdCBE mediates efficient and inheritable modifications in mouse mitochondrial genome. *Mol. Ther. Nucl. Acids*. **27**, 73–80.
  72. Schon E.A., DiMauro S., Hirano M. (2012) Human mitochondrial DNA: roles of inherited and somatic mutations. *Nat. Rev. Genet.* **13**, 878–890.
  73. Gammage P.A., Moraes C.T., Minczuk M. (2018) Mitochondrial genome engineering: the revolution may not be CRISPR-ized. *Trends Genet.* **34**, 101–110.
  74. Cho S.I., Lee S., Mok Y.G., Lim K., Lee J., Lee J.M., Chung E., Kim J.S. (2022) Targeted A-to-G base editing in human mitochondrial DNA with programmable deaminases. *Cell*. **185**, 1764–1776.e12.
  75. Fu J., Li Q., Liu X., Tu T., Lv X., Yin X., Lv J., Song Z., Qu J., Zhang J., Li J., Gu F. (2021) Human cell based directed evolution of adenine base editors with improved efficiency. *Nat. Commun.* **12**, 5897.
  76. Jeong Y.K., Lee S., Hwang G.H., Hong S.A., Park S.E., Kim J.S., Woo J.S., Bae S. (2021) Adenine base editor engineering reduces editing of bystander cytosines. *Nat. Biotechnol.* **39**, 1426–1433.
  77. Wang L., Xue W., Zhang H., Gao R., Qiu H., Wei J., Zhou L., Lei Y.N., Wu X., Li X., Liu C., Wu J., Chen Q., Ma H., Huang X., Cai C., Zhang Y., Yang B., Yin H., Yang L., Chen J. (2021) Eliminating base-editor-induced genome-wide and transcriptome-wide off-target mutations. *Nat. Cell Biol.* **23**, 552–563.
  78. Zhang X., Zhu B., Chen L., Xie L., Yu W., Wang Y., Li L., Yin S., Yang L., Hu H., Han H., Li Y., Wang L., Chen G., Ma X., Geng H., Huang W., Pang X., Yang Z., Wu Y., Siwko S., Kurita R., Nakamura Y., Yang L., Liu M., Li D. (2020) Dual base editor catalyzes both cytosine and adenine base conversions in human cells. *Nat. Biotechnol.* **38**, 856–860.
  79. Li C., Zhang R., Meng X., Chen S., Zong Y., Lu C., Qiu J.L., Chen Y.H., Li J., Gao C. (2020) Targeted, random mutagenesis of plant genes with dual cytosine and adenine base editors. *Nat. Biotechnol.* **38**, 875–882.
  80. Grünewald J., Zhou R., Lareau C.A., Garcia S.P., Iyer S., Miller B.R., Langner L.M., Hsu J.Y., Aryee M.J., Joung J.K. (2020) A dual-deaminase CRISPR base editor enables concurrent adenine and cytosine editing. *Nat. Biotechnol.* **38**, 861–864.
  81. Sakata R.C., Ishiguro S., Mori H., Tanaka M., Tatsuno K., Ueda H., Yamamoto S., Seki M., Masuyama N., Nishida K., Nishimasu H., Arakawa K., Kondo A., Nureki O., Tomita M., Aburatani H., Yachie N. (2020) Base editors for

- simultaneous introduction of C-to-T and A-to-G mutations. *Nat. Biotechnol.* **38**, 865–869.
82. Xie J., Huang X., Wang X., Gou S., Liang Y., Chen F., Li N., Ouyang Z., Zhang Q., Ge W., Jin Q., Shi H., Zhuang Z., Zhao X., Lian M., Wang J., Ye Y., Quan L., Wu H., Wang K., Lai L. (2020) ACBE, a new base editor for simultaneous C-to-T and A-to-G substitutions in mammalian systems. *BMC Biol.* **18**, 131.
  83. Liang Y., Xie J., Zhang Q., Wang X., Gou S., Lin L., Chen T., Ge W., Zhuang Z., Lian M., Chen F., Li N., Ouyang Z., Lai C., Liu X., Li L., Ye Y., Wu H., Wang K., Lai L. (2022) AGBE: a dual deaminase-mediated base editor by fusing CGBE with ABE for creating a saturated mutant population with multiple editing patterns. *Nucl. Acids Res.* **50**, 5384–5399.
  84. Kuscic C., Parlak M., Tufan T., Yang J., Szlachta K., Wei X., Mammadov R., Adli M. (2017) CRISPR-STOP: gene silencing through base-editing-induced nonsense mutations. *Nat. Methods.* **14**, 710–712.
  85. Billon P., Bryant E.E., Joseph S.A., Nambiar T.S., Hayward S.B., Rothstein, R., Ciccio A. (2017) CRISPR-mediated base editing enables efficient disruption of eukaryotic genes through induction of STOP codons. *Mol. Cell.* **67**, 1068–1079.e4.
  86. Wang X., Liu Z., Li G., Dang L., Huang S., He L., Ma Y., Li C., Liu M., Yang G., Huang X., Zhou F., Ma X. (2020) Efficient gene silencing by adenine base editor-mediated start codon mutation. *Mol. Ther.* **28**, 431–440.
  87. Lee C., Hyun Jo D., Hwang G.H., Yu J., Kim J.H., Park S.E., Kim J.S., Kim J.H., Bae S. (2019) CRISPR-Pass: gene rescue of nonsense mutations using adenine base editors. *Mol. Ther.* **27**, 1364–1371.
  88. Gapinske M., Luu A., Winter J., Woods W.S., Kostan K.A., Shiva N., Song J.S., Perez-Pinera P. (2018) CRISPR-SKIP: programmable gene splicing with single base editors. *Genome Biol.* **19**, 107.
  89. Anzalone A.V., Randolph P.B., Davis J.R., Sousa A.A., Koblán L.W., Levy J.M., Chen P.J., Wilson C., Newby G.A., Raguram A., Liu D.R. (2019) Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature.* **576**, 149–157.
  90. Landrum M.J., Lee J.M., Benson M., Brown G., Chao C., Chitipiralla S., Gu B., Hart J., Hoffman D., Hoover J., Jang W., Katz K., Ovetsky M., Riley G., Sethi A., Tully R., Villamarin-Salomon R., Rubinstein W., Maglott D.R. (2016) ClinVar: public archive of interpretations of clinically relevant variants. *Nucl. Acids Res.* **44**, D862–D868.
  91. Anzalone A.V., Koblán L.W., Liu D.R. (2020) Genome editing with CRISPR-Cas nucleases, base editors, transposases and prime editors. *Nat. Biotechnol.* **38**, 824–844.
  92. Marzec M., Brąszewska-Zalewska A., Hensel G. (2020) Prime editing: a new way for genome editing. *Trends Cell Biol.* **30**, 257–259.
  93. 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**, 257.
  94. Jang H., Jo D.H., Cho C.S., Shin J.H., Seo J.H., Yu G., Gopalappa R., Kim D., Cho S.R., Kim J.H., Kim H.H. (2022) Application of prime editing to the correction of mutations and phenotypes in adult mice with liver and eye diseases. *Nat. Biomed. Eng.* **6**, 181–194.
  95. Schene I.F., Joore I.P., Oka R., Mokry M., van Vugt A.H.M., van Boxtel R., van der Doef H.P.J., van der Laan L.J.W., Versteegen M.M.A., van Hasselt P.M., Nieuwenhuis E.E.S., Fuchs S.A. (2020) Prime editing for functional repair in patient-derived disease models. *Nat. Commun.* **11**, 5352.
  96. Petri K., Zhang W., Ma J., Schmidts A., Lee H., Horng J.E., Kim D.Y., Kurt I.C., Clement K., Hsu J.Y., Pinello L., Maus M.V., Joung J.K., Yeh J.J. (2022) CRISPR prime editing with ribonucleoprotein complexes in zebrafish and primary human cells. *Nat. Biotechnol.* **40**, 189–193.
  97. Bosch J.A., Birchak G., Perrimon N. (2021) Precise genome engineering in *Drosophila* using prime editing. *Proc. Natl. Acad. Sci. USA.* **118**, e2021996118.
  98. Qian Y., Zhao D., Sui T., Chen M., Liu Z., Liu H., Zhang T., Chen S., Lai L., Li Z. (2021) Efficient and precise generation of Tay-Sachs disease model in rabbit by prime editing system. *Cell Discov.* **7**, 50.
  99. Lin Q., Zong Y., Xue C., Wang S., Jin S., Zhu Z., Wang Y., Anzalone A.V., Raguram A., Doman J.L., Liu D.R., Gao C. (2020) Prime genome editing in rice and wheat. *Nat. Biotechnol.* **38**, 582–585.
  100. Lu Y., Tian Y., Shen R., Yao Q., Zhong D., Zhang X., Zhu J.K. (2021) Precise genome modification in tomato using an improved prime editing system. *Plant Biotechnol.* **19**, 415–417.
  101. Ferreira da Silva J., Oliveira G.P., Arasa-Verge E.A., Kagiou C., Moretton A., Timelthaler G., Jiricny J., Loizou J.I. (2022) Prime editing efficiency and fidelity are enhanced in the absence of mismatch repair. *Nat. Commun.* **13**, 760.
  102. Chen P.J., Hussmann J.A., Yan J., Knipping F., Ravisankar P., Chen P.F., Chen C., Nelson J.W., Newby G.A., Sahin M., Osborn M.J., Weissman J.S., Adamson B., Liu D.R. (2021) Enhanced prime editing systems by manipulating cellular determinants of editing outcomes. *Cell.* **184**, 5635–5652.e29.
  103. Aida T., Wilde J.J., Yang L., Hou Y., Li M., Xu D., Lin J., Qi P., Lu Z., Feng G. (2020) Prime editing primarily induces undesired outcomes in mice. *bioRxiv.* 08.06.239723.
  104. Averina O.A., Permyakov O.A., Grigorieva O.O., Starshin A.S., Mazur A.M., Prokhortchouk E.B., Dontsova O.A., Sergiev P.V. (2021) Comparative analysis of genome editors efficiency on a model of mice zygotes microinjection. *J. Mol. Sci.* **22**, 10221.
  105. Park S.J., Jeong T.Y., Shin S.K., Yoon D.E., Lim S.Y., Kim S.P., Choi J., Lee H., Hong J.I., Ahn J., Seong J.K., Kim K. (2021) Targeted mutagenesis in mouse cells and embryos using an enhanced prime editor. *Genome Biol.* **22**, 170.
  106. Nelson J.W., Randolph P.B., Shen S.P., Everette K.A., Chen P.J., Anzalone A.V., An M., Newby G.A., Chen J.C., Hsu A., Liu D.R. (2022) Engineered pegRNAs improve prime editing efficiency. *Nat. Biotechnol.* **40**, 402–410.
  107. Velimirovic M., Zanetti L.C., Shen M.W., Fife J.D., Lin L., Cha M., Akinci E., Barnum D., Yu T., Sherwood R.I. (2022) Peptide fusion improves prime editing efficiency. *Nat. Commun.* **13**, 3512.
  108. Song M., Lim J.M., Min S., Oh J.S., Kim D.Y., Woo J.S., Nishimasu H., Cho S.R., Yoon S., Kim H.H. (2021) Generation of a more efficient prime editor 2 by addition of the Rad51 DNA-binding domain. *Nat. Commun.* **12**, 5617.
  109. Kweon J., Yoon J.K., Jang A.H., Shin H.R., See J.E., Jang G., Kim J.I., Kim Y. (2021) Engineered prime editors with PAM flexibility. *Mol. Ther.* **29**, 2001–2007.

110. Li X., Zhou L., Gao B.Q., Li G., Wang X., Wang Y., Wei J., Han W., Wang Z., Li J., Gao R., Zhu J., Xu W., Wu J., Yang B., Sun X., Yang L., Chen J. (2022) Highly efficient prime editing by introducing same-sense mutations in pegRNA or stabilizing its structure. *Nat. Commun.* **13**, 1669.
111. Liu Y., Yang G., Huang S., Li X., Wang X., Li G., Chi T., Chen Y., Huang X., Wang X. (2021) Enhancing prime editing by Csy4-mediated processing of pegRNA. *Cell Res.* **31**, 1134–1136.
112. Zhang G., Liu Y., Huang S., Qu S., Cheng D., Yao Y., Ji Q., Wang X., Huang X., Liu J. (2022) Enhancement of prime editing via xrRNA motif-joined pegRNA. *Nat. Commun.* **13**, 1856.
113. Li X., Wang X., Sun W., Huang S., Zhong M., Yao Y., Ji Q., Huang X. (2022) Enhancing prime editing efficiency by modified pegRNA with RNA G-quadruplexes. *J. Mol. Cell. Biol.* **14**, mjac022.
114. Anzalone A.V., Gao X.D., Podracky C.J., Nelson A.T., Koblan L.W., Raguram A., Levy J.M., Mercer J.A.M., Liu D.R. (2022) Programmable deletion, replacement, integration and inversion of large DNA sequences with twin prime editing. *Nat. Biotechnol.* **40**, 731–740.
115. Yuan Q., Gao X. (2022) Multiplex base- and prime-editing with drive-and-process CRISPR arrays. *Nat. Commun.* **13**, 2771.
116. Fiumara M., Ferrari S., Omer-Javed A., Beretta S., Albano L., Canarutto D., Varesi A., Gaddoni C., Brombin C., Cugnata F., Zonari E., Naldini M.M., Barcella M., Gentner B., Merelli I., Naldini L. (2023) Genotoxic effects of base and prime editing in human hematopoietic stem cells. *Nat. Biotechnol.* Sep 7. doi: 10.1038/s41587-023-01915-4. Epub ahead of print. Erratum in: *Nat. Biotechnol.* 2024 Jan 25; PMID: 37679541.
117. Testa L.C., Musunuru K. (2023) Base editing and prime editing: potential therapeutic options for rare and common diseases. *BioDrugs: Clin. Immunotherapy, Biopharm. Gene Therapy.* **37**, 453–462.

## Current Knowledge of Base Editing and Prime Editing

**O. A. Averina<sup>1,2,3</sup>, S. A. Kuznetsova<sup>1,\*</sup>, O. A. Permyakov<sup>1,3</sup>, P. V. Sergiev<sup>1,2,3</sup>**

<sup>1</sup>*Institute of Functional Genomics, Lomonosov Moscow State University, Moscow, 119991 Russia*

<sup>2</sup>*Belozersky Institute of Physico-Chemical Biology, Lomonosov Moscow State University, Moscow, 119991 Russia*

<sup>3</sup>*Department of Chemistry, Lomonosov Moscow State University, Moscow, 119991 Russia*

*\*e-mail: svetlana@belozersky.msu.ru*

Modern genetic engineering technologies, such as base editing (BE) and prime editing (PE), have proven to be effective and reliable genome editing tools that do not require the introduction of double-strand breaks in DNA and the presence of donor templates. Relatively new, they quickly gained recognition for their accuracy, simplicity and multiplexing capabilities. This review summarizes new literature on these technologies: architecture and methods for creating editors, specificity, efficiency and versatility. The advantages, disadvantages and prospects for using these editors in basic and applied research are discussed. The information presented in the review may be useful for planning genome editing studies and for analyzing their results when solving various problems in fundamental biology, biotechnology, medicine and agriculture.

**Keywords:** CRISPR/Cas technology, base editing, nucleoside deaminases, prime editing, reverse transcriptase, repair systems

УДК 577.21

## КАК СМЕСТИТЬ РАВНОВЕСИЕ РЕПАРАЦИИ РАЗРЫВОВ ДНК В ПОЛЬЗУ ГОМОЛОГИЧНОЙ РЕКОМБИНАЦИИ

© 2024 г. О. А. Аверина<sup>a, b, c</sup>, С. А. Кузнецова<sup>a, \*</sup>, О. А. Пермяков<sup>a, c</sup>, П. В. Сергиев<sup>a, b, c</sup><sup>a</sup>Институт функциональной геномики, Московский государственный университет им. М.В. Ломоносова, Москва, 119991 Россия<sup>b</sup>Научно-исследовательский институт физико-химической биологии им. А.Н. Белозерского, Московский государственный университет им. М.В. Ломоносова, Москва, 119991 Россия<sup>c</sup>Химический факультет, Московский государственный университет им. М.В. Ломоносова, Москва, 119991 Россия

\*e-mail: svetlana@belozersky.msu.ru

Поступила в редакцию 15.11.2023 г.

После доработки 28.12.2023 г.

Принята к публикации 28.01.2024 г.

Появление технологии CRISPR/Cas направленного редактирования генома позволило с высокой эффективностью проводить генно-инженерные манипуляции с эукариотическими геномами. Один из ключевых этапов этой технологии – целенаправленная индукция сайт-специфических расщеплений (разрывов) ДНК. Репарация этих разрывов проходит по одному из двух путей: негомологичного соединения концов или гомологичной рекомбинации. Выбор пути репарации обусловлен архитектурой участков в месте разрыва ДНК, образующихся в результате резекции концов, и зависит от фазы клеточного цикла. Негомологичное соединение концов – основной путь репарации двухцепочечных разрывов ДНК в клетках млекопитающих – включает неспецифическую реакцию лигирования, точность которой зависит от структуры концов разрыва, а результатом могут быть различные вставки или делеции в целевом участке генома. Встраивание в геном нужной последовательности проходит по пути гомологичной рекомбинации, для реализации которого необходима матрица с участками гомологии по обеим сторонам от двухцепочечного разрыва. Введение генетической конструкции в заданное место генома является важной, но на данный момент сложной и трудозатратной задачей. Вместе с тем, выбор пути репарации может иметь принципиальное значение для фундаментальных исследований функции генов и создания трансгенных животных, моделирующих заболевания человека, с целью разработки методов их терапии. В настоящем обзоре предпринята попытка объединить и структурировать информацию о подходах к увеличению эффективности репарации ДНК с участием гомологичной рекомбинации. Перечислены наиболее эффективные стратегии, позволяющие сдвинуть равновесие в сторону гомологичной рекомбинации, такие как использование ингибиторов негомологичного соединения концов, регуляция ключевых факторов гомологичной рекомбинации, контроль клеточного цикла, статуса хроматина, конструирование матриц для гомологичной рекомбинации.

**Ключевые слова:** редактирование генома, технология CRISPR/Cas, нуклеазы, системы репарации, гомологичная рекомбинация

DOI: 10.31857/S0026898424040029, EDN: INFKTE

### ВВЕДЕНИЕ

Новые стратегии и подходы к направленному редактированию генома, разработанные за последние десятилетия, активно используются

для решения различных задач фундаментальной биологии, биотехнологии, медицины и сельского хозяйства. Появление геномного редактирования позволило изучать функциональную роль генов млекопитающих, создавать животные модели

Сокращения: CRISPR – короткие палиндромные повторы, разделенные уникальными последовательностями-спейсерами (Clustered Regularly Interspaced Short Palindromic Repeats); CRISPR/Cas-ассоциированная нуклеаза; PAM – смежный протоспейсерный мотив (Protospacer Adjacent Motif); PFS – последовательность нуклеотидов, фланкирующая протоспейсер (Protospacer Flanking Sequence); NHEJ – негомологичное соединение концов (Nonhomologous DNA End Joining); MMEJ – соединение концов ДНК на основе микрогомологии (Microhomology-Mediated End Joining); HDR – гомологичная

рекомбинация (Homology Directed Recombination); rPHK – гидовая РНК (sgRNA – single-guide RNA), tracrPHK – транс-активирующая РНК (trans-activating RNA); crPHK – направляющая РНК (crisprRNA); ОЦР – одноцепочечные разрывы; ДЦР – двухцепочечные разрывы; RuvC – каталитический домен нуклеазы Cas9, вносящий одноцепочечный разрыв в целевой участок ДНК (протоспейсер); HNH – каталитический домен нуклеазы Cas9, вносящий одноцепочечный разрыв в целевой участок ДНК и связывающийся со спейсером crPHK; аминокислотные остатки – а.к. (при цифре).

генетических заболеваний человека и разрабатывать подходы к их диагностике и лечению, усовершенствовать значимые для сельского хозяйства виды растений. Конструирование генетических редакторов, позволяющих избирательно активировать, модифицировать или выключать целевые гены-мишени, в том числе ассоциированные с патологическими процессами и генетическими заболеваниями, является одним из приоритетных направлений развития генетических технологий. Знание принципов создания генетических редакторов, обладающих высокой эффективностью, точностью редактирования, безопасностью и доступностью для широкого круга исследователей, позволит ученым успешно решать задачи высокоточного редактирования генома растений, животных и человека. В настоящем обзоре суммированы сведения об усовершенствовании и разработке новых высокоточных геномных редакторов CRISPR/Cas с повышенной эффективностью гомологичной рекомбинации (HDR), что позволит вводить целевые генетические конструкции в заданные участки генома.

#### ЭНДОНУКЛЕАЗНАЯ СИСТЕМА РЕДАКТИРОВАНИЯ ГЕНОМА CRISPR/CAS

Изначально CRISPR/Cas — это система адаптивного иммунитета архей и бактерий, нацеленная на чужеродные генетические элементы вторгшихся фагов. Система CRISPR/Cas включает геномные локусы CRISPR, состоящие из коротких палиндромных повторов, разделенных уникальными последовательностями-спейсерами, встроенными в геном в ходе адаптивного иммунного ответа, и белков Cas, кодируемых фланкирующими локусы генами. Белки Cas обеспечивают реализацию молекулярного механизма адаптивного иммунитета, который можно разделить на три стадии: адаптации — встраивания в CRISPR-кассету новых спейсеров; процессинга, во время которого происходит транскрипция CRISPR-массива в пре-crPHK (предшественник crPHK) и ее специфичный гидролиз с образованием зрелых crPHK; интерференции — распознавания целевой нуклеиновой кислоты и ее разрушения. На стадии интерференции белки Cas эффекторного комплекса, сформированного с участием CRISPR-PHK (crPHK), расщепляют ДНК или РНК атакующих фагов, содержащих участок, комплементарный последовательности спейсера, и тем самым дезактивируют их. Специфичность в отношении целевого участка нуклеиновой кислоты фагов достигается за счет малых некодирующих РНК, известных как направляющие, или гидовые РНК (гРНК), которые составляют клеточную память о прошлых инфекциях. гРНК представляет собой молекулу РНК, которая содержит переменную спейсерную последовательность, направляющую Cas в гомологичную целевую область

протоспейсера в геноме инфекционных элементов. гРНК может содержать два сегмента: сегмент нацеливания ДНК (crPHK) и сегмент, связывающий белок (tracrPHK).

Этот механизм лег в основу системы редактирования генома CRISPR/Cas для изучения функций генов, создания мутантных клеточных линий и животных, методов диагностики и терапии различных заболеваний.

CRISPR-системы подразделяют на два класса в зависимости от архитектуры их эффекторных комплексов, в каждом из которых выделяют по три типа (табл. 1) [1]:

- в системах первого класса несколько белковых единиц образуют эффекторный комплекс с crPHK для распознавания и расщепления целевой последовательности;

- системы второго класса состоят из односубъединичных эффекторных нуклеаз с разнообразными функциональными свойствами.

Белки Cas играют важнейшую роль на всех этапах работы этой защитной системы. Они обладают различными функциональными свойствами и ферментативной активностью. Так, белки Cas1 и Cas2, задействованные в большинстве систем CRISPR/Cas на этапе адаптации, формируют комплекс, где белки Cas1 функционируют как металлзависимые оцДНК/дцДНК-эндонуклеазы, которые вырезают фрагмент из вирусной ДНК и вносят разрывы в локус CRISPR, а также как интегразы, которые встраивают новый спейсер в бактериальный геном. Белки Cas2 обладают рибонуклеазной (PHKаза, специфичная к U-богатым участкам) или дезоксирибонуклеазной (дцДНКаза) активностью в зависимости от системы, к которой принадлежат, и играют структурную роль. Белок Csn системы CRISPR/Cas класса 2 (тип 2) участвует во встраивании новых спейсеров. Белки Cas3—Cas10, Cas12 и Cas13 являются компонентами эффекторного комплекса и участвуют в интерференции. Так, белки Cas3' содержат хеликазный домен и обладают хеликазной активностью, благодаря чему расплетают двойную спираль ДНК-мишени, а также имеют нуклеазный домен (HD-домен), который разрезает оцДНК-мишень. Белки Cas3'' — это металлзависимые дцДНК-эндонуклеазы, а белки Cas4 — нуклеазы с трехцистеиновым C-концевым кластером, обладающие 5'-экзонуклеазной активностью (оцДНКаза). Рибонуклеазы Cas5 и Cas6 участвуют в процессинге пре-crPHK, Cas7 участвует в связывании crPHK, а Cas8 — в узнавании PAM. Нуклеазы Cas9 (дцДНК-эндонуклеазы) и Cas12 (оцДНК/дцДНК-эндонуклеазы) участвуют в разрезании ДНК-мишени, а Cas13 расщепляют оцPHK. Cas10 содержат Palm-домен (PHK-распознающий домен) и обладают ДНКазной активностью.

Таблица 1. Классификация и организация системы CRISPR/Cas

Класс	Тип	Подтип	Белки Cas	Организация оперонов для типа Cas в зависимости от вида бактерий и архей
Класс 1	Тип 1	I-A, I-B, I-C, I-D, I-E, I-F, I-U	Cas1, Cas2, Cas3', Cas3'', Cas4, Cas5, Cas6, Cas7, Cas8	I-A: <i>Archaeoglobus fulgidus</i> I-B: <i>Clostridium kluyveri</i> I-C: <i>Bacillus halodurans</i> I-D: <i>Cyanothece</i> sp. I-E: <i>Escherichia coli</i> K12 I-F: <i>Yersinia pseudotuberculosis</i> I-U: <i>Geobacter sulfurreducens</i>
	Тип 3	III-A, III-B, III-C, III-D	Cas1, Cas2, Cas5, Cas6, Cas7, Cas9, Cas10	III-A: <i>Staphylococcus epidermidis</i> III-B: <i>Pyrococcus furiosus</i> III-C: <i>Methanothermobacter thermautotrophicus</i> III-D: <i>Synechocystis</i> sp.
	Тип 4		Cas1, Cas2, Cas5, Cas6, Cas7	IV: <i>Thioalkalivibrio</i> sp. K90mix
Класс 2	Тип 2	II-A, II-B, II-C	Cas1, Cas2, Cas4, Cas9, Csn	II-A: <i>Streptococcus pyogenes</i> II-B: <i>Legionella pneumophila</i> str. Paris II-C: <i>Neisseria lactamica</i> 020-06
	Тип 5	V-A, V-B, V-C, V-D, V-E	Cas1, Cas2, Cas4, Cas12	V-A: <i>Francisella</i> cf. <i>novicida</i> Fx1 V-B: <i>Alicyclobacillus acidoterrestris</i> V-C: <i>Oleiphilus</i> sp. V-D: <i>Bacterium</i> CG09_39_24 V-E: <i>Deltaproteobacteria bacterium</i>
	Тип 6	VI-A, VI-C, VI-B1, VI-B2	Cas1, Cas2, Cas13	VI-A: <i>Leptotrichia shahii</i> VI-C: <i>Fusobacterium prefoetens</i> VI-B1: <i>Prevotella buccae</i> VI-B2: <i>Bergeyella zoohelcum</i>

Для создания мутантных животных наиболее широко используется система CRISPR/Cas9 второго класса подтипа II-A *Streptococcus pyogenes* (SpCas9), поскольку она относительно просто устроена и хорошо изучена [2]. Для адресации Cas9 на целевой участок генома в некоторых исследованиях используют гРНК с отдельными компонентами сгРНК и тасгРНК. Этот вариант хорош, когда необходимо одновременно использовать множество различных сгРНК, сочетаемых с одной тасгРНК. Для единичного разрезания более удобно использовать единую гидовую РНК (егРНК), где специально разработанная короткая последовательность сгРНК связана линкерной петлей (тетрапетля) с каркасной последовательностью тасгРНК (рис. 1) [3]. В итоге егРНК стала наиболее популярным форматом направляющих РНК, поэтому термины егРНК и гРНК часто применяют в сообществе CRISPR в одном и том же значении.

Для более эффективного использования и повышения стабильности по отношению к нуклеазной деградации сгРНК и тасгРНК модифицируют, включая в их структуру неприродные фрагменты. Так, в обзоре Филипповой и соавт. подробно описаны стратегии создания искусственных гРНК, содержащих модифицированный углеводофосфатный остов, нативные структурные мотивы и метки

для визуализации. Описаны различные подходы к улучшению функции синтетической гРНК [4]. Отмечается, что химические модификации компонентов гРНК могут повышать ее устойчивость к гидролизу, изменять термодинамическую стабильность комплексов РНК-белок и РНК-ДНК, снижать иммуногенный и цитотоксический эффекты. В целом, можно проводить широкомаштабные химические модификации, изменяя около 70% нуклеотидов, при этом избегая модификации 2'-ОН и фосфатных групп в егРНК, которые взаимодействуют с белком Cas9. Продемонстрировано, что химически модифицированные направляющие РНК сохраняют в среднем высокую специфичность [5, 6]. При этом неудачный дизайн сгРНК может привести к более низкой специфичности и более высокому проценту нецелевых мутаций. Так, например, удаление с 5'-конца сгРНК 2–3 нуклеотидов значительно снижает количество нецелевых мутаций при сохранении эффективности, но если егРНК укоротить до 16 нуклеотидов или менее, то ее активность резко падает, как и при удалении с 3'-конца или добавлении на 5'-конец динуклеотида GG [7]. Для конструирования оптимальной структуры сгРНК используют различные вычислительные инструменты, такие как CRISPR-P 2.0, E-CRISP и CasFinder [8]. Таким образом, оптимальный дизайн направляющей РНК

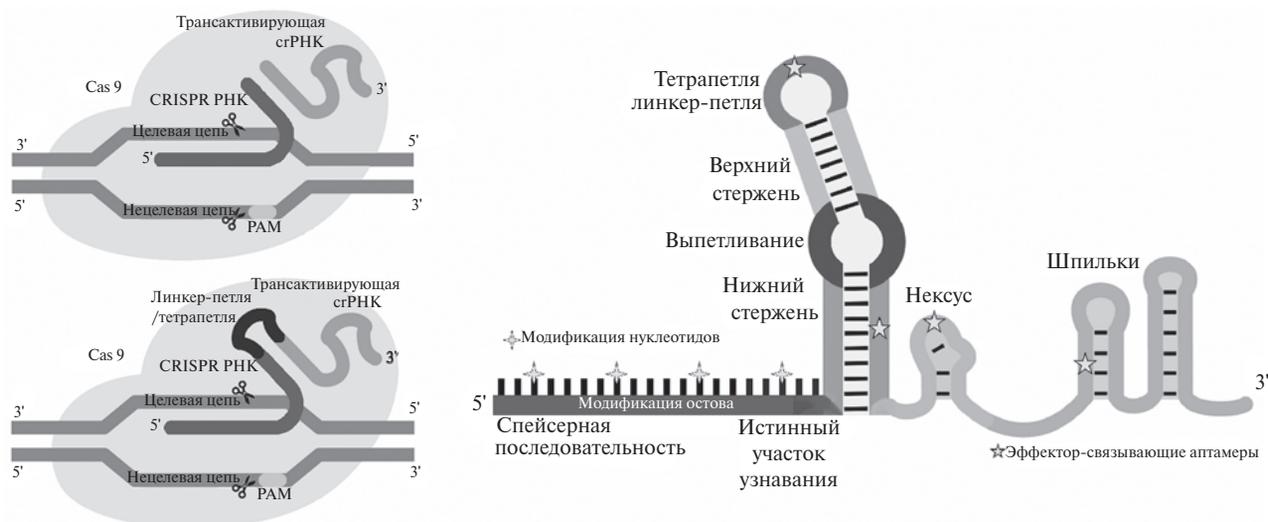


Рис. 1. Варианты гидовой РНК.

является ключом к CRISPR/Cas-опосредованному точному редактированию генома.

Впервые систему CRISPR/Cas9 применили в 2013 году для локального редактирования генома в клетках млекопитающих. В настоящее время с использованием CRISPR/Cas9 создают животные модели с точечными мутациями в геноме, вставками или делециями нескольких нуклеотидов без добавления иных генетических элементов вблизи места редактирования, например, маркеров для селекции клеток по чувствительности к препарату в культуральной среде и рекомбинантных элементов, таких как loxP- или FRT-сайты [9]. Внедрение системы CRISPR/Cas9 позволило получать мутантных мышей на генетическом фоне, недоступном для редактирования генома с использованием более ранних подходов, например, иммунодефицитной линии NOD/Scid-ILgamma (NSG) [10]. Также CRISPR/Cas9 была адаптирована для биаллельного редактирования в разных локусах и создания мышей с мутациями в двух генах [11–14]. Прямое сравнение системы редактирования генома CRISPR/Cas с предыдущими платформами, такими как ZFN (введение двухцепочечных разрывов (ДЦР) в ДНК с помощью нуклеаз с мотивом цинковых пальцев), TALEN (введение разрывов с использованием нуклеаз TALE (transcription activator-like effector nucleases)) (рис. 2) и методами направленной модификации генома с помощью эмбриональных стволовых клеток (ES), ясно показывает, что CRISPR/Cas обладает рядом преимуществ, таких как более высокая эффективность, простота использования, относительно низкая стоимость, компактность и ряд других [9].

Высокая эффективность SpCas9 позволила идентифицировать потенциально важные гены с помощью полногеномного функционального скрининга. Так, используя лентивирусные гРНК

для генерации большого количества клеток с нокаутом, после проведения положительной и отрицательной селекции можно идентифицировать важные гены и мишени для лекарственных средств [15–17].

Метод SpCas9 проявил себя мощным и универсальным инструментом генной инженерии, совершившим революцию в биологических и биомедицинских исследованиях, однако он имеет ряд ограничений. Так, иногда могут образовываться нежелательные ДЦР в нецелевом геномном локусе [18]. Специфичность расщепления ДНК определяется структурой Cas9, комплементарностью гРНК к ДНК и наличием последовательности PAM, прилегающей к целевой последовательности (рис. 3). Однако в процессе гибридизации возможно несовпадение нуклеотидов между гРНК и ДНК [19] или ограничения в доступности и распознавании сайтов PAM, так как SpCas9 может узнавать последовательности PAM, отличные от 5'-NGG-3'. Эффективность расщепления в этом случае становится значительно ниже (как, например, в случае 5'-NAG-3') [20]. Другим фактором может стать количество ошибок при спаривании оснований, когда при возникновении двух и более ошибок активность эндонуклеазы падает [21]. После возникновения незапланированных ДЦР могут появиться случайные мутации, обусловленные запуском подверженного ошибкам пути репарации ДНК – негомлогичного соединения концов (NHEJ) [22]. Эти нежелательные нецелевые ДЦР также могут привести к аномалиям хромосомного масштаба [23]. Так, при работе с Cas9 непреднамеренно возникали крупные делеции (тысячи пар нуклеотидов), нацеленные на различные локусы в разных хромосомах, которые вызывают большую интерстициальную “потерю гетерозиготности” [24–26]. Самое опасное, что такие непредвиденные большие делеции трудно обнаружить

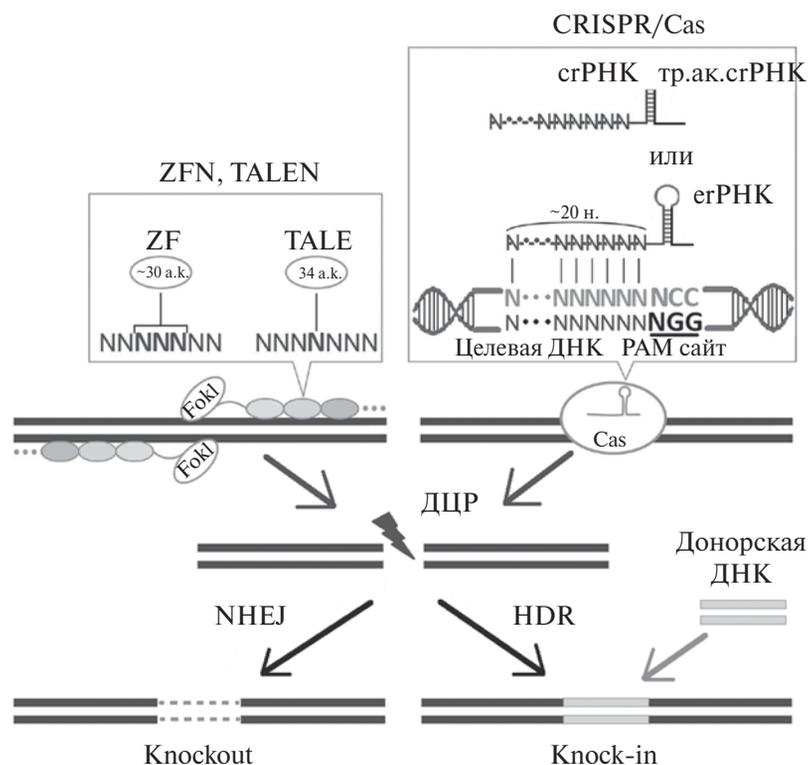
с помощью классических методов скрининга. Описаны случаи обнаружения делеции в 5.2 т.п.н. с помощью количественной ПЦР и флуоресцентной гибридизации *in situ* [25] и от 100 п.н. до 9.5 т.п.н. – с помощью ПЦР длинных фрагментов [27].

К недостаткам SpCas9 относится его большой размер, что ограничивает выбор вирусных векторов для доставки. Так, емкость аденоассоциированных вирусов (ААВ) равна приблизительно 4.7 т.п.н., тогда как SpCas9 вместе с гРНК занимает около 4.2 т.п.н. [28]. Клиническое применение технологии сдерживается ее иммуногенностью, отсутствием безопасной и эффективной системы доставки, нецелевыми разрезами и этическими проблемами [29].

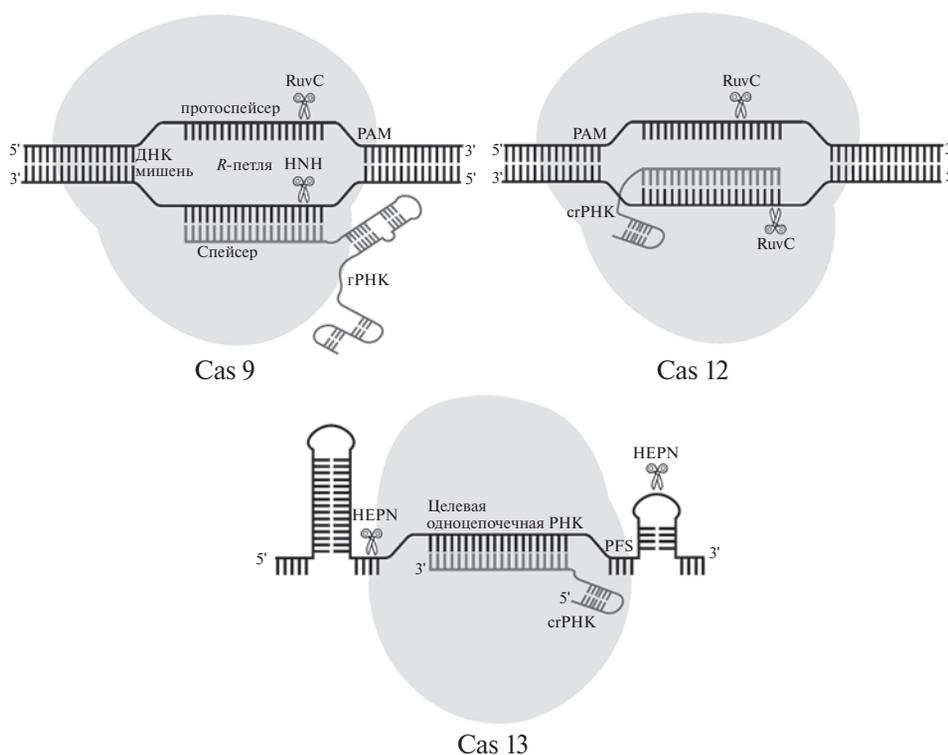
Для решения обозначенных проблем разработаны усовершенствованные системы и технологии.

Во-первых, сконструирована серия высокоточных вариантов SpCas9 (eSpCas9 1.0 & 1.1,

SpCas9-HF1, HypaCas9, Sniper-Cas9, SpG-HF1, SpRY-HF1, HiFi-iSpyMac), а также вариантов, способных распознавать широкий спектр последовательностей PAM (SpCas9-EQR/VQR/VRER/VRQR/NRRH/NRCH/NRTH, xCas9 [30]). Во-вторых, в качестве альтернативы SpCas9 можно использовать нуклеазы Cas9 *Staphylococcus aureus* (SaCas9) и *Campylobacter jejuni* (CjCas9), которые имеют пониженную нецелевую нуклеазную активность, поскольку распознают более длинные последовательности PAM, 5'-NNGRRT-3' и 5'-NNNRYAC-3', соответственно [31, 32], или варианты Cas10 и Cas14, которым вообще не требуется последовательность PAM [33]. Мутантный вариант *Acidaminococcus* sp. *BV3L6* Cas12a (AsCas12a) – E174R/S542R/K548R распознает различные PAM, включая 5'-TTYN-3', 5'-VTTV-3' и 5'-TRTV-3', а другие мутанты AsCas12a – RR и RVR, распознают TYCV и 5'-TATV-3' PAM соответственно [34]. Кроме того, мутации G532R/K595R были



**Рис. 2.** Технологии редактирования генома ZFN, TALEN, CRISPR/Cas. ZFN – сайт-специфическое расщепление ДНК химерной ZF-нуклеазой, сконструированной на основе ДНК-связывающего домена типа “цинковых пальцев” (30 а.к., взаимодействующих с триплетом нуклеотидов в ДНК) и каталитического домена эндонуклеазы рестрикции FokI. TALEN – расщепление ДНК химерной нуклеазой TALE, сконструированной путем присоединения к ДНК-узнающим белкам TALE каталитического домена эндонуклеазы FokI; один мономер ДНК-связывающего домена TALEN (34 а.к.) узнает один нуклеотид целевой последовательности ДНК. CRISPR/Cas – узнавание целевой последовательности ДНК происходит с помощью гРНК, состоящей из направляющей и транскрибирующей РНК (либо с помощью егРНК, в которой направляющая и транскрибирующая РНК связаны линкером), а расщепление – нуклеазой Cas. PAM – короткий NGG-мотив, необходимый для связывания Cas и внесения разрывов в ДНК. Репарация ДНК проходит по одному из двух путей: негомологичного соединения концов (NHEJ) или гомологичной рекомбинации (HDR) с участием донорной ДНК-матрицы. В результате репарации разрывов в ДНК происходит нокаут определенного гена (Knockout) или встраивание в геном целевой генетической конструкции (Knock-in). тр.ак.сrPHK – транскрибирующая сrPHK.



**Рис. 3.** Организация систем CRISPR/Cas9, CRISPR/Cas12 и CRISPR/Cas13. Стрелками обозначены места одноцепочечных разрывов ДНК каталитическими доменами RuvC и HNH (Cas9 и Cas12) и места разрыва РНК белком HEPN (Cas13).

введены в *Lachnospiraceae bacterium* ND Cas12a (LbCas12a) для создания варианта LbCas12a-RR, который распознает 5'-TGYCV-3' PAM, расширяя область нацеливания [35]. В-третьих, во избежание появления хромосомных транслокаций, нецелевых и неконтролируемых мутаций, образующихся в результате ДЦР (см. выше), разработаны дополнительные инструменты редактирования генов, в которых не происходит образования ДЦР. Так, инактивация одного из нуклеазных доменов Cas9 (как расщепляющего цепь, на которую направлена гРНК, так и расщепляющего нецелевую цепь ДНК), приводит к образованию нуклеазы Cas9 (nCas9), осуществляющей только одноцепочечный разрыв (ОЦР) ДНК [3, 36]. Инактивация обоих нуклеазных доменов приводит к образованию каталитически полностью неактивной формы Cas9 (dead Cas9, или dCas9) [37]. Белки nCas9 и dCas9 составляют строительные блоки редакторов в усовершенствованных технологиях редактирования CRISPR/Cas: редактирования оснований и праймированного редактирования. В последующих поколениях генетических редакторов белок dCas9 был объединен с белками-регуляторами транскрипции или ферментами, модифицирующими хроматин, чтобы регулировать только уровень транскрипции без необратимого изменения генома. Это привело к созданию систем интерференции (CRISPRi) и активации (CRISPRa). Так, в системах CRISPRi

репрессор KRAB (Krüppel-associated box), объединенный с dCas9, взаимодействует с комплексами, образующими гетерохроматин. Эти комплексы могут индуцировать метилирование и деацетилирование гистонов и, как следствие, ингибирование связывания РНК-полимераз с энхансерными или промоторными областями и инактивацию транскрипции [38]. Напротив, системы CRISPRa активируют транскрипцию с помощью доменов активации VP16 или более эффективных *in vivo* VP64 и VP192 путем взаимодействия с TATA-связывающим белком, транскрипционным фактором TFIIB и гистон-ацетилазой SAGA [39, 40]. Технологии CRISPRi и CRISPRa представляют интерес для проведения генетического скрининга, иммунотерапии опухолей и перепрограммирования плюрипотентных стволовых клеток [41–43].

Получить устойчивую наследуемую модификацию метилирования ДНК, подавляющую транскрипцию генов, можно с помощью временной экспрессии эпигенетического редактора CRISPRoff, созданного путем слияния dCas9 с KRAB и ДНК-(цитозин-5)-метилтрансферазой 3A (DNMT3A). Ранее было показано, что комплекс dCas9-DNMT3A может индуцировать сайт-специфическое метилирование CpG, приводящее к сайленсингу транскрипции. CRISPRoff может найти применение в скрининге

генома, клеточной инженерии, подавлении энхансеров и генетическом механизме эпигенетической модификации [44].

Уменьшить нецелевые эффекты CRISPR/Cas9 можно также путем снижения чрезмерной и длительной экспрессии этой системы. С этой целью используют механизм, защищающий фаги от CRISPR/Cas-опосредованного иммунитета бактерий, а именно, белки Aсг, нацеленные на нуклеазы Cas. В общей сложности идентифицировано 44 белка Aсг для CRISPR/Cas типа I, II и V, но не для подтипа II-V и типа VI [45]. Известны три механизма Aсг-опосредованного ингибирования: подавление связывания Cas9 с ДНК, воздействие на связывание Cas9 с гРНК и блокирование активности Cas9 [46, 47]. В настоящее время для снижения нецелевых эффектов используют только AсгIIA2 и AсгIIA4. Конкурируя с сайтом PAM и/или другими участками ДНК за Cas9, белки Aсг блокируют расщепляющую активность Cas9, предотвращая чрезмерную и длительную активацию CRISPR/Cas9, и уменьшают таким образом нецелевые эффекты [48].

Решить проблемы, связанные с большим размером нуклеазы SpCas9, можно, используя ее ортологи меньшего размера, которые легче упаковать в вирусные векторы. Однако в этом случае необходимо учитывать архитектуру нуклеаз, разнообразие необходимых последовательностей PAM и относительную эффективность редактирования. Так, белки CjCas9 размером 2.95 т.п.н., SaCas9 (3.16 т.п.н.), CasX/Cas12e (2.94 т.п.н.), Cas14 (2.85 т.п.н.), Cas13d (2.79 т.п.н.) и даже более крупный белок Cas12a размером 3.6–3.9 т.п.н. можно упаковать в аденоассоциированные вирусные векторы [49–52].

Значительные успехи в применении системы CRISPR/Cas9 дали толчок к изучению других систем CRISPR/Cas для достижения потенциально более высокой эффективности редактирования и новых приложений.

К наиболее хорошо изученным, после Cas9, относятся белки Cas12 и Cas13 (рис. 3). Cas12 представляет собой эффекторный белок класса 2 типа V. У этого белка отсутствует каталитический домен HNH, вместо него для индукции расположенных в шахматном порядке “ступенчатых” разрывов с липкими концами за пределами сайта PAM используется домен RuvC-подобной нуклеазы, тогда как Cas9 имеет оба домена и создает разрывы с тупыми концами в области PAM. Показано, что Cas12 успешно распознает дцДНК (Cas12a/Cpf1) и может расщеплять оцДНК (Cas12f/Cas14). Кроме того, Cas12 регулируется единственной относительно компактной сгРНК и не нуждается в tracrРНК. Cas12 удобно использовать для мультиплексного редактирования, поскольку она обладает активностью РНКазы III (катализирующий процессинг пре-сгРНК в зрелые сгРНК), что

позволяет получать множество сгРНК из единого предшественника. В отличие от Cas9, нуклеаза Cas12 может расщеплять целевую последовательность (цис-расщепление), а также активировать свою способность расщеплять нецелевые последовательности (транс- или коллатеральное расщепление) [53]. Благодаря транс-расщепляющей активности, Cas12 можно использовать для обнаружения нуклеиновых кислот [54]. Белок Cas12 был перепрофилирован в нацеленный на ДНК-эндонуклеазу CRISPR трансрепортер (DETECTR), способный обнаруживать присутствие оцДНК. Эта технология применима для детекции ДНК-вирусов в клинических образцах [53]. Известно множество вариантов семейства Cas12: Cas12a (Cpf1), Cas12b, Cas12d (CasY), Cas12e (CasX), Cas12f (Cas14), Cas12j (CasF), Cas12h, Cas12i и Cas12c [33]. Эффективность наиболее известного из них, Cas12a, сравнима с эффективностью SpCas9, однако широкого применения в клетках млекопитающих этот вариант не нашел, в основном из-за строгой потребности в последовательности смежного мотива протоспейсера TTTV (PAM). Напротив, его ортолог Mb3Cas12a (*Moraxella bovoculi* AAX11\_00205) может эффективно редактировать геном мыши (последовательность TTV PAM). Более того, использование Mb3Cas12a, меченного мономерным стрептавидином, в сочетании с биотинилированной донорной ДНК-матрицей для HDR приводит в 40% случаев к рождению мышей F0 с нужной отредактированной последовательностью [55]. Недавно при реконструкции эволюции эндонуклеаз Cas9 и Cas12 как в прокариотических, так и в эукариотических клетках обнаружили семейство кодируемых транспозонами РНК-управляемых систем, названных OMEGA. Хотя эти программируемые нуклеазы, включая IscB, IsrB и TnpB, имеют большие перспективы для нацеливания и редактирования генома, их активность и биологическая значимость до сих пор не полностью охарактеризованы [56].

Cas13 представляет собой белок класса II и типа VI. Члены семейства CRISPR/Cas13 работают как двухкомпонентные системы, в которых сгРНК без участия tracrРНК образует комплекс с белком Cas13, состоящим из двух нуклеотидсвязывающих доменов HEPN. Каждый из этих доменов содержит сайт расщепления РНК. Для Cas13, также как для Cas9 и Cas12, требуются фланкирующие последовательности протоспейсеров (PFS), которые действуют аналогично PAM Cas9 и, следовательно, необходимы для расщепления РНК. Недавно обнаружили, что PFS нужны не для всех систем типа VI–Cas13 (например, для Cas13d (930 а.к.)), что обеспечивает большую гибкость для нацеливания и расщепления любой последовательности нуклеиновых кислот. Кроме того, Cas13 – единственная известная прокариотическая система CRISPR/Cas, нацеленная на одиночную РНК,

т.е. нуклеаза Cas13, обладая уникальной двойной РНКазной активностью, способна катализировать HEPN-независимый процессинг пре-сгРНК в зрелую сгРНК и производить HEPN-зависимое расщепление РНК [33, 51]. Cas13 также можно перепрофилировать в редактор оснований РНК. Так, каталитически неактивный ортолог Cas13 из *Prevotella* sp. (dPspCas13b) был объединен с дезаминазой ADAR2 для разработки редактора REPAIR, который может преобразовать А→I в РНК. В процессе транскрипции и сплайсинга инозин читается как гуанин, поэтому инструмент REPAIR может восстанавливать патогенные мутации G→A. Впоследствии получили систему RESCUE, которая не только сохраняет исходную активность дезаминазы А→I, но также осуществляет переход С→U. Эффективность редактирования и специфичность редакторов оснований РНК пока невысоки, поэтому необходимы дальнейшие исследования [57, 58]. Как уже упоминалось, существуют различные варианты Cas13 (Cas13a (1250 а.к.), Cas13b (1150 а.к.), Cas13d (930 а.к.)), одним из самых перспективных из которых считается Cas13d (CasRx). Cas13d имеет ряд преимуществ, особенно при работе с млекопитающими, в том числе:

1. небольшой размер, подходящий для упаковки в вирусные векторы;
2. низкая нейротоксичность;
3. двойная нуклеазная активность;
4. для расщепления РНК не требуется фланкирующая последовательность протоспейсера;
5. значимо низкая нецелевая активность, более высокая эффективность и специфичность в расщеплении целевой РНК.

Характеристики Cas13d делают его важным кандидатом для пространственно-временной инженерии транскриптома, обнаружения нуклеиновых кислот, мультиплексной регуляции генов, посттранскрипционного сайленсинга генов, альтернативного сплайсинга, отслеживания и эпигенетической регуляции РНК [51].

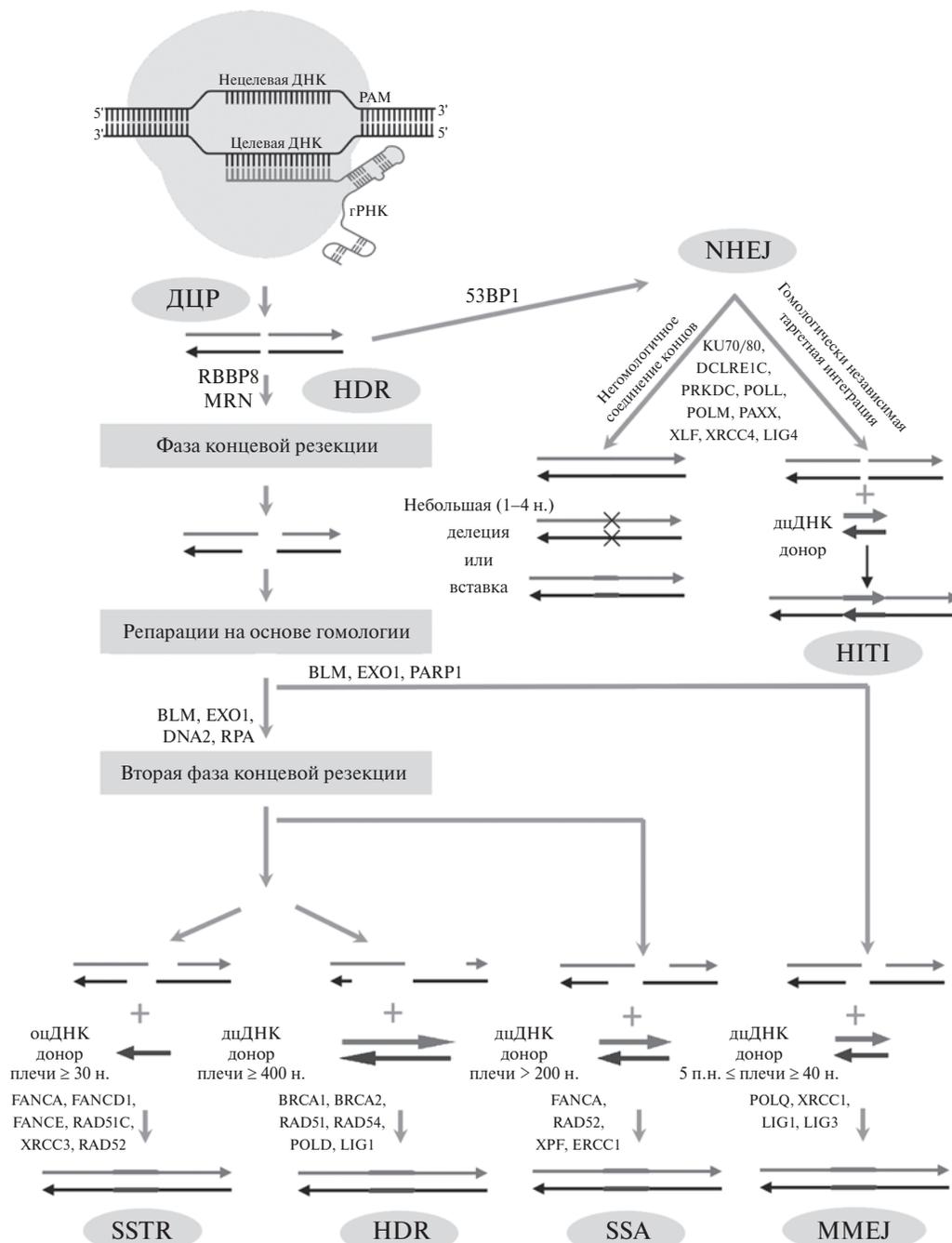
## РАЗРЫВЫ ДНК

Клеточная ДНК постоянно подвергается многочисленным спонтанным повреждениям (до  $10^5$  на клетку ежедневно), потенциально нарушающим стабильность генома, провоцирующим хромосомные перестройки, дестабилизирующим структуру и функции генов и приводящим к развитию патологических состояний [59–61]. Известны различные виды повреждений ДНК, включая модификации гетероциклических оснований и углеводородного остова, разрыв N-гликозидных связей, образование ковалентных связей между цепями ДНК, а также между ДНК и белками (например, гистонами). Наиболее радикальными среди них

являются ДЦР. В этом случае оба конца ДНК оказываются топологически разделенными, что может вызвать изменения кариотипа. Разрывы ДНК занимают центральное место в важнейших биологических процессах, включая мейоз [62] и V(D)J-рекомбинацию, при которой гены, кодирующие компоненты антител, рекомбинируют с образованием огромного разнообразия конечных продуктов [63]. Целенаправленная индукция сайт-специфических расщеплений ДНК стала краеугольным камнем технологий редактирования генома. Так, нуклеазы Cas вносят различные разрывы, например, Cas9 может производить ДЦР с тупыми концами, Cas12a (Cpf1) – с ‘липкими’, а pCas9 вносит парные или одиночные ОЦР [64].

В свою очередь, клетки развили способность восстанавливать повреждения ДНК, инициируя сложный сигнальный ответ, который модулирует клеточный цикл, изменяет хроматиновое окружение и рекрутирует факторы репарации [64]. Считается, что оба варианта разрывов ДНК репарируются с использованием различных путей, хотя на практике граница между восстановлением ОЦР и ДЦР размыта, так как ОЦР могут превращаться в ДЦР, например, в результате прохождения репликативной вилки или когда два соседних ОЦР спонтанно образуют ДЦР с липкими концами [65].

Клетки млекопитающих используют два основных механизма репарации разрывов ДНК: NHEJ и репарацию с помощью HDR. Восстановление ДЦР на протяжении всего клеточного цикла происходит преимущественно путем NHEJ. Этот путь может приводить к мутациям, чаще всего к небольшим вставкам и/или делециям (InDel) размером 1–4 п.н. Однако при наличии гомологичной матрицы в виде сестринской хроматиды или искусственно добавленных донорных ДНК-матриц в поздней S-фазе и в фазе G2 запускается путь HDR. HDR может пойти по нескольким направлениям в зависимости от типа и размера матричной ДНК. Так, в присутствии донорной дцДНК-матрицы со средними плечами гомологии >200 п.н. репарация может осуществляться путем одноцепочечного отжига по прямым повторам (SSA, single strand annealing). При более длинных плечах ( $\geq 400$  п.н.) восстановление пойдет по пути HDR, а если плечи гомологии равны примерно 30 н. и используется донорная оцДНК-матрица, то выбирается путь синтеззависимого отжига цепей (SDSA, synthesis-dependent strand annealing). Донорная дцДНК-матрица без плеч гомологии может быть встроена в место ДЦР ДНК путем гомологичной независимой целевой интеграции – этот путь обозначен как NITI (Homology-Independent Targeted Integration). Альтернативным путем репарации является путь микрогомологичного соединения концов (MMEJ), который имеет сходство как с NHEJ, так и с HDR. Подобно NHEJ,



**Рис. 4.** Репарация разрывов ДНК. Приведены пути репарации HDR, NHEJ, MMEJ, проходящие с участием белковых факторов и ферментов нуклеинового обмена (указаны рядом со стрелками, указывающими направление пути репарации), и соответствующие донорные ДНК-матрицы.

MMEJ активен в большинстве стадий клеточного цикла, однако, как и HDR, нуждается в первоначальной резекции концов ДЦР, при этом используются чрезвычайно короткие плечи гомологии – 5–40 п.н. (рис. 4) [66].

Понимание детерминант, которые направляют восстановление Cas-индуцированного разрезания ДНК по определенному пути репарации, имеет решающее значение для прогнозирования

и контроля результата редактирования генома; на эффективность этого процесса влияет также регуляция ключевых факторов конкурирующих путей репарации ДНК.

#### ИНГИБИТОРЫ NHEJ

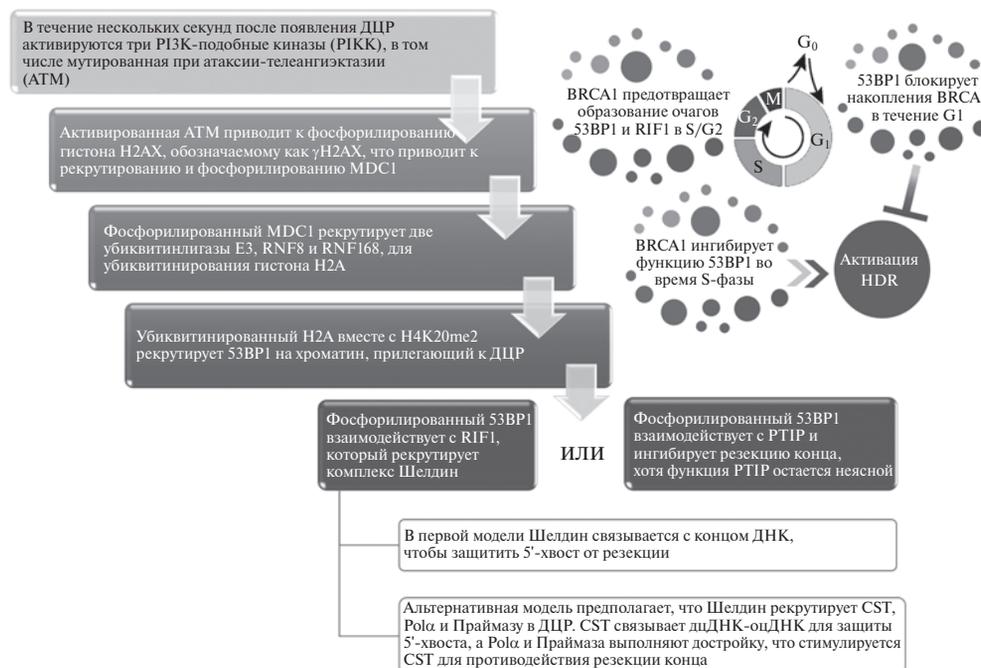
Выбор пути репарации ДЦР ДНК может иметь принципиальное значение при создании

трансгенных животных. Пути NHEJ и HDR — это конкурирующие пути репарации Cas9-индуцируемых ДЦР. При этом NHEJ является основным путем восстановления ДЦР в клетках млекопитающих, поскольку HDR зависит от фаз клеточного цикла и наличия эндогенных матриц. Если необходимо только нарушить функцию гена, “сбив” рамку считывания, то путь NHEJ считается оптимальным, однако, если требуется внести точные правки в целевые участки гена, то необходимо задействовать путь HDR. В этом случае нужно искусственно сместить баланс в сторону HDR. Самым очевидным решением становится подавление пути NHEJ. Известно несколько регуляторов ключевых факторов NHEJ. Сравним их эффективность, достоинства и недостатки.

В основе пути NHEJ лежит механизм прямого лигирования разъединенных концов ДНК независимо от степени гомологии между нуклеотидными последовательностями. На данном этапе подавить NHEJ можно с помощью регуляции каскада (комплекса) хроматинсвязывающего белка 53BP1 с белками-эффекторами RIF1, Шелдин и CST, участвующими в репарации ДЦР (53BP1–PTIP/RIF1–Шелдин–CST). Этот комплекс ингибирует резекцию концов ДНК в фазе G1 клеточного цикла (рис. 5). Ключевым фактором каскада является белок 53BP1; вычленив его из процесса репарации, можно подавить NHEJ-путь. Существуют два подхода к контролю 53BP1. Один из них заключается в подавлении его рекрутирования в ДЦР. Так, E3 убиквитин-протеин-лигаза RAD18 связывает

убиквитинированный гистон H2AK15Ub с большей аффинностью, чем белок 53BP1, тем самым ингибируя рекрутирование 53BP1 в ДЦР. Аналогичным образом белок DN1S, доминантно-негативный мутант 53BP1, конкурирует с эндогенным 53BP1 или вытесняет его, уменьшая привлечение эффекторов NHEJ [67–69]. Другой подход предполагает удаление 53BP1 из участка ДЦР, например, полиубиквитинированием 53BP1, опосредованным убиквитинлигазой E3 SPOP [70].

На следующем этапе пути NHEJ гетеродимеры белка Ku70/Ku80 (ДНК-связывающий компонент ДНК-зависимой протеинкиназы (ДНК-ПК, Ku) распознают концы ДЦР и связываются с ними, защищая свободные концы ДНК в месте разрыва и предотвращая резекцию 5'-конца. Комплекс Ku70/Ku80 рекрутирует и активирует ДНК-зависимую каталитическую субъединицу, активность которой необходима для реализации пути NHEJ. Несмотря на решающую роль субъединиц Ku на ранних этапах пути NHEJ, разработано ограниченное количество ингибиторов Ku70/80, а их эффективность дает неоднозначные результаты. Так, использование Ku-специфической малой интерферирующей РНК (siРНК) подавляло Ku70/Ku80 и в 2 раза увеличивало CRISPR/Cas9-опосредованную эффективность HDR в фибробластах свиной [71]. Многообещающие результаты получены с использованием низкомолекулярных ингибиторов Ku70/80. Показано, что использование арилалкиловых эфиров производных арилпиразолон-карбоновой кислоты уменьшало связывание Ku до



**Рис. 5.** Схема активации каскада 53BP1-PTIP-RIF1-Шелдин-CST. Стрелками обозначена последовательность этапов активации, в прямоугольных вставках приведено их описание. Вверху справа приведены сведения о роли белка BRCA1, вовлеченного в процесс HDR, на разных фазах клеточного цикла.

50% [72], а замена сложноэфирной связи на амидную приводила к четырехкратному увеличению способности подавлять связывание Ku-ДНК. В результате дальнейших модификаций этих соединений создан отдельный класс ингибиторов Ku-DBi – высокоэффективных и специфичных ингибиторов связывания Ku-ДНК, которые напрямую взаимодействовали с Ku и подавляли NHEJ [73]. Ингибирование Ku70 и Ku80 с помощью CRISPR значительно повышало эффективность HDR [74]. Так, активация системы CRISPRi (CRISPR interference) приводила к увеличению эффективности HDR в 1.85 раза при ингибировании Ku70 и в 1.58 раза в случае Ku80 [75]. Вместе с тем, в некоторых исследованиях показано, что подавление Ku не влияет на эффективность HDR [76], а также может приводить к снижению эффективности обеих путей репарации ДНК [77].

Интересно отметить, что ингибирование Ku не всегда может сместить равновесие в сторону HDR. Например, в качестве ингибитора гетеродимеров Ku70/Ku80 предложено низкомолекулярное соединение L – STL127705, идентифицированное с помощью скрининга библиотеки *in silico* [78]. Хотя нет отчетов о проверке влияния непосредственно этого ингибитора на эффективность HDR, изучение его 4-фторфенильного аналога STL127685 не влияло на снижение эффективности пути NHEJ или увеличение эффективности HDR [79].

Низкомолекулярные ингибиторы комплекса Ku70/Ku80 устраняют до 90% активности холофермента ДНК-ПК, содержащего помимо ДНК-связывающего комплекса Ku70/Ku80 еще и каталитическую субъединицу [72]. Каталитическая субъединица ДНК-ПК отвечает за распознавание ДЦР, фосфорилирование компонента X семейства гистонов H2A (H2AX) и рекрутирование/активацию механизма репарации для индукции пути NHEJ [80]. Обработка клеток HEK293T ингибиторами ДНК-ПК NU7441 (или KU57788) и KU-0060648 приводила к двукратному увеличению эффективности HDR [81], в эмбрионах рыбок *Danio rerio* соединение NU7441 увеличивало эффективность HDR в 13.4 раза [82]. Эти данные согласуются с результатами, полученными на плюрипотентных стволовых клетках, где еще один ингибитор ДНК-ПК, NU7026, повышал эффективность интеграции донорных олигонуклеотидов примерно в 1.6 раза [79]. Более мощным ингибитором ДНК-ПК считается M3814, его активность вызывает увеличение HDR в 4 раза и позволяет проводить редактирование нескольких локусов одновременно [83]. Комбинация нескольких ингибиторов может иметь синергический эффект. Например, ингибитор M3814 в сочетании с трихостатином А в 3 раза снижал эффективность NHEJ [84], а в сочетании с линеаризацией донорной плазмиды и транзиентного нокдауна ДНК-полимеразы  $\theta$  (метод ViPoD) почти

полностью подавлял NHEJ/ММЕJ и обеспечивал создание биаллельных knock-in в эмбриональных стволовых клетках мышей [13].

ДНК-ПК фосфорилирует множество субстратов, включая Artemis, XRCC4, ДНК-лигазу IV и XLF, что способствует удержанию концов ДЦР ДНК и облегчает рекрутирование ферментов процессинга и лигирования.

Использование ингибиторов ДНК-лигазы IV, играющей важную роль в лигировании концов ДЦР в процессе NHEJ, крайне актуально для подавления этого пути и смещения равновесия в сторону HDR. Наиболее хорошо охарактеризованным ингибитором ДНК-лигазы IV является низкомолекулярное производное дигидроптерина – соединение SCR7, механизм действия которого подробно рассмотрен в работах [85, 86]. Отмечено, что совместная инъекция компонентов CRISPR/Cas9 и SCR7 в зиготы повышает эффективность HDR-опосредованного инсерционного мутагеназа, что приводит к появлению мышей со вставками в нескольких локусах. Применение SCR7 смещает баланс в сторону HDR в 10 раз в эмбрионах мышей и на 46% у крыс, а также в широком диапазоне клеточных линий, включая клетки рака легкого, молочной железы, толстой кишки, меланомы, и первичные фибробласты. В случае вставок длинных фрагментов ДНК использование SCR7 повышает результативность до 13 раз, а с короткими вставками – до 19 раз [76]. Комбинация SCR7 с другими факторами может значительно повысить эффективность пути HDR, например, при использовании Rad52 до 40% [87], а при совместной инъекции с компонентами CRISPR/Cas9 в зиготы можно получить вставки в нескольких локусах [88]. Показано, что применение SCR7 увеличивает результативность стратегии повышения эффективности HDR до 30 раз [89]. Однако вопрос о роли SCR7 в повышении эффективности HDR за счет снижения эффективности NHEJ остается спорным. В целом, активность SCR7, по-видимому, является ткане- и организмспецифичной. Об этом свидетельствуют сообщения о слабом влиянии SCR7 на подавление NHEJ или усиление HDR в локусах ROSA26-like (RLL) генома кролика [90], MALAT1 клеток HEK293T [91], pROSA26 в эмбриональных фибробластах свиньи [92], при обработке линии эмбриональных стволовых клеток H1 [93], клеток U2OS DR-GFP [94], HSPC [95] и линии клеток миелогенного лейкоза K562 [96], при создании CRISPR/Cas12a-опосредованного нокаута в плюрипотентных стволовых клетках человека [97], а также в эмбрионах *D. rerio* [82]. Представлены и модифицированные варианты SCR7. Так, более стабильная циклическая форма SCR7 может быть дополнительно окислена до SCR7-пиразина с сохранением способности подавлять NHEJ [98]. Показано, что SCR7-пиразин увеличивает эффективность

генетического редактирования на 50% в клетках MCF7 [99], но в создании нокаута в *Xenopus* это соединение было менее эффективным [100], что также свидетельствует о ткане- и видоспецифичности SCR7. Другая модификация SCR7, идентифицированная как SCR130, индуцирует апоптоз и обладает противоопухолевой активностью [101]. Блокировать активность ДНК-лигазы IV можно также путем shPНК-опосредованного (shPНК – образующая шпилечные структуры малая PНК) сайленсинга генов и совместной экспрессии системы Cas9 и белков E1B55K и E4orf6 аденовируса 4 вплоть до повышения эффективности HDR до 8 раз и практически полного подавления NHEJ в клеточных линиях человека и мыши [74].

Когда концы ДЦР невозможно напрямую лигировать с помощью комплексов XLF-XRCC4 и ДНК-лигазы IV (например, из-за выступающих фрагментов нескольких нуклеотидов), такие нуклеазы, как Artemis, полинуклеотидкиназа-3'-фосфатаза (PNKP) и ДНК-полимеразы (Pol $\mu$  и Pol $\lambda$ ) могут модифицировать места разрывов оцДНК-дцДНК с образованием совместимых тупых концов [85, 102].

Если не задействовать описанные механизмы подавления NHEJ, то на заключительном этапе этого пути лигирование выполняет комплекс XRCC4-ДНК-лигаза IV-XLF. Основания ДНК случайным образом добавляются и удаляются ДНК-полимеразами и нуклеазами, что приводит к образованию вставок, небольших по сравнению с исходной геномной матрицей, которая составляет основу подверженного ошибкам редактирования на основе NHEJ.

Повысить эффективность направленного редактирования генома можно не только с помощью ингибирования NHEJ, но и путем прямой активации HDR. Последнее может быть более безопасным для клеток, так как не приводит к их гибели из-за нерепарируемых разрывов ДНК.

## РЕГУЛЯЦИЯ КЛЮЧЕВЫХ ФАКТОРОВ ПУТИ HDR

Репарация ДНК по пути HDR протекает значительно медленнее, чем по пути NHEJ. Этот процесс продолжается в течение 7 ч и более и приводит к резекции не менее  $10^3$  нуклеотидов с 5'-конца от места разрыва с образованием 3'-выступающего конца ДНК. Концевая резекция – это важная часть механизмов MMEJ и HDR. Наличие участка оцДНК необходимо для гибридизации с комплементарной последовательностью матрицы для HDR, чтобы при удлинении этого конца ДНК-полимеразой можно было точно восстановить последовательность, утраченную, возможно, в месте разрыва.

Большинство вариантов пути HDR, таких как канонический путь, путь синтеззависимого отжига

цепей (SDSA), путь индуцированного разрыва репликации (BIR) и одноцепочечного отжига по прямым повторам (SSA), протекают при наличии матрицы, и ДНК-трансфераза RAD51 опосредует поиск донорной ДНК-матрицы в процессе HDR. Так, сборка нуклеопротеиновых филаментов RAD51-оцДНК необходима для реализации путей HDR, SDSA и некоторых BIR, но путь SSA не зависит от RAD51 [103].

Недавно было показано, что компоненты пути анемии Фанкони (Fanconi Anemia, FA) – пути репарации межцепочечных сшивок в ДНК, участвуют и в репарации ДЦР ДНК. Анемия Фанкони – генетическое заболевание, вызываемое мутациями в генах, регулирующих зависимое от репликации удаление межцепочечных сшивок ДНК, характеризуется недостаточностью костного мозга и предрасположенностью к раку. Сигнальный путь FA координирует сложный механизм, в котором задействованы элементы трех классических путей репарации ДНК: HDR, эксцизионной репарации нуклеотидов и транслезионного синтеза ДНК. Для этого в пути FA используется уникальный ядерно-белковый комплекс, который убиквитинирует белки FANCD2 и FANCI, приводя к образованию структур репарации ДНК. В сигнальном пути FA задействованы 19 белков – от FANCA до FANCT, а также множество ассоциированных белков. Считают, что HDR является одним из этапов пути FA, в котором реализуется SSTR-путь рекомбинации с использованием олигонуклеотидных оцДНК-матриц (Single-Stranded Templated Repair). Белки FA – FANCD1 (BRCA2), FANCR (RAD51) и FANCD2, участвуют в инициации репарации ДЦР ДНК по механизму HDR. Изучение взаимодействия между FA и путями репарации ДЦР приводит к появлению все большего числа доказательств того, что FA не только участвует в репарации межцепочечных поперечных сшивок дцДНК, но также играет роль в ДЦР-опосредованном редактировании генома [104].

Как уже отмечено, основным шагом, который определяет выбор HDR в качестве пути репарации ДЦР, является 5'-3'-резекция конца ДНК с образованием 3'-выступающего одноцепочечного фрагмента. Этот процесс инициируется белковым комплексом MRN (MRX у дрожжей), состоящим из белков MRE11, RAD50 и NBS1, который имеет решающее значение для короткой концевой резекции [105]. Одной из стратегий, применяемых для смещения равновесия репарации разрывов ДНК в сторону HDR, стала разработка химерной конструкции Cas9 для привлечения комплекса MRN к участку ДЦР. В этой системе белок SpCas9 слит с доменом из 126 а.к. щелочной нуклеазы HSV-1 (UL12), который и рекрутирует комплекс MRN [106].

Неоднозначная ситуация наблюдается и с участием в выборе пути репарации ДНК белка PARP1, представителя семейства поли(ADP-рибозо)полимераз, участвующего в репарации повреждений ДНК и ремоделировании хроматина. С одной стороны, этот белок может способствовать привлечению комплекса MRN, что стимулирует пути MMEJ, SSA или HDR [107], однако PARP1 быстро рекрутируется в новые ДЦР и может ингибировать резекию и привлекать Ku-комплекс, способствуя реализации пути NHEJ [108].

Установлено, что фосфорилированный фактор резекии CtIP и комплекс BRCA1-BARD1 стимулируют процесс резекии с помощью MRN, в ходе которого субединица MRN NBS1 генерирует короткие выступающие концы [109, 110]. На данном этапе сместить равновесие в пользу HDR можно, используя нуклеазу Cas9, гибридованную с HE, N-концевым фрагментом CtIP, за счет рекрутирования эндогенного CtIP в месте расщепления. Этот эффект показан на клеточных линиях человека, iPSC-клетках и зиготах крыс. Его результативность зависела от сrПНК и варьировала в зависимости от локуса [111].

Короткие одноцепочечные фрагменты, генерируемые MRN, служат местом связывания таких ферментов, как экзонуклеаза 1 (EXO1), хеликаза Блума (BLM) и нуклеаза DNA2 [112]. EXO1 генерирует длинные 3'-одноцепочечные последовательности ДНК за счет своей процессирующей 5'-экзонуклеазной активности. Стимулировать резекию с помощью EXO1 может гетеротримерный комплекс SOSS1, связывающий оцДНК [113]. Сам комплекс MRN также рекрутирует и усиливает процессивность EXO1, а хеликаза BLM увеличивает средство EXO1 к местам разрыва ДНК [112]. Отмечено, что репликативный белок A (RPA) может как стимулировать, так и ингибировать активность EXO1 в зависимости от того, фосфорилирован он или нет [114]. Комплекс BLM/DNA2 обладает как 5'-, так и 3'-эндонуклеазной активностью: хеликаза BLM разделяет нити ДНК, позволяя нуклеазе DNA2 расщеплять оцДНК [112]. Хеликазная активность BLM увеличивается при взаимодействии с фактором резекии CtIP, который усиливает также DNA2-опосредованное расщепление ДНК [115]. RPA ингибирует 3'-эндонуклеазную активность DNA2, одновременно стимулируя ее 5'-эндонуклеазную активность, чтобы обеспечить 5'-3'-направление резекии [114]. Как уже упоминалось, регуляторный механизм 53BP1-PTIP/RIF1-Шелдин-CST ингибирует дальнюю резекию концов (Long Range Resection), причем хроматинсвязывающий белок 53BP1 блокирует привлечение белка BRCA1 к ДЦР во время G1-фазы, в то время как белок BRCA1 предотвращает образование очагов 53BP1 и RIF1 в ДЦР во время фазы S/G2 [116]. Антагонистические отношения между

BRCA1-BARD1 (BARD1 – белок, ассоциированный с BRCA1, BRCA1-associated RING domain protein 1) и 53BP1 могут быть связаны с двумя состояниями посттрансляционной модификации гистонов H2A и H4: убиквитинированием лизина-15 гистона H2A и метилированием лизина-20 гистона H4, что влияет на выбор пути репарации ДЦР [117]. Также RAD18, а именно его улучшенный вариант e18, подавляя локализацию 53BP1 в ДЦР, стимулирует CRISPR-опосредованную HDR, тем самым повышая эффективность точного редактирования генома [118].

Поскольку образующиеся одноцепочечные фрагменты ДНК очень нестабильны, они связываются с белком RPA, который защищает оцДНК от нуклеазной деградации и предотвращает образование вторичных структур. С помощью “посредников” рекомбинации, таких как белки BRCA1, BRCA2, а также партнера BRCA2 – белка PALB2, RPA вытесняется белком RAD51, который затем образует протяженные спиральные нуклеопротеиновые филаменты на оцДНК путем блокирования гидролиза АТФ. RPA представляет собой барьер для сборки филаментов RAD51. В G2/S-фазе циклин-зависимая киназа 1 (CDK1) фосфорилирует RAD51, что усиливает способность RAD51 конкурировать с RPA за оцДНК [119]. Установлено, что небольшая молекула RS-1, идентифицированная в результате скрининга библиотеки из 10 000 соединений, может усиливать ДНК-связывающую активность RAD51, что в несколько раз увеличивает эффективность Cas9-опосредованной HDR в клетках HEK293A, плюрипотентных стволовых клетках человека, в эмбрионах кроликов и *D. rerio*. После формирования нуклеопротеинового филамента RAD51-оцДНК происходит поиск комплементарной ДНК-матрицы для HDR. Инвазия 3'-выступающего конца ДНК в матрицу приводит к образованию D-петли, за которым следует удлинение 3'-выступающего конца ДНК с помощью ДНК-полимеразы  $\delta$ . Способствовать инвазии ДНК может сверхэкспрессия дрожжевого RAD52 ( $\gamma$ RAD52), но не его человеческого аналога, HDRAD52. Показано, что применение слитого белка  $\gamma$ RAD52-Cas9 приводит к увеличению эффективности HDR в клетках разных организмов. Предполагается, что после образования D-петли процесс репарации может пойти тремя путями, в зависимости от того, образуются одна или две структуры Холлидея. Затем действует группа ферментов, называемых резольвазами, которые завершают процесс репарации и восстанавливают хромосому до ее неповрежденного состояния [68].

Скрининг около 4000 малых молекул выявил две небольшие молекулы, L755507 и брэфельдин А, способных повысить эффективность HDR. Так, использование соединения L755507, агониста  $\beta$ 3-адренергических рецепторов, приводило к увеличению

эффективности вставки GFP в 3 раза по сравнению с контрольными клетками. Применение брефельдина А, ингибитора внутриклеточного транспорта белка из эндоплазматического ретикулума в аппарат Гольджи, также приводило к повышению эффективности встраивания в 2 раза [120].

Следует отметить, что активность некоторых ключевых факторов, участвующих в HDR в клетках дрожжей и млекопитающих, регулируется с помощью высококонсервативного белка SUMO (небольшой белок примерно из 100 а.к.), способного регулировать активность того белка, к которому он присоединяется [121]. Как правило, сумоилирование положительно воздействует на концевую резекцию. Например, SUMO оказывает сильное влияние на функционирование CtIP [122] и способствует стабильности белка EXO1 в клетках млекопитающих [123]. Следовательно, можно предположить, что сумоилирование этих ключевых факторов может обеспечить более высокую эффективность HDR.

Из-за неравномерного распределения активности NHEJ и HDR по клеточному циклу, когда NHEJ активен на протяжении всего цикла, а HDR ограничена лишь поздней фазой G2 и S, когда происходит репликация ДНК перед делением клетки, ингибиторы клеточного цикла используют для повышения эффективности HDR, останавливая или задерживая клетки-мишени в фазах S или G2.

### КОНТРОЛЬ КЛЕТОЧНОГО ЦИКЛА

Поскольку репарация разрывов в ДНК по пути HDR ограничена фазами S и G2 клеточного цикла, синхронизация клеток в этих фазах с помощью химических соединений может значительно повысить ее эффективность. Такими соединениями стали блокаторы G1/S: афидиколин, гидроксимочевина, мимозин и тимидин, повышающие вероятность восстановления ДНК по пути HDR в неонатальных фибробластах [124]. Нокодазол и АВТ-751 останавливают клетки в фазе G2/M, ингибируя полимеризацию микротрубочек. Нокодазол способствует увеличению эффективности HDR в 1.38 раза в клетках HEK293T [124], а АВТ-751 увеличивает в 6 раз эффективность редактирования генома в линиях стволовых клеток человека [93]. Некоторые блокаторы M/G1 не всегда показывают высокие результаты в стимуляции HDR, как например, ловастатин в неонатальных фибробластах [124]. Недавно было показано, что низкомолекулярное соединение XL413, ингибитор киназы CDC7, участвующей в инициации репликации ДНК, замедляет S-фазу, удлиняя фазы S/G2/M, и увеличивает эффективность HDR до 3.5 раз [96]. Тиазольное производное хинона, соединение RO-3306, может временно останавливать переход от фазы G2 к фазе M путем ингибирования функции циклин-зависимой киназы (CDK1), увеличивая количество

клеток в фазах S и G2 [125]. Соединение RO-3306, ингибитор CDK1, применяемое в сочетании с белком Cas9, нуклеазная активность которого снижена в фазе G1, может улучшить соотношение путей HDR/NHEJ в 20 раз в пользу HDR [126]. Вместе с тем, эффекты блокаторов клеточного цикла зависят от типа клеток-мишеней и могут иметь ограничения для использования *in vivo* из-за потенциальной токсичности. Зависимая от фазы клеточного цикла экспрессия Cas9 может быть дополнительным подходом к синхронизации образования ДЦР с фазами клеточного цикла. Стратегия контроля экспрессии Cas9 в фазах S и G2 основана на создании комплекса Cas9-hGem (Cas9-геминин). Слияние первых 110 а.к. геминина с Cas9 приводит к протеолитической деструкции такого гибридного белка, вызванной APC/CDH1-опосредованным убиквитинированием в поздних фазах M и G1, и в итоге увеличивает эффективность HDR до 87% в клетках HEK293T [91]. Эффективность этого подхода подтверждена на плюрипотентных стволовых клетках человека и фибробластах свиньи, что делает его многообещающим инструментом для повышения эффективности HDR [127, 128]. Кроме того, существуют не прямые методы замедления клеточного цикла. Так, например, инкубация эмбрионов *D. rerio* на льду увеличивает результативность HDR в 1.5 раза [82].

### СТАТУС ХРОМАТИНА

Статус хроматина, его компоненты и структура значительно влияют на выбор пути репарации ДЦР, однако детальные механизмы этого процесса остаются неясными. Система CRISPR-Cas функционирует в различных хромосомных контекстах, таких как эухроматин или гетерохроматин.

Фундаментальной единицей хроматина является нуклеосома, состоящая из двух копий каждого корового гистона (H3, H4, H2A и H2B), обернутых 147 парами оснований ДНК. Неструктурированные концевые области коровых гистонов подвергаются множеству различных посттрансляционных модификаций, среди которых наиболее хорошо изучены фосфорилирование, метилирование, ацетилирование и убиквитинирование. Посттрансляционные модификации гистонов, а также метилирование ДНК помогают сформировать эухроматин и гетерохроматин. Эухроматин соответствует открытому состоянию хроматина, он связан с активной транскрипцией и обогащен метилированными (например, H3K4me3, H3K36me3) и ацетилированными гистонами. Напротив, гетерохроматин более компактен, менее доступен для аппарата транскрипции и обогащен репрессивными гистоновыми посттрансляционными модификациями и, таким образом, транскрипционно неактивен. Гетерохроматин разделяют на факультативный, конститутивный и центромерный. Последний включает

вариант гистона H3, СENP-A, который эпигенетически определяет расположение центромер и перемежается с активными гистоновыми посттрансляционными модификациями H3K4me3, H3K36me3, но лишен репрессивного H3K9me3, характерного для конститутивного гетерохроматина [129, 130].

Установлено, что триметилирование остатка лизина-36 гистона H3 (H3K36me3) играет критическую роль в репарации ДНК по пути HDR в эухроматине [130]. Триметилированный гистон H3K36me3 обнаружен также в участках, которые репарируются по пути NHEJ [131]. Установлено, что содержание аналогичного гистона, содержащего диметилированный лизин-36 (H3K36me2), снижено в ДЦР, которые репарируются по пути HDR, и характерно для ДЦР, которые репарируются по пути NHEJ [130]. Модифицированный гистон H3, содержащий диметилированный остаток лизина-4 (H3K4me2) в центромерах, способствует репарации по пути HDR в фазе G1 клеточного цикла [132], триметилирование лизина-9 гистона H3 (H3K9me3) также способствует восстановлению разрывов в ДНК по пути HDR и уменьшает эффективность репарации по пути NHEJ.

В среднем для гетерохроматина более характерны пути репарации HDR и MMEJ. Так, использование системы KRAV (Küppel-associated box), которая рекрутирует компоненты ремоделирования хроматина и нуклеосом, создавая локальные гетерохроматиновые области в геноме, в конечном счете приводит к смещению баланса от NHEJ к HDR [133, 134]. Показано подобное смещение равновесия и от NHEJ к MMEJ [129]. Такое распределение выглядит логичным, так как NHEJ по умолчанию (особенно на этапе G1) более активен, чем MMEJ или HDR и, следовательно, в эухроматине может реализоваться в основном NHEJ. В то же время ремоделирование участка разрыва ДНК в гетерохроматине может быть медленным, что дает время для активации MMEJ и HDR. Использование ингибиторов гистоновых деацетилаз HDAC1 или HDAC2, которые поддерживают компактную природу гетерохроматина, показало значительное увеличение скорости как NHEJ, так и HDR – примерно в 2–3 раза [135]. Так, применение трихостатина А, ингибитора HDAC класса I/II, увеличивало примерно в 2.2 раза эффективность редактирования генов в экспрессирующих pCas9 клетках, а применение вальпроевой кислоты, также ингибитора HDAC класса I/II, в сочетании со сверхэкспрессией RAD51 значительно повысило эффективность биаллельной HDR в клетках ES/iPS человека [136]. Следует отметить, что в более раннем исследовании не зафиксировано очевидных различий в соотношении HDR/NHEJ в различных состояниях хроматина [137].

Таким образом, выбор пути репарации ДЦР контролируется статусом метилирования H3K36,

а посттрансляционные модификации H3K4me2 и H3K9me3 способствуют реализации HDR [138]. На сегодняшний день вопрос о точном механизме того, как структуры хроматина влияют на выбор пути репарации Cas9-опосредованного ДЦР, остается открытым. В отношении репарации опосредованных Cas9 разрывов ДНК существует обобщенное мнение, что структура хроматина влияет на активность Cas9 как *in vivo*, так и *in vitro*, при этом гетерохроматин препятствует CRISPR-Cas-опосредованному редактированию генома [139].

## ДНК-МАТРИЦЫ ДЛЯ ГОМОЛОГИЧНОЙ РЕКОМБИНАЦИИ

Выбор правильного варианта ДНК-матрицы для репарации по пути HDR может значительно повлиять на результативность редактирования генома. Существуют три основных варианта донорных матриц: дцДНК, оцДНК, а также линейная оцДНК аденоассоциированного вируса.

Матрицы дцДНК удобно получать, они имеют низкую стоимость и поэтому широко используются. HDR с использованием в качестве доноров дцДНК менее эффективна, чем с оцДНК-матрицами, однако дцДНК позволяет создавать вставки больших фрагментов (до 7.1 т.п.н.) [66]. Плечи гомологии, необходимые для дцДНК-матриц, обычно относительно длинные (0.5–2 т.п.н.), но успешно используются и плечи гомологии <100 п.н. [140]. Размер плеча является ограничивающим фактором, поскольку эффективность knock-in HDR с введением в целевое место генома заданной последовательности) положительно коррелирует с длиной плеча гомологии дцДНК и отрицательно коррелирует с размером вставки [76].

Матрицы оцДНК имеют короткие плечи гомологии (70–100 нуклеотидов), их легко синтезировать с введением различных модификаций для повышения эффективности трансформации и стабильности *in vivo*. Общий уменьшенный размер оцДНК делает ее более подходящей для точечного редактирования со стратегией короткой вставки (<100 нуклеотидов) вместо целых генов [66].

Один из вариантов оцДНК – короткий одноцепочечный олигонуклеотид (оцОНК) длиной около 200 нуклеотидов. оцОНК содержит плечо гомологии ≥30 нуклеотидов как на 5'-, так и на 3'-конце, где более длинное плечо гомологии может повысить эффективность редактирования, но при этом проявляет более выраженную цитотоксичность. Ограниченная длина оцОНК не позволяет делать крупные вставки, хотя использование нескольких оцОНК с перекрывающимися областями увеличивает возможность вставок крупных фрагментов ДНК [66]. Повышение результативности редактирования при использовании оцОНК достигнуто путем фосфориоатной модификации, которая стабилизировала

матрицу за счет уменьшения вне- и внутриклеточной деградации нуклеазами и увеличивала ее способность проникать через плазматическую мембрану клетки-мишени [141]. В настоящее время оцОНК стали рутинным инструментом редактирования как *in vitro*, так и *in vivo*, особенно в экспериментах по изменению нескольких генов [142].

Для расширения использования оцОНК разработана система Easi-CRISPR, в которой в качестве донорных матриц используется длинная оцДНК [143], которая позволяет получать вставки более крупных фрагментов, чем оцОНК, заменять гены в выбранных местах генома и получать условные аллели. Как и в случае оцОНК, расстояние между участком целевой модификации и сайтом Cas-расщепления должно быть как можно меньше. Средняя длина такой ДНК составляет 0.2–2 т.п.н. (максимальная длина 5 т.п.н., после чего ДНК склонны к спонтанному разрывам). При этом используется короткое плечо гомологии, обычно 55–329 нуклеотидов на каждом конце. Длинная оцДНК обладает рядом недостатков: она более токсична, чем оцОНК, а ее получение дороже и сложнее, кроме того, репарация с помощью длинной оцДНК приводит к большей частоте ошибок, чем в случае дцДНК [66, 119].

В дополнение к оцОНК и длинным оцДНК в качестве матриц для HDR используется линейная оцДНК аденоассоциированного вируса (ААВ), упакованная в капсид после удаления всех вирусных генов. Сообщается, что оцДНК-ААВ-матрица обладает высокой эффективностью точечных мутаций и вставок больших фрагментов ДНК,

поскольку вирус обладает естественной способностью стимулировать HDR [144, 145]. В целом, длина плеча гомологии оцДНК-ААВ-матрицы составляет 400–800 нуклеотидов. Средняя длина вставки при использовании оцДНК-ААВ матрицы составляет 771–3300 нуклеотидов и не может превышать 3800. Однако с помощью CRISPR-Cas в HSPC и Т-клетки можно вставить ДНК длиной 5700 оснований [146].

HDR с использованием донорной дцДНК-матрицы проходит в основном по пути HDR, тогда как при использовании оцДНК матрицы репарация проходит по пути одноцепочечного отжига (SSA) и синтез-зависимого отжига цепи (SDSA) [103]. Однако донорные дцДНК-матрицы как плазмидные, так и линейные, задействованы в системах репарации HMEJ, Tild и PITCh в качестве альтернативы HDR. В качестве донорных ДНК-матриц в HMEJ использовали суперскрученные плазмиды, несущие сайты-мишени синтетической гРНК, и два плеча гомологии (600–900 п.н.) [147], которые в стратегии Tild были заменены линейными дцДНК с двумя плечами гомологии по 800 п.н. [148]. При этом оба типа донорных ДНК-матриц могли эффективно и точно встраивать экзогенную ДНК в сайт-мишень. В системе PITCh также используются сверхскрученные плазмиды, несущие сайты-мишени синтетической гРНК, но с микрогомологичными последовательностями в качестве плеч гомологии (5–25 п.н.). В итоге донорная последовательность встраивалась в геном по MMEJ-пути [149], что позволило работать системе во время G1/ранних S-фаз клеточного цикла в отличие от HDR. В целом

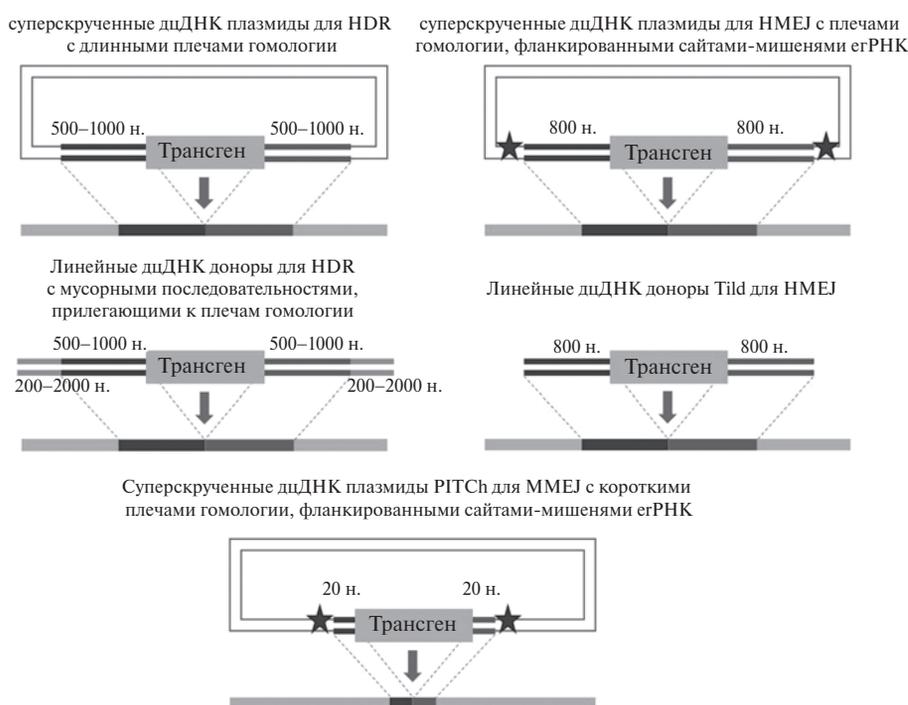


Рис. 6. Схемы гомологичной репарации (HDR) ДЦР ДНК с использованием в качестве донорных матриц дцДНК.

разработка этих стратегий позволила в разы увеличить эффективность HDR в следующем порядке Tild > HMEJ > PITCh ≥ HDR, где разрыв между Tild и HDR достигал шести порядков (рис. 6) [148].

Также показано, что встроить донорную последовательность можно и с использованием пути NHEJ. В этом случае ДЦР вводятся как в целевой геном, так и в донорную матрицу. Сначала донорную плазмиду линейризуют и разрезают (при этом донорная матрица с двойным разрезом менее эффективна, чем с единичным, так как в первом случае образуются два фрагмента ДНК, конкурирующих за геномную интеграцию), а затем напрямую лигируют с концами разрезанной геномной ДНК по пути NHEJ. NHEJ эффективно повторно лигирует концы ДНК без ошибок и не требует областей гомологии для точной вставки. Однако донорная матрица может быть вставлена в любой ориентации относительно концов разорванной геномной ДНК, и нецелевые ДЦР также могут привести к непреднамеренной вставке в случайные участки генома. Чтобы обойти эту проблему, в донорную плазмиду вводили короткую гомологичную последовательность – “приманку”. Оказалось, что вставка не зависит от гомологии между локусом-мишенью и “приманкой” в донорной плазмиде [150, 151]. Аналогичный результат можно получить с использованием белка Cas12a (Cpf1), который в отличие от Cas9 генерирует ДЦР с липкими концами, потенциально способствующими точному редактированию генома [152]. Недавно обнаружено, что матрица репарации короткой дцДНК, содержащая липкий конец, соответствующий одному из концов Cpf1-опосредованного ДЦР, и гомологичное плечо, разделяющее гомологию с геномной областью, примыкающей к другому концу ДЦР, обеспечивает точную репарацию и вводит желаемую нуклеотидную замену. Эта стратегия задействует механизмы как HDR, так и MMEJ, и названа HDR с помощью лигирования [153].

Таким образом, оптимальные результаты редактирования могут быть получены путем выбора подходящей донорной ДНК-матрицы в соответствии с экспериментальными требованиями. Однако важны также формат предоставления матрицы донорной ДНК, ее доступность и концентрация в ядре. Обычно матрица добавляется в виде очищенной или “голой” ДНК, но в нормальной клеточной среде спираль ДНК обернута вокруг кластера белков – гистонов. Показано, что сконструированные ДНК-матрицы в форме, имитирующей хроматин, наряду со сниженной цитотоксичностью демонстрируют в 2.3–7.4 раза более высокую эффективность HDR по сравнению с развернутой ДНК [154]. Для сближения компонентов CRISPR с донорной ДНК-матрицей внутри клеток разработано несколько систем. Во-первых, модульная стратегия РНК-аптамер-стрептавидин (S1mplex), состоящая из еРНК, дополненной

стрептавидинсвязывающим РНК-аптамером, самого стрептавидина и белка Cas, созданная для формирования комплекса рибонуклеопротеинов CRISPR-Cas9 с биотинилированной донорной оцДНК-матрицей (bio-оцДНК) [155]. Во-вторых, получен вариант Cas9, слитый с авидином через гибкий линкер и соединенный мостиком с модифицированной биотином оцДНК (bio-оцДНК) для увеличения сродства [156], а также с использованием SNAP-метки сконструирован комплекс рибонуклеопротеинов CRISPR-Cas9 и bio-оцДНК [157].

## ЗАКЛЮЧЕНИЕ

Эволюция развития технологии CRISPR/Cas привела к получению гибкой, простой в использовании молекулярной платформы для редактирования генома различных организмов. За последние годы разработаны разнообразные стратегии CRISPR-опосредованного редактирования генома, включая новые эффективные подходы к редактированию генома CRISPR-Cas9 с повышенной эффективностью HDR для встраивания в геномы протяженных фрагментов ДНК. Это позволяет добиться значительных достижений в фундаментальных и биомедицинских исследованиях и расширяет понимание биологии живых организмов и патогенетических основ болезней человека.

Работа поддержана Российским научным фондом (грант № 21-64-00006) и частично профинансирована из бюджета Института функциональной геномики МГУ имени М.В. Ломоносова.

Настоящая статья не содержит описания исследований, выполненных с использованием человека и животных в качестве объектов.

Авторы заявляют об отсутствии конфликта интересов, связанных с публикацией настоящей статьи.

## СПИСОК ЛИТЕРАТУРЫ

1. 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., Siksny V., Terns M.P., Venclovas Č., White M.F., Yakunin A.F., Yan W., Koonin E.V. (2020) Evolutionary classification of CRISPR-Cas systems: a burst of class 2 and derived variants. *Nat. Rev. Microbiol.* **18**, 67–83. doi: 10.1038/s41579-019-0299-x
2. Liu Z., Dong H., Cui Y., Cong L., Zhang D. (2020) Application of different types of CRISPR/Cas-based systems in bacteria. *Microb. Cell Fact.* **19**, 1–14. doi: 10.1186/s12934-020-01431-z
3. Jinek M., Chylinski K., Fonfara I., Hauer M., Doudna J.A., Charpentier E. (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*. **337**, 816–821. doi: 10.1126/science.1225829

4. Filippova J., Matveeva A., Zhuravlev E., Stepanov G. (2019) Guide RNA modification as a way to improve CRISPR/Cas9-based genome-editing systems. *Biochimie*. **167**, 49–60. doi: 10.1016/j.biochi.2019.09.003
5. Hendel A., Bak R.O., Clark J.T., Kennedy A.B., Ryan D.E., Roy S., Steinfeld I., Lunstad B.D., Kaiser R.J., Wilkens A.B., Bacchetta R., Tsalenko A., Dellinger D., Bruhn L., Porteus M.H. (2015) Chemically modified guide RNAs enhance CRISPR-Cas genome editing in human primary cells. *Nat. Biotechnol.* **33**, 985–989. doi: 10.1038/nbt.3290
6. Yin H., Song C.Q., Suresh S., Wu Q., Walsh S., Rhym L.H., Mintzer E., Bolukbasi M.F., Zhu L.J., Kauffman K., Mou H., Oberholzer A., Ding J., Kwan S.Y., Bogorad R.L., Zatselin T., Koteliensky V., Wolfe S.A., Xue W., Langer R., Anderson D.G. (2017) Structure-guided chemical modification of guide RNA enables potent non-viral *in vivo* genome editing. *Nat. Biotechnol.* **35**, 1179–1187. doi: 10.1038/nbt.4005
7. Fu Y., Sander J.D., Reyon D., Cascio V.M., Joung J.K. (2014) Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat. Biotechnol.* **32**, 279–284. doi: 10.1038/nbt.2808
8. Liu G., Zhang Y., Zhang T. (2020) Computational approaches for effective CRISPR guide RNA design and evaluation. *Comput. Struct. Biotechnol. J.* **18**, 35–44. doi: 10.1016/j.csbj.2019.11.006
9. Gurumurthy C.B., Lloyd K.C.K. (2019) Generating mouse models for biomedical research: technological advances. *Dis. Model. Mechan.* **12**, dmm029462. doi:10.1242/dmm.029462
10. Li F., Cowley D.O., Banner D., Holle E., Zhang L., Su L. (2014) Efficient genetic manipulation of the NOD-Rag1-/-IL2RgammaC-null mouse by combining *in vitro* fertilization and CRISPR/Cas9 technology. *Sci. Rep.* **4**, 5290. doi: 10.1038/srep05290
11. Wang H., Yang H., Shivalila C.S., Dawlaty M.M., Cheng A.W., Zhang F., Jaenisch R. (2013) One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell*. **153**, 910–918. doi: 10.1016/j.cell.2013.04.025
12. Sunagawa G.A., Sumiyama K., Ukai-Tadenuma M., Perrin D., Fujishima H., Ukai H., Nishimura O., Shi S., Ohno R.I., Narumi R., Shimizu Y., Tone D., Ode K.L., Kuraku S., Ueda H.R. (2016) Mammalian reverse genetics without crossing reveals Nr3a as a short-sleeper gene. *Cell Rep.* **14**, 662–677. doi: 10.1016/j.celrep.2015.12.052
13. Arai D., Nakao Y. (2021) Efficient biallelic knock-in in mouse embryonic stem cells by *in vivo*-linearization of donor and transient inhibition of DNA polymerase  $\theta$ /DNA-PK. *Sci Rep.* **11**, 18132. doi:10.1038/s41598-021-97579-8
14. Li X., Sun B., Qian H., Ma J., Paolino M., Zhang Z. (2022) A high-efficiency and versatile CRISPR/Cas9-mediated HDR-based biallelic editing system. *J. Zhejiang Univ. Sci. B*. **23**, 141–152. doi: 10.1631/jzus.B2100196
15. Shalem O., Sanjana N.E., Hartenian E., Shi X., Scott D.A., Mikkelsen T., Heckl D., Ebert B.L., Root D.E., Doench J.G., Zhang F. (2014) Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science*. **343**, 84–87. doi: 10.1126/science.1247005
16. Wang T., Wei J.J., Sabatini D.M., Lander E.S. (2014) Genetic screens in human cells using the CRISPR-Cas9 system. *Science*. **343**, 80–84. doi: 10.1126/science.1246981
17. Zhou Y., Zhu S., Cai C., Yuan P., Li C., Huang Y., Wei W. (2014) High-throughput screening of a CRISPR/Cas9 library for functional genomics in human cells. *Nature*. **509**, 487–491. doi: 10.1038/nature13166
18. Lino C.A., Harper J.C., Carney J.P., Timlin J.A. (2018) Delivering crispr: a review of the challenges and approaches. *Drug Deliv.* **25**, 1234–1257. doi: 10.1080/10717544.2018.1474964
19. Slaymaker I.M., Gao L., Zetsche B., Scott D.A., Yan W.X., Zhang F. (2016) Rationally engineered Cas9 nucleases with improved specificity. *Science*. **351**, 84–88. doi: 10.1126/science.aad5227
20. Doench J.G., Fusi N., Sullender M., Hegde M., Vaimberg E.W., Donovan K.F., Smith I., Tothova Z., Wilen C., Orchard R., Virgin H.W., Listgarten J., Root D.E. (2016) Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat. Biotechnol.* **34**, 184–191. doi: 10.1038/nbt.3437
21. Hsu P.D., Scott D.A., Weinstein J.A., Ran F.A., Konermann S., Agarwala V., Li Y., Fine E.J., Wu X., Shalem O., Cradick T.J., Marraffini L.A., Bao G., Zhang F. (2013) DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.* **31**, 827–832. doi: 10.1038/nbt.2647
22. 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. doi: 10.1126/science.1231143
23. Turchiano G., Andrieux G., Klermund J., Blattner G., Pennucci V., El Gaz M., Monaco G., Poddar S., Mussolino C., Cornu T.I., Boerries M., Cathomen T. (2021) Quantitative evaluation of chromosomal rearrangements in gene-edited human stem cells by CAST-Seq. *Cell Stem Cell*. **28**, 1136–1147.e5. doi: 10.1016/j.stem.2021.02.002
24. Parikh B.A., Beckman D.L., Patel S.J., White J.M., Yokoyama W.M. (2015) Detailed phenotypic and molecular analyses of genetically modified mice generated by CRISPR-Cas9-mediated editing. *PLoS One*. **10**, e0116484. doi: 10.1371/journal.pone.0116484
25. Traxler E.A., Yao Y., Wang Y.D., Woodard K.J., Kurita R., Nakamura Y., Hughes J.R., Hardison R.C., Blobel G.A., Li C., Weiss M.J. (2016) A genome-editing strategy to treat  $\beta$ -hemoglobinopathies that recapitulates a mutation associated with a benign genetic condition. *Nat. Med.* **22**, 987–990. doi: 10.1038/nm.4170
26. Cullot G., Boutin J., Toutain J., Prat F., Pennamen P., Rooryck C., Teichmann M., Rousseau E., Lamrissi-Garcia I., Guyonnet-Duperat V., Bibeyran A., Lalanne M., Prouzet-Mauléon V., Turcq B., Ged C., Blouin J.M., Richard E., Dabernat S., Moreau-Gaudry F., Bedel A. (2019) CRISPR-Cas9 genome editing induces megabase-scale chDromosomal truncations. *Nat. Commun.* **10**, 1136. doi: 10.1038/s41467-019-09006-2
27. Kosicki M., Tomberg K., Bradley A. (2018) Repair of double-strand breaks induced by CRISPR-Cas9 leads to large

- deletions and complex rearrangements. *Nat. Biotechnol.* **36**, 765–771.  
doi: 10.1038/nbt.4192
28. Xu C.L., Ruan M.Z.C., Mahajan V.B., Tsang S.H. (2019) Viral delivery systems for CRISPR. *Viruses.* **11**, 28.  
doi: 10.3390/v11010028
  29. Charlesworth C.T., Deshpande P.S., Dever D.P., Camarena J., Lemgart V.T., Cromer M.K., Vakulskas C.A., Collingwood M.A., Zhang L., Bode N.M., Behlke M.A., Dejene B., Cieniewicz B., Romano R., Lesch B.J., Gomez-Ospina N., Mantri S., Pavel-Dinu M., Weinberg K.I., Porteus M.H. (2019) Identification of preexisting adaptive immunity to Cas9 proteins in humans. *Nat. Med.* **25**, 249–254.  
doi: 10.1038/s41591-018-0326-x
  30. Song M., Koo T. (2021) Recent advances in CRISPR technologies for genome editing. *Arch. Pharm. Res.* **44**, 537–552.  
doi: 10.1007/s12272-021-01336-4
  31. Walton R.T., Christie K.A., Whittaker M.N., Kleinstiver B.P. (2020) Unconstrained genome targeting with near-PAMless engineered CRISPR-Cas9 variants. *Science.* **368**, 290–296.  
doi: 10.1126/science.aba8853
  32. Wang Y., Wang B., Xie H., Ren Q., Liu X., Li F., Lv X., He X., Cheng C., Deng R., Li J., Zhao J., Song Z., Gu F. (2019) Efficient human genome editing using SaCas9 ribonucleoprotein complexes. *Biotechnol. J.* **14**, e1800689.  
doi: 10.1002/biot.201800689
  33. Görücü Yılmaz S. (2021) Genome editing technologies: CRISPR, LEAPER, RESTORE, ARCUT, SATI, and RESCUE. *EXCLI J.* **20**, 19–45.  
doi: 10.17179/excli2020-3070
  34. Kleinstiver B.P., Sousa A.A., Walton R.T., Tak Y.E., Hsu J.Y., Clement K., Welch M.M., Horng J.E., Malagon-Lopez J., Scarfò I., Maus M.V., Pinello L., Aryee M.J., Joung J.K. (2019) Engineered CRISPR–Cas12a variants with increased activities and improved targeting ranges for gene, epigenetic and base editing. *Nat. Biotechnol.* **37**, 276–282.  
doi: 10.1038/s41587-018-0011-0
  35. Li S., Zhang X., Wang W., Guo X., Wu Z., Du W., Zhao Y., Xia L. (2018) Expanding the scope of CRISPR/Cpf1-mediated genome editing in rice. *Mol. Plant.* **11**, 995–998.  
doi: 10.1016/j.molp.2018.03.009
  36. Gasiunas G., Barrangou R., Horvath P., Siksnys V. (2012) Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc. Natl. Acad. Sci. USA.* **109**, 2579–2586.  
doi: 10.1073/pnas.1208507109
  37. Qi L.S., Larson M.H., Gilbert L.A., Doudna J.A., Weissman J.S., Arkin A.P., Lim W.A. (2013) Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell.* **152**, 1173–1183.  
doi: 10.1016/j.cell.2013.02.022
  38. Gilbert L.A., Larson M.H., Morsut L., Liu Z., Brar G.A., Torres S.E., Stern-Ginossar N., Brandman O., Whitehead E.H., Doudna J.A., Lim W.A., Weissman J.S., Qi L.S. (2013) XCRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell.* **154**, 442–451.  
doi: 10.1016/j.cell.2013.06.044
  39. Maeder M.L., Linder S.J., Cascio V.M., Fu Y., Ho Q.H., Joung J.K. (2013) CRISPR RNA-guided activation of endogenous human genes. *Nat. Methods.* **10**, 977–979.  
doi: 10.1038/nmeth.2598
  40. Balboa D., Weltner J., Euroola S., Trokovic R., Wartiovaara K., Otonkoski T. (2015) Conditionally stabilized dCas9 activator for controlling gene expression in human cell reprogramming and differentiation. *Stem Cell Rep.* **5**, 448–459.  
doi: 10.1016/j.stemcr.2015.08.001
  41. Gilbert L.A., Horlbeck M.A., Adamson B., Villalta J.E., Chen Y., Whitehead E.H., Guimaraes C., Panning B., Ploegh H.L., Bassik M.C., Qi L.S., Kampmann M., Weissman J.S. (2014) Genome-scale CRISPR-mediated control of gene repression and activation. *Cell.* **159**, 647–661.  
doi: 10.1016/j.cell.2014.09.029
  42. Wang G., Chow R.D., Bai Z., Zhu L., Errami Y., Dai X., Dong M.B., Ye L., Zhang X., Renauer P.A., Park J.J., Shen L., Ye H., Fuchs C.S., Chen S. (2019) Multiplexed activation of endogenous genes by CRISPR a elicits potent antitumor immunity. *Nat. Immunol.* **20**, 1494–1505.  
doi: 10.1038/s41590-019-0500-4
  43. Liu P., Chen M., Liu Y., Qi L.S., Ding S. (2018) CRISPR-based chromatin remodeling of the endogenous Oct4 or Sox2 locus enables reprogramming to pluripotency. *Cell Stem Cell.* **22**, 252–261.e4.  
doi: 10.1016/j.stem.2017.12.001
  44. Nuñez J.K., Chen J., Pommier G.C., Cogan J.Z., Replogle J.M., Adriaens C., Ramadoss G.N., Shi Q., Hung K.L., Samelson A.J., Pogson A.N., Kim J.Y.S., Chung A., Leonetti M.D., Chang H.Y., Kampmann M., Bernstein B.E., Hovestadt V., Gilbert L.A., Weissman J.S. (2021) Genome-wide programmable transcriptional memory by CRISPR-based epigenome editing. *Cell.* **184**, 2503–2519.e17.  
doi: 10.1016/j.cell.2021.03.025
  45. Zhang F., Song G., Tian Y. (2019) Anti-CRISPRs: The natural inhibitors for CRISPR-Cas systems. *Animal Model Exp. Med.* **2**, 69–75.  
doi: 10.1002/ame2.12069
  46. Malone L.M., Birkholz N., Fineran P.C. (2021) Conquering CRISPR: how phages overcome bacterial adaptive immunity. *Curr. Opin. Biotechnol.* **68**, 30–36.  
doi: 10.1016/j.copbio.2020.09.008
  47. Harrington L.B., Doxzen K.W., Ma E., Liu J.J., Knott G.J., Edraki A., Garcia B., Amrani N., Chen J.S., Cofsky J.C., Kranzusch P.J., Sontheimer E.J., Davidson A.R., Maxwell K.L., Doudna J.A. (2017) A broad-spectrum inhibitor of CRISPR-Cas9. *Cell.* **170**, 1224–1233.e15.  
doi: 10.1016/j.cell.2017.07.037
  48. Liu L., Yin M., Wang M., Wang Y. (2019) Phage AcrIIA2 DNA mimicry: structural basis of the CRISPR and anti-CRISPR arms race. *Mol. Cell.* **73**, 611–620.e3.  
doi: 10.1016/j.molcel.2018.11.011
  49. Jo D.H., Koo T., Cho C.S., Kim J.H., Kim J.S., Kim J.H. (2019) Long-term effects of *in vivo* genome editing in the mouse retina using *Campylobacter jejuni* Cas9 expressed via adeno-associated virus. *Mol. Therapy.* **27**, 130–136.  
doi: 10.1016/j.ymthe.2018.10.009
  50. Harrington L.B., Burstein D., Chen J.S., Paez-Espino D., Ma E., Witte I.P., Cofsky J.C., Kyripides N.C., Banfield J.F., Doudna J.A. (2018) Programmed DNA destruction by miniature CRISPR-Cas14 enzymes. *Science.* **362**, 839–842.  
doi: 10.1126/science.aav4294
  51. Gupta R., Ghosh A., Chakravarti R., Singh R., Ravichandiran V., Swarnakar S., Ghosh D. (2022) Cas13d: a new molecular scissor for transcriptome engineering. *Front. Cell Dev. Biol.* **10**, 866800.  
doi: 10.3389/fcell.2022.866800

52. Koo T., Park S.W., Jo D.H., Kim D., Kim J.H., Cho H.Y., Kim J., Kim J.H., Kim J.S. (2018) CRISPR-LbCpf1 prevents choroidal neovascularization in a mouse model of age-related macular degeneration. *Nat. Commun.* **9**, 1855. doi:10.1038/s41467-018-04175-y
53. Chen J.S., Ma E., Harrington L.B., Da Costa M., Tian X., Palefsky J.M., Doudna J.A. (2018) CRISPR-Cas12a target binding unleashes indiscriminate single-stranded DNase activity. *Science*. **360**, 436–439. doi: 10.1126/science.aar6245
54. Leung R.K., Cheng Q.X., Wu Z.L., Khan G., Liu Y., Xia H.Y., Wang J. (2022) CRISPR-Cas12-based nucleic acids detection systems. *Methods*. **203**, 276–281. doi: 10.1016/j.ymeth.2021.02.018
55. Wang Z., Wang Y., Wang S., Gorzalski A.J., McSwiggin H., Yu T., Castaneda-Garcia K., Prince B., Wang H., Zheng H., Yan W. (2020) Efficient genome editing by CRISPR-Mb-3Cas12a in mice. *J. Cell Sci.* **133**, jcs240705. doi: 10.1242/jcs.240705
56. Altae-Tran H., Kannan S., Demircioglu F.E., Oshiro R., Nety S.P., McKay L.J., Dlakić M., Inskoop W.P., Makarova K.S., Macrae R.K., Koonin E.V., Zhang F. (2021) The widespread IS200/IS605 transposon family encodes diverse programmable RNA-guided endonucleases. *Science*. **374**, 57–65. doi: 10.1126/science.abj6856
57. Cox D.B.T., Gootenberg J.S., Abudayyeh O.O., Franklin B., Kellner M.J., Joung J., Zhang F. (2017) RNA editing with CRISPR-Cas13. *Science*. **358**, 1019–1027. doi: 10.1126/science.aaq0180
58. Abudayyeh O.O., Gootenberg J.S., Franklin B., Koob J., Kellner M.J., Ladha A., Joung J., Kirchgatterer P., Cox D.B.T., Zhang F. (2019) A cytosine deaminase for programmable single-base RNA editing. *Science*. **365**, 382–386. doi: 10.1126/science.aax7063
59. Hoeijmakers J.H.J. (2009) DNA damage, aging, and cancer. *NEJM*. **361**, 1475–1485. <https://doi.org/10.1056/NEJMra0804615>
60. Lieber M.R. (2010) The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annu. Rev. Biochem.* **79**, 181–211. doi: 10.1146/annurev.biochem.052308.093131
61. Sirbu B.M., Cortez D. (2013) DNA damage response: tHdR levels of DNA repair regulation. *Cold Spring Harb. Perspect. Biol.* **5**, a012724. doi: 10.1101/cshperspect.a012724
62. Baudat F., Imai Y., De Massy B. (2013) Meiotic recombination in mammals: localization and regulation. *Nat. Rev. Genet.* **14**, 794–806. doi: 10.1038/nrg3573
63. Schatz D.G., Ji Y. (2011) Recombination centres and the orchestration of V(D)J recombination. *Nat. Rev. Immunol.* **11**, 251–263. doi: 10.1038/nri2941
64. van de Kooij B., van Attikum H. (2021) Genomic reporter constructs to monitor pathway-specific repair of DNA double-strand breaks. *Front Genet.* **12**, 809832. doi: 10.3389/fgene.2021.809832
65. Cannan W.J., Pederson D.S. (2016) Mechanisms and consequences of double-strand DNA break formation in chromatin. *J. Cell Physiol.* **231**, 3–14. doi: 10.1002/jcp.25048
66. Zhang X., Li T., Ou J., Huang J., Liang P. (2022) Homology-based repair induced by CRISPR-Cas nucleases in mammalian embryo genome editing. *Protein Cell*. **13**, 316–335. doi: 10.1007/s13238-021-00838-7
67. Denes C.E., Cole A.J., Aksoy Y.A., Li G., Neely G.G., Hesselson D. (2021) Approaches to enhance precise CRISPR/Cas9-mediated genome editing. *Int. J. Mol. Sci.* **22**, 8571. doi: 10.3390/ijms22168571
68. Yang H., Ren S., Yu S., Pan H., Li T., Ge S., Zhang J., Xia N. (2020) Methods favoring homology-directed repair choice in response to CRISPR/Cas9 induced-double strand breaks. *Int. J. Mol. Sci.* **21**, 6461. doi: 10.3390/ijms21186461
69. Sun W., Liu H., Yin W., Qiao J., Zhao X., Liu Y. (2022) Strategies for enhancing the homology-directed repair efficiency of CRISPR-Cas systems. *CRISPR J.* **5**, 7–18. doi: 10.1089/crispr.2021.0039
70. Ma J., Zhou Y., Pan P., Yu H., Wang Z., Li L.L., Wang B., Yan Y., Pan Y., Ye Q., Liu T., Feng X., Xu S., Wang K., Wang X., Jian Y., Ma B., Fan Y., Gao Y., Huang H., Li L. (2023) TRABID overexpression enables synthetic lethality to PARP inhibitor via prolonging 53BP1 retention at double-strand breaks. *Nat. Commun.* **14**, 1810. doi: 10.1038/s41467-023-37499-5
71. Li G., Liu D., Zhang X., Quan R., Zhong C., Mo J., Huang Y., Wang H., Ruan X., Xu Z., Zheng E., Gu T., Hong L., Li Z., Wu Z., Yang H. (2018) Suppressing Ku70/Ku80 expression elevates homology-directed repair efficiency in primary fibroblasts. *Int. J. Biochem. Cell Biol.* **99**, 154–160. doi: 10.1016/j.biocel.2018.04.011
72. Pawelczak K.S., Gavande N.S., VanderVere-Carozza P.S., Turchi J.J. (2018) Modulating DNA repair pathways to improve precision genome engineering. *ACS Chem. Biol.* **13**, 389–396. doi: 10.1021/acscchembio.7b00777
73. Gavande N.S., VanderVere-Carozza P.S., Pawelczak K.S., Mendoza-Munoz P., Vernon T.L., Hanakahi L.A., Summerlin M., Dynlacht J.R., Farmer A.H., Sears C.R., Nasrallah N.A., Garrett J., Turchi J.J. (2020) Discovery and development of novel DNA-PK inhibitors by targeting the unique Ku-DNA interaction. *Nucl. Acids Res.* **48**, 11536–11550. doi: 10.1093/nar/gkaa934
74. Chu V.T., Weber T., Wefers B., Wurst W., Sander S., Rajewsky K., Kühn R. (2015) Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. *Nat. Biotechnol.* **33**, 543–548. doi: 10.1038/nbt.3198
75. Li G., Quan R., Wang H., Ruan X., Mo J., Zhong C., Yang H., Li Z., Gu T., Liu D., Wu Z., Cai G., Zhang X. (2019) Inhibition of KU70 and KU80 by CRISPR interference, not NgAgo interference, increases the efficiency of homologous recombination in pig fetal fibroblasts. *J. Integr. Agric.* **18**, 438–448. doi: 10.1016/S2095-3119(18)62150-1
76. Shy B.R., MacDougall M.S., Clarke R., Merrill B.J. (2016) Co-incident insertion enables high efficiency genome engineering in mouse embryonic stem cells. *Nucl. Acids Res.* **44**, 7997–8010. doi: 10.1093/nar/gkw685
77. Yu W., Li L., Wang G., Zhang W., Xu J., Liang A. (2018) KU70 inhibition impairs both non-homologous end joining and homologous recombination DNA damage repair

- tHDRough SHP-1 induced dephosphorylation of SIRT1 in adult T-cell leukemia-lymphoma cells. *Cell. Physiol. Biochem.* **49**, 2111–2123.  
doi: 10.1159/000493815
78. Weterings E., Gallegos A.C., Dominick L.N., Cooke L.S., Bartels T.N., Vagner J., Matsunaga T.O., Mahadevan D. (2016) A novel small molecule inhibitor of the DNA repair protein Ku70/80. *DNA Repair.* **43**, 98–106.  
doi: 10.1016/j.dnarep.2016.03.014
79. Riesenbergs S., Maricic T. (2018) Targeting repair pathways with small molecules increases precise genome editing in pluripotent stem cells. *Nat. Commun.* **9**, 2164.  
doi: 10.1038/s41467-018-04609-7
80. Mohiuddin I.S., Kang M.H. (2019) DNA-PK as an emerging therapeutic target in cancer. *Front Oncol.* **9**, 635.  
doi: 10.3389/fonc.2019.00635
81. Robert F., Barbeau M., Éthier S., Dostie J., Pelletier J. (2015) Pharmacological inhibition of DNA-PK stimulates Cas9-mediated genome editing. *Genome Med.* **7**, 93.  
doi: 10.1186/s13073-015-0215-6
82. Aksoy Y.A., Nguyen D.T., Chow S., Chung R.S., Guillemin G.J., Cole N.J., Hesselson D. (2019) Chemical reprogramming enhances homology-directed genome editing in zebrafish embryos. *Commun. Biol.* **2**, 198.  
doi: 10.1038/s42003-019-0444-0
83. Riesenbergs S., Chintalapati M., Macak D., Kanis P., Maricic T., Pääbo S. (2019) Simultaneous precise editing of multiple genes in human cells. *Nucl. Acids Res.* **47**, e116.  
doi: 10.1093/nar/gkz669
84. Fu Y.W., Dai X.Y., Wang W.T., Yang Z.X., Zhao J.J., Zhang J.P., Wen W., Zhang F., Oberg K.C., Zhang L., Cheng T., Zhang X.B. (2021) Dynamics and competition of CRISPR-Cas9 ribonucleoproteins and AAV donor-mediated NHEJ, MMEJ and HDR editing. *Nucl. Acids Res.* **49**, 969–985.  
doi: 10.1093/nar/gkaa1251
85. Ray U., Raghavan S.C. (2020) Modulation of DNA double-strand break repair as a strategy to improve precise genome editing. *Oncogene.* **39**, 6393–6405.  
doi: 10.1038/s41388-020-01445-2
86. Ray U., Vartak S.V., Raghavan S.C. (2020) NHEJ inhibitor SCR7 and its different forms: Promising CRISPR tools for genome engineering. *Gene.* **763**, 144997.  
doi: 10.1016/j.gene.2020.144997
87. Shao S., Ren C., Liu Z., Bai Y., Chen Z., Wei Z., Wang X., Zhang Z., Xu K. (2017) Enhancing CRISPR/Cas9-mediated homology-directed repair in mammalian cells by expressing *Saccharomyces cerevisiae* Rad52. *Int. J. Biochem. Cell Biol.* **92**, 43–52.  
doi: 10.1016/j.biocel.2017.09.012
88. Maruyama T., Dougan S.K., Truttmann M.C., Bilate A.M., Ingram J.R., Ploegh H.L. (2015) Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. *Nat. Biotechnol.* **33**, 538–542.  
doi: 10.1038/nbt.3190
89. Aird E.J., Lovendahl K.N., St Martin A., Harris R.S., Gordon W.R. (2018) Increasing Cas9-mediated homology-directed repair efficiency tHDRough covalent tethering of DNA repair template. *Commun. Biol.* **1**, 54.  
doi: 10.1038/s42003-018-0054-2
90. Song J., Yang D., Xu J., Zhu T., Chen Y.E., Zhang J. (2016) RS-1 enhances CRISPR/Cas9- and TALEN-mediated knock-in efficiency. *Nat. Commun.* **7**, 10548.  
doi: 10.1038/ncomms10548
91. Gutschner T., Haemmerle M., Genovese G., Draetta G.F., Chin L. (2016) Post-translational regulation of Cas9 during G1 enhances homology-directed repair. *Cell Rep.* **14**, 1555–1566.  
doi: 10.1016/j.celrep.2016.01.019
92. Xie Z., Pang D., Wang K., Li M., Guo N., Yuan H., Li J., Zou X., Jiao H., Ouyang H., Li Z., Tang X. (2017) Optimization of a CRISPR/Cas9-mediated knock-in strategy at the porcine Rosa26 locus in porcine foetal fibroblasts. *Sci. Rep.* **7**, 3036.  
doi: 10.1038/s41598-017-02785-y
93. Yang D., Scavuzzo M.A., Chmielowiec J., Sharp R., Bajic A., Borowiak M. (2016) Enrichment of G2/M cell cycle phase in human pluripotent stem cells enhances HDR-mediated gene repair with customizable endonucleases. *Sci. Rep.* **6**, 21264.  
doi: 10.1038/srep21264
94. Canny M.D., Moatti N., Wan L.C.K., Fradet-Turcotte A., Krasner D., Mateos-Gomez P.A., Zimmermann M., Orthwein A., Juang Y.C., Zhang W., Noordermeer S.M., Seclen E., Wilson M.D., Vorobyov A., Munro M., Ernst A., Ng T.F., Cho T., Cannon P.M., Sidhu S.S., Durocher D. (2018) Inhibition of 53BP1 favors homology-dependent DNA repair and increases CRISPR-Cas9 genome-editing efficiency. *Nat. Biotechnol.* **36**, 95–102.  
doi: 10.1038/nbt.4021
95. Jayavaradhan R., Pillis D.M., Malik P.A. (2019) Versatile tool for the quantification of CRISPR/Cas9-induced genome editing events in human hematopoietic cell lines and hematopoietic stem/progenitor cells. *J. Mol. Biol.* **431**, 102–110.  
doi: 10.1016/j.jmb.2018.05.005
96. Wienert B., Nguyen D.N., Guenther A., Feng S.J., Locke M.N., Wyman S.K., Shin J., Kazane K.R., Gregory G.L., Carter M.A.M., Wright F., Conklin B.R., Marson A., Richardson C.D., Corn J.E. (2020) Timed inhibition of CDC7 increases CRISPR-Cas9 mediated templated repair. *Nat. Commun.* **11**, 2109.  
doi: 10.1038/s41467-020-15845-1
97. Ma X., Chen X., Jin Y., Ge W., Wang W., Kong L., Ji J., Guo X., Huang, J., Feng X.H., Fu J., Zhu S. (2018) Small molecules promote CRISPR-Cpf1-mediated genome editing in human pluripotent stem cells. *Nat. Commun.* **9**, 1303.  
doi: 10.1038/s41467-018-03760-5
98. Vartak S.V., Swarup H.A., Gopalakrishnan V., Gopinatha V.K., Ropars V., Nambiar M., John F., Kothanahally S.K.S., Kumari R., Kumari N., Ray U., Radha G., Dinesh D., Pandey M., Ananda H., Karki S.S., Srivastava M., Charbonnier J.B., Choudhary B., Mantelingu K., Raghavan S.C. (2018) Autocyclized and oxidized forms of SCR7 induce cancer cell death by inhibiting nonhomologous DNA end joining in a ligase IV dependent manner. *FEBS J.* **285**, 3959–3976.  
doi: 10.1111/febs.14661
99. Killian T., Dickopf S., Haas A.K., Kirstenpfad C., Mayer K., Brinkmann U. (2017) Disruption of diphthamide synthesis genes and resulting toxin resistance as a robust technology for quantifying and optimizing CRISPR/Cas9-mediated gene editing. *Sci. Rep.* **7**, 15480.  
doi: 10.1038/s41598-017-15206-x

100. Aslan Y., Tadjuidje E., Zorn A.M., Cha S.W. (2017) High-efficiency non-mosaic CRISPR-mediated knock-in and In-Del mutation in F0 *Xenopus*. *Development*. **144**, 2852–2858. doi: 10.1242/dev.152967
101. Ray U., Raul S.K., Gopinatha V.K., Ghosh D., Rangappa K.S., Mantelingu K., Raghavan S.C. (2020) Identification and characterization of novel SCR7-based small-molecule inhibitor of DNA end-joining, SCR130 and its relevance in cancer therapeutics. *Mol. Carcinog.* **59**, 618–628. doi: 10.1002/mc.23186
102. Scully R., Panday A., Elango R., Willis N.A. (2019) DNA double-strand break repair-pathway choice in somatic mammalian cells. *Nat. Rev. Mol. Cell. Biol.* **20**, 698–714. doi: 10.1038/s41580-019-0152-0
103. Yeh C.D., Richardson C.D., Corn J.E. (2019) Advances in genome editing tHDRough control of DNA repair pathways. *Nat. Cell Biol.* **21**, 1468–1478. doi: 10.1038/s41556-019-0425-z
104. Richardson C.D., Kazane K.R., Feng S.J., Zelin E., Bray N.L., Schäfer A.J., Floor S.N., Corn J.E. (2018) CRISPR-Cas9 genome editing in human cells occurs via the Fanconi anemia pathway. *Nat. Genet.* **50**, 1132–1139. doi: 10.1038/s41588-018-0174-0
105. Syed A., Tainer J.A. (2018) The MRE11-RAD50-NBS1 complex conducts the orchestration of damage signaling and outcomes to stress in DNA replication and repair. *Annu. Rev. Biochem.* **87**, 263–294. doi: 10.1146/annurev-biochem-062917-012415
106. Reuven N., Adler J., Broennimann K., Myers N., Shaul Y. (2019) Recruitment of DNA repair MRN complex by intrinsically disordered protein domain fused to Cas9 improves efficiency of CRISPR-mediated genome editing. *Biomolecules*. **9**, 584. doi: 10.3390/biom9100584
107. Lee K.J., Saha J., Sun J., Fattah K.R., Wang S.C., Jakob B., Chi L., Wang S.Y., Taucher-Scholz G., Davis A.J., Chen D.J. (2015) Phosphorylation of Ku dictates DNA double-strand break (DSB) repair pathway choice in S phase. *Nucl. Acids Res.* **44**, 1732–1745. doi: 10.1093/nar/gkv1499
108. Caron M.C., Sharma A.K., O’Sullivan J., Myler L.R., Ferreira M.T., Rodrigue A., Coulombe Y., Ethier C., Gagné J.P., Langelier M.F., Pascal J.M., Finkelstein I.J., Hendzel M.J., Poirier G.G., Masson J.Y. (2019) Poly-(ADP-ribose) polymerase-1 antagonizes DNA resection at double-strand breaks. *Nat. Commun.* **10**, 2954. doi: 10.1038/s41467-019-10741-9
109. Huertas P., Jackason S.P. (2009) Human CtIP mediates cell cycle control of DNA end resection and double strand break repair. *J. Biol. Chem.* **284**, 9558–9565. doi: 10.1074/jbc.M808906200
110. Anand R., Ranjha L., Cannavo E., Cejka P. (2016) Phosphorylated CtIP functions as a co-factor of the MRE11-RAD50-NBS1 endonuclease in DNA end resection. *Mol. Cell.* **64**, 940–950. doi: 10.1016/j.molcel.2016.10.017
111. Charpentier M., Khedher A.H.Y., Menoret S., Brion A., Lamribet K., Dardillac E., Boix C., Perrouault L., Tesson L., Geny S., De Cian A., Itier J.M., Anegon I., Lopez B., Giovannangeli C., Concordet J.P. (2018) CtIP fusion to Cas9 enhances transgene integration by homology-dependent repair. *Nat. Commun.* **9**, 1–11. doi: 10.1038/s41467-018-03475-7
112. Nimonkar A.V., Genschel J., Kinoshita E., Polaczek P., Campbell J.L., Wyman C., Modrich P., Kowalczykowski S.C. (2011) BLM-DNA2-RPA-MRN and EXO1-BLM-RPA-MRN constitute two DNA end resection machineries for human DNA break repair. *Genes Dev.* **25**, 350–362. doi: 10.1101/gad.2003811
113. Myler L.R., Gallardo I.F., Zhou Y., Gong F., Yang S.H., Wold M.S., Miller K.M., Paull T.T., Finkelstein I.J. (2016) Single-molecule imaging reveals the mechanism of Exo1 regulation by single-stranded DNA binding proteins. *Proc. Natl. Acad. Sci. USA.* **113**, E1170–9. doi: 10.1073/pnas.1516674113
114. Symington L.S. (2016) Mechanism and regulation of DNA end resection in eukaryotes. *Crit. Rev. Biochem. Mol. Biol.* **51**, 195–212. doi: 10.3109/10409238.2016.1172552
115. Daley J.M., Jimenez-Sainz J., Wang W., Miller A.S., Xue X., Nguyen K.A., Jensen R.B., Sung P. (2017) Enhancement of BLM-DNA2-mediated long-range DNA end resection by CtIP. *Cell Rep.* **21**, 324–332. doi: 10.1016/j.celrep.2017.09.048
116. Tarsounas M., Sung P. (2020) The antitumorigenic roles of BRCA1-BARD1 in DNA repair and replication. *Nat. Rev. Mol. Cell Biol.* **21**, 284–299. doi: 10.1038/s41580-020-0218-z
117. Becker J.R., Bonnet C., Clifford G., Groth A., Wilson M.D., Chapman J.R. (2020) BARD1 links histone H2A lysine-15 ubiquitination to initiation of BRCA1-dependent homologous recombination. *bioRxiv. Cold Spring Harb. Lab.* doi: https://doi.org/10.1101/2020.06.01.127951
118. Nambiar T.S., Billon P., Diedenhofen G., Hayward S.B., Tagliatela A., Cai K., Huang J.W., Leuzzi G., Cuella-Martin R., Palacios A., Gupta A., Egli D., Ciccia A. (2019) Stimulation of CRISPR-mediated homology-directed repair by an engineered RAD18 variant. *Nat. Commun.* **10**, 3395. doi: 10.1038/s41467-019-11105-z
119. Xue C., Greene E.C. (2021) DNA repair pathway choices in CRISPR-Cas9-mediated genome editing. *Trends Genet.* **37**, 639–656. doi: 10.1016/j.tig.2021.02.008
120. Yu C., Liu Y., Ma T., Liu K., Xu S., Zhang Y., Liu H., La Russa M., Xie M., Ding S., Qi L.S. (2015) Small molecules enhance CRISPR genome editing in pluripotent stem cells. *Cell Stem Cell.* **16**, 142–147. doi: 10.1016/j.stem.2015.01.003
121. Dhingra N., Zhao X. (2019) Intricate SUMO-based control of the homologous recombination machinery. *Genes Dev.* **33**, 1346–1354. doi: 10.1101/gad.328534.119
122. Soria-Bretones I., Cepeda-García C., Checa-Rodríguez C., Heyer V., Reina-San-Martin B., Soutoglou E., Huertas P. (2017) DNA end resection requires constitutive sumoylation of CtIP by CBX4. *Nat. Commun.* **8**, 113. doi: 10.1038/s41467-017-00183-6
123. Bologna S., Altmannova V., Valtorta E., Koenig C., Liberali P., Gentili C., Anrather D., Ammerer G., Pelkmans L., Krejci L., Ferrari S. (2015) Sumoylation regulates EXO1 stability and processing of DNA damage. *Cell Cycle.* **14**, 2439–2450. doi: 10.1080/15384101.2015.1060381

124. Lin S., Staahl B.T., Alla R.K., Doudna J.A. (2014) Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. *Elife*. **3**, e04766. doi: 10.7554/eLife.04766
125. Vassilev L.T. (2006) Cell cycle synchronization at the G2/M phase border by reversible inhibition of CDK1. *Cell Cycle*. **5**, 2555–2556. doi: 10.4161/cc.5.22.3463
126. Lomova A., Clark D.N., Campo-Fernandez B., Flores-Björström C., Kaufman M.L., Fitz-Gibbon S., Wang X., Miyahira E.Y., Brown D., DeWitt M.A., Corn J.E., Hollis R.P., Romero Z., Kohn D.B. (2019) Improving gene editing outcomes in human hematopoietic stem and progenitor cells by temporal control of DNA repair. *Stem Cells*. **37**, 284–294. doi: 10.1002/stem.2935
127. Gerlach M., Kraft T., Brenner B., Petersen B., Niemann H., Montag J. (2018) Efficient knock-in of a point mutation in porcine fibroblasts using the CRISPR/Cas9-GMNN fusion gene. *Genes*. **9**, 296. doi: 10.3390/genes9060296
128. Howden S.E., McColl B., Glaser A., Vadolas J., Petrou S., Little M.H., Elefanti A.G., Stanley E.G. (2016) A Cas9 variant for efficient generation of InDel-free knockin or gene-corrected human pluripotent stem cells. *Stem Cell Rep.* **7**, 508–517. doi: 10.1016/j.stemcr.2016.07.001
129. Schep R., Brinkman E.K., Leemans C., Vergara X., van der Weide R.H., Morris B., van Schaik T., Manzo S.G., Peric-Hupkes D., van den Berg J., Beijersbergen R.L., Medema R.H., van Steensel B. (2021) Impact of chromatin context on Cas9-induced DNA double-strand break repair pathway balance. *Mol. Cell*. **81**, 2216–2230.e10. doi: 10.1016/j.molcel.2021.03.032
130. Chen Z., Tyler J.K. (2022) The chromatin landscape channels DNA double-strand breaks to distinct repair pathways. *Front. Cell Dev. Biol.* **10**, 909696. doi: 10.3389/fcell.2022.909696
131. Clouaire T., Rocher V., Lashgari A., Arnould C., Aguirreben-goia M., Biernacka A., Skrzypczak M., Aymard F., Fongang B., Dojer N., Iacovoni J.S., Rowicka M., Ginalski K., Côté J., Legube G. (2018) Comprehensive mapping of histone modifications at DNA double-strand breaks deciphers repair pathway chromatin signatures. *Mol. Cell*. **72**, 250–262.e6. doi: 10.1016/j.molcel.2018.08.020
132. Yilmaz D., Furst A., Meaburn K., Lezaja A., Wen Y., Altmeyer M., Reina-San-Martin B., Soutoglou E. (2021) Activation of homologous recombination in G1 preserves centromeric integrity. *Nature*. **600**, 748–753. doi: 10.1038/s41586-021-04200-z
133. Chen X., Rinsma M., Janssen J.M., Liu J., Maggio I., Gonçalves M.A. (2016) Probing the impact of chromatin conformation on genome editing tools. *Nucl. Acids Res.* **44**, 6482–6492. doi: 10.1093/nar/gkw524
134. Janssen J.M., Chen X., Liu J., Gonçalves M.A.F.V. (2019) The chromatin structure of CRISPR-Cas9 target DNA controls the balance between mutagenic and homology-directed gene-editing events. *Mol. Ther. Nucl. Acids*. **16**, 141–154. doi: 10.1016/j.omtn.2019.02.009
135. Liu B., Chen S., Rose A., Chen D., Cao F., Zwinderman M., Kiemel D., Aïssi M., Dekker F.J., Haisma H.J. (2020) Inhibition of histone deacetylase 1 (HDAC1) and HDAC2 enhances CRISPR/Cas9 genome editing. *Nucl. Acids Res.* **48**, 517–532. doi: 10.1093/nar/gkx1136
136. Takayama K., Igai K., Hagihara Y., Hashimoto R., Hanawa M., Sakuma T., Tachibana M., Sakurai F., Yamamoto T., Mizuguchi H. (2017) Highly efficient biallelic genome editing of human ES/iPS cells using a CRISPR/Cas9 or TALEN system. *Nucl. Acids Res.* **45**, 5198–5207. doi: 10.1093/nar/gkx130
137. Knight S.C., Xie L., Deng W., Guglielmi B., Witkowsky L.B., Bosanac L., Zhang E.T., El Beheiry M., Masson J.B., Dahan M., Liu Z., Doudna J.A., Tjian R. (2015) Dynamics of CRISPR-Cas9 genome interrogation in living cells. *Science*. **350**, 823–826. doi: 10.1126/science.aac6572
138. Caron P., Pobega E., Polo S.E. (2021) DNA double-strand break repair: all roads lead to heterochromatin marks. *Front. Genet.* **12**, 730696. doi: 10.3389/fgene.2021.730696
139. Chechik L., Martin O., Soutoglou E. (2020) Genome editing fidelity in the context of DNA sequence and chromatin structure. *Front. Cell. Dev. Biol.* **8**, 319. doi: 10.3389/fcell.2020.00319
140. Song F., Stieger K. (2017) Optimizing the DNA donor template for homology-directed repair of double-strand breaks. *Mol. Ther. Nucl. Acids*. **7**, 53–60. doi: 10.1016/j.omtn.2017.02.006
141. Renaud J.B., Boix C., Charpentier M., De Cian A., Cochennec J., Duvernois-Berthet E., Perrouault L., Tesson L., Edouard J., Thinarth R., Cherif Y., Menoret S., Fontanière S., de Croz N., Fraichard A., Sohm F., Anegón I., Concordet J.P., Giovannangeli C. (2016) Improved genome editing efficiency and flexibility using modified oligonucleotides with TALEN and CRISPR-Cas9 nucleases. *Cell Rep.* **14**, 2263–2272. doi: 10.1016/j.celrep.2016.02.018
142. Lim D., Sreekanth V., Cox K.J., Law B.K., Wagner B.K., Karp J.M., Choudhary A. (2020) Engineering designer beta cells with a CRISPR-Cas9 conjugation platform. *Nat. Commun.* **11**, 4043. doi: 10.1038/s41467-020-17725-0
143. Shola D.T.N., Yang C., Han C., Norinsky R., Peraza R.D. (2021) Generation of mouse model (KI and CKO) via Easi-CRISPR. *Methods Mol. Biol.* **2224**, 1–27. doi: 10.1007/978-1-0716-1008-4\_1
144. Yoon Y., Wang D., Tai P.W.L., Riley J., Gao G., Rivera-Pérez J.A. (2018) Streamlined *ex vivo* and *in vivo* genome editing in mouse embryos using recombinant adeno-associated viruses. *Nat. Commun.* **9**, 412. doi: 10.1038/s41467-017-02706-7
145. Chen S., Sun S., Moonen D., Lee C., Lee A.Y., Schaffer D.V., He L. (2019) CRISPR-READI: efficient generation of knockin mice by CRISPR RNP electroporation and AAV donor infection. *Cell Rep.* **27**, 3780–3789.e4. doi: 10.1016/j.celrep.2019.05.103
146. Bak R.O., Porteus M.H. (2017) CRISPR-mediated integration of large gene cassettes using AAV donor vectors. *Cell Rep.* **20**, 750–756. doi: 10.1016/j.celrep.2017.06.064
147. Yao X., Wang X., Hu X., Liu Z., Liu J., Zhou H., Shen X., Wei Y., Huang Z., Ying W., Wang Y., Nie Y.H., Zhang C.C., Li S., Cheng L., Wang Q., Wu Y., Huang P., Sun Q., Shi L., Yang H. (2017) Homology-mediated end joining-based targeted integration using CRISPR/Cas9. *Cell Res.* **27**, 801–814. doi: 10.1038/cr.2017.76
148. Yao X., Zhang M., Wang X., Ying W., Hu X., Dai P., Meng F., Shi L., Sun Y., Yao N., Zhong W., Li Y., Wu K.,

- Li W., Chen Z.J., Yang H. (2018) Tild-CRISPR allows for efficient and precise gene knock-in in mouse and human cells. *Dev. Cell.* **45**, 526–536.e5. doi: 10.1016/j.devcel.2018.04.021
149. Sakuma T., Nakade S., Sakane Y., Suzuki K.T., Yamamoto T. (2016) MMEJ-assisted gene knock-in using TALENs and CRISPR-Cas9 with the PITCh systems. *Nat. Protoc.* **11**, 118–133. doi: 10.1038/nprot.2015.140
150. Artegiani B., Hendriks D., Beumer J., Kok R., Zheng X., Joore I., Chuva de Sousa Lopes S., van Zon J., Tans S., Clevers H. (2020) Fast and efficient generation of knock-in human organoids using homology-independent CRISPR–Cas9 precision genome editing. *Nat. Cell Biol.* **22**, 321–331. doi: 10.1038/s41556-020-0472-5
151. Lau C.H., Tin C., Suh Y. (2020) CRISPR-based strategies for targeted transgene knock-in and gene correction. *Fac. Rev.* **9**, 20. doi: 10.12703/r/9-20
152. 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. doi: 10.1016/j.cell.2015.09.038
153. Zhao Z., Shang P., Sage F., Geijsen N. (2022) Ligation-assisted homologous recombination enables precise genome editing by deploying both MMEJ and HDR. *Nucl. Acids Res.* **50**, e62. doi: 10.1093/nar/gkac118
154. Cruz-Becerra G., Kadonaga J.T. Enhancement of homology-directed repair with chromatin donor templates in cells. *Elife.* 2020 Apr 28;9:e55780. doi: 10.7554/eLife.55780
155. Carlson-Stevermer J., Abdeen A.A., Kohlenberg L., Goerland M., Molugu K., Lou M., Saha K. (2017) Assembly of CRISPR ribonucleoproteins with biotinylated oligonucleotides via an RNA aptamer for precise gene editing. *Nat. Commun.* **8**, 1711. doi: 10.1038/s41467-017-01875-9
156. Ma M., Zhuang F., Hu X., Wang B., Wen X.Z., Ji J.F., Xi J.J. (2017) Efficient generation of mice carrying homozygous double-floxed alleles using the Cas9-avidin/biotin-donor DNA system. *Cell Res.* **27**, 578–581. doi: 10.1038/cr.2017.29
157. Savic N., Ringnalda F.C., Lindsay H., Berk C., Bargsten K., Li Y., Neri D., Robinson M.D., Ciaudo C., Hall J., Jinek M., Schwank G. (2018) Covalent linkage of the DNA repair template to the CRISPR–Cas9 nuclease enhances homology-directed repair. *Elife.* **7**, e33761. doi: 10.7554/eLife.33761

## How To Shift The Equilibrium Of Dna Break Repair In Favor Of Homology Recombination

O. A. Averina<sup>1, 2, 3</sup>, S. A. Kuznetsova<sup>1, \*</sup>, O. A. Permyakov<sup>1, 3</sup>, P. V. Sergiev<sup>1, 2, 3</sup>

<sup>1</sup>Institute of Functional Genomics, Lomonosov Moscow State University, Moscow, 119991 Russia

<sup>2</sup>Belozersky Institute of Physico-Chemical Biology, Lomonosov Moscow State University, Moscow, 119991 Russia

<sup>3</sup>Department of Chemistry, Lomonosov Moscow State University, Moscow, 119991 Russia

\*e-mail: svetlana@belozersky.msu.ru

With the practical implementation of the CRISPR/Cas technology for targeted genome editing, it has become possible to carry out genetic engineering manipulations with eukaryotic genomes with high efficiency. One of the key stages of this technology is the targeted induction of site-specific DNA cleavages (breaks). The cell repairs these breaks via one of two pathways: nonhomologous end joining or homologous recombination. The choice of DNA repair pathway is determined by the architecture of the sites at the DNA break area formed as a result of terminal resection and depends on the phases of the cell cycle. Nonhomologous end joining is the main pathway for repair of double-stranded DNA breaks in mammalian cells. It involves a nonspecific ligation reaction, the accuracy of which depends on the structure of the ends of the break, and can result in various insertions or deletions in the target region of the genome. Integration of the desired sequence into the genome occurs along the path of homologous recombination, the implementation of which requires a matrix with homology regions on both sides of the double-strand break. The introduction of a genetic construct into a given location in the genome is an important, but currently complex and labor-intensive task. At the same time, for fundamental studies of gene function and the creation of animal models of human diseases, the choice of the repair pathway can be of fundamental importance. This review is an attempt to combine and structure all known information on approaches to increasing the efficiency of DNA repair involving homologous recombination. The article lists the most effective strategies to shift the balance towards homologous repair, such as the use of inhibitors of the non-homologous end joining mechanism, regulation of key factors of homologous recombination, control of the cell cycle, chromatin status, construction of templates for homologous recombination.

**Keywords:** genome editing, CRISPR/Cas technology, nucleases, repair systems, homologous recombination

УДК 575.113:575.8

## ПАТТЕРНЫ АМИНОКИСЛОТНЫХ ЗАМЕН В БЕЛКАХ Е6 И Е7 ВИРУСА ПАПИЛЛОМЫ ЧЕЛОВЕКА 16 ТИПА: ФИЛОГЕОГРАФИЯ И ЭВОЛЮЦИЯ

© 2024 г. Е. Е. Зеленова<sup>a, b, \*</sup>, А. А. Карлсен<sup>c, d</sup>, Д. В. Авдошина<sup>e</sup>, К. К. Кюрегян<sup>c, d</sup>,  
М. Г. Беликова<sup>c, e, f</sup>, И. Д. Троценко<sup>c, g</sup>

<sup>a</sup>Институт молекулярной биологии им. В. А. Энгельгардта Российской академии наук, Москва, 119991 Россия

<sup>b</sup>Национальный медицинский исследовательский центр онкологии им. Н. Н. Блохина, Москва, 115478 Россия

<sup>c</sup>Российский университет дружбы народов, Москва, 117198 Россия

<sup>d</sup>Институт вакцин и сывороток им. И. И. Мечникова, Москва, 105064 Россия

<sup>e</sup>Федеральный научный центр исследований и разработки иммунобиологических препаратов им. М. П. Чумакова  
Российской академии наук, Москва, 117218 Россия

<sup>f</sup>Научный центр эпидемиологии и микробиологии им. Н. Ф. Гамалеи  
Министерства здравоохранения Российской Федерации, Москва, 123098 Россия

<sup>g</sup>Московское онкологическое общество, Москва, 111123 Россия

\*e-mail: zelenovayeye@gmail.com

Поступила в редакцию 05.10.2023 г.

После доработки 28.12.2023 г.

Принята к публикации 29.12.2023 г.

Белки Е6 и Е7 вируса папилломы человека (ВПЧ) играют ключевую роль в онкогенезе папилломавирусной инфекции. Данные по изменчивости этих онкобелков ограничены, а факторы, влияющие на их изменчивость, малоизучены. Нами проанализирована вариабельность известных к настоящему времени последовательностей белков Е6 и Е7 ВПЧ типа 16 (ВПЧ16) с учетом их географического происхождения и года сбора образцов, а также направления их эволюции в основных географических регионах мира. Из базы данных GenBank NCBI 6 октября 2022 года были выгружены все последовательности, относящиеся к фрагментам генома ВПЧ16, кодирующим онкобелки Е6 и Е7. Образцы отфильтровали по следующим параметрам: последовательность включает хотя бы одну из двух интересующих рамок считывания целиком, известны дата сбора и страна происхождения. Всего отобрали 3651 полногеномную нуклеотидную последовательность, кодирующую белок Е6, и 4578 полногеномных нуклеотидных последовательностей, кодирующих белок Е7. Полученные после выборки и выравнивания нуклеотидные последовательности переводили в аминокислотные и анализировали с применением программ MEGA11, R, RStudio, Jmodeltest 2.1.20, BEAST v1.10.4, Fastcov и Biostrings. Наибольшая изменчивость первичной структуры белка Е6 зарегистрирована в позициях 17, 21, 32, 85 и 90 а.о., а для Е7 – в позициях 28, 29, 51 и 77 а.о. Выборка разделилась по географическому признаку на 5 гетерогенных групп: африканскую, европейскую, американскую, регион Юго-Западной и Южной Азии и регион Юго-Восточной Азии. Для ряда стран были выявлены уникальные аминокислотные замены (АК-замены) в белках Е6/Е7 ВПЧ16, предположительно, характерные для отдельных этнических групп. В основном они локализируются в сайтах известных В- и Т-клеточных эпитопов и относительно редко в структурно-функциональных доменах. Выявленные различия в АК-заменах у различных этнических групп и их колокализация с кластерами В- и Т-клеточных эпитопов позволяют предположить их возможную взаимосвязь с географическим распределением аллелей и гаплотипов главного комплекса гистосовместимости (HLA). Это может приводить к распознаванию различного набора В- и Т-клеточных эпитопов вируса, что обуславливает региональные различия в направлении эпитопного дрейфа. В результате филогенетического анализа нуклеотидных последовательностей, кодирующих белок Е6 ВПЧ16, выявлен общий предок, подтверждена региональная кластеризация последовательностей гена белка Е6 по набору наиболее распространенных АК-замен, а также обнаружены случаи реверсии отдельных АК-замен при изменении региона распространения вируса. Для белка Е7 аналогичный анализ был невозможен из-за высокой гомологии последовательностей. Ковариационный анализ совокупной выборки выявил отсутствие взаимосвязи между аминокислотными остатками в белке Е6, в белке Е7, а также между Е6 и Е7. Полученные в работе данные важны для разработки терапевтических вакцин против ВПЧ высокого канцерогенного риска.

**Ключевые слова:** вирус папилломы человека, белок Е6, белок Е7, первичная структура, эволюция, филогенетический анализ, ковариационный анализ, В-клеточные эпитопы, Т-клеточные эпитопы

DOI: 10.31857/S0026898424040036, EDN: INFZCJ

Сокращения: АК-замены – замены аминокислотных остатков;  
ВПЧ – вирус папилломы человека; РШМ – рак шейки матки.

Вирус папилломы человека (ВПЧ) – ДНК-содержащий вирус диаметром 45–55 нм, семейства Papillomaviridae. На сегодняшний день известно более 200 штаммов папилломавирусов, патогенных для человека, которые можно разделить на три основные группы по степени онкогенного риска [1]:

- неонкогенные типы (1, 2, 3 и 5);
- типы низкого канцерогенного риска (6, 11, 42, 43 и 44);
- типы высокого канцерогенного риска (16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 68, 73 и 82).

Штаммы ВПЧ высокого канцерогенного риска относятся к основным этиологическим агентам развития рака шейки матки (РШМ), а также других злокачественных новообразований: ротоглотки, полового члена, анальной области, вульвы и влагалища [2]. Продукты трансляции ранних генов: онкобелки E5, E6 и E7 – вовлечены в пролиферацию и малигнизацию инфицированных клеток.

Белок E6 принимает непосредственное участие в трансформации инфицированных клеток в неопластические. Он ингибирует остановку клеточного цикла и апоптоз путем блокады активности главного онкосупрессора – белка p53, – способствует потере полярности клеток и инициирует гиперплазию [3]. Белок E6 также подавляет интерфероновый ответ организма через взаимодействие с IRF3 (interferon regulatory factor 3) [4].

Онкобелок E7 способствует дестабилизации генома и накоплению хромосомных аберраций. Взаимодействие белка E7 с IRF1 (interferon regulatory factor 1) приводит к уклонению инфицированных клеток от интерферонового ответа иммунной системы организма и тем самым поддерживает длительную персистенцию ВПЧ [5], что считается ведущим звеном в ВПЧ-ассоциированном канцерогенезе. Также белок E7 ингибирует один из главных онкосупрессоров – белок RB1 (retinoblastoma 1), – способствуя бесконтрольному делению инфицированных клеток [6].

В отличие от капсидных белков L1 и L2, онкобелки E6 и E7 высококонсервативны [7]. Тем не менее, согласно ряду исследований, определенные аминокислотные замены (АК-замены) в этих белках ассоциированы с хронической персистенцией ВПЧ и последующей малигнизацией инфицированных клеток [8–10].

В силу их роли в ВПЧ-ассоциированном канцерогенезе онкобелки E6 и E7 рассматривают как основу для разработки терапевтических вакцин против ВПЧ высокого канцерогенного риска, поэтому большое значение имеет информация относительно их вариабельности и ее определяющих параметров.

Нами проведен анализ глобальной изменчивости первичной структуры онкобелков E6 и E7 ВПЧ типа 16 (ВПЧ16) и направлений их эволюции

с привязкой к основным географическим регионам мира. Выбор ВПЧ16 обусловлен тем, что он преобладает среди генотипов высокого канцерогенного риска и ассоциирован с развитием ряда злокачественных опухолей. Кроме того, полногеномные последовательности ВПЧ16 широко представлены в международных базах данных, что важно для корректного филогенетического анализа.

## ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

**Отбор образцов из баз данных.** Из международной базы GenBank NCBI (<https://www.ncbi.nlm.nih.gov/nucleotide>) 6 октября 2022 года были выгружены все имеющиеся на тот момент фрагменты генома ВПЧ16, кодирующие белки E6 и E7. С использованием программы MEGA11 (<https://www.megasoftware.net/>) были отобраны образцы, содержащие полную нуклеотидную последовательность в пределах хотя бы одной из двух интересующих нас рамок считывания (ORF). Полученные после выборки и выравнивания нуклеотидные последовательности были переведены в аминокислотные в программе MEGA11. Оценку стабильности генома проводили на основе сравнения с референсными последовательностями белков E6 и E7 ВПЧ16: NP\_041325.1 (GeneID: 1489078) и NP\_041326.1 (GeneID: 1489079) соответственно.

**Филогенетический анализ.** Для выявления популяционноспецифических АК-замен ВПЧ16 из исследуемой выборки были отобраны образцы с известной датой сбора биоматериала и страной происхождения. Всего была выгружена 3651 последовательность гена E6 и 4578 последовательностей гена E7. Для филогенетического анализа базы образцов проверяли на избыточность и исключали все последовательности, идентичность которых превышала 99.8%. Для построения филогенетических деревьев для гена E6 было отобрано 356 нуклеотидных последовательности, для E7 – 143.

Для оценки временного сигнала полученных выборок использовали программу TempEst (<https://www.beast.community/tempest>). Параметры для расчета модели замен для построения филогенетических деревьев выбирали с помощью программы Jmodeltest 2.1.20 (<https://www.mybiosoftware.com/tag/jmodeltest>). Оптимальным оказался вариант GTR (General Time Reversible) с Gamma, равным 4. Для построения филогенетического дерева использовали программы из пакета BEAST v1.10.4 ([https://www.beast.community/2018-11-14.10.4\\_released.html](https://www.beast.community/2018-11-14.10.4_released.html)). Вначале были сгенерированы тестовые запуски, позволившие на основании АИС-критерия (Akaike Information Criterion) выбрать оптимальный вариант комбинации параметров, а именно: Fixed local clock, Skyline Growth. Длина MCMC (Markov chain Monte Carlo) составила 100 млн генераций. Для дальнейшей визуализации использована

программа FigTree-1.4.4 (<https://www.mybiosoftware.com/figtree-1-3-1-produce-figures-phylogenetic-trees.html>). Выделение кластеров для последующего анализа было основано на принципе гомологии нуклеотидных последовательностей.

**Ковариационный анализ.** Для анализа ковариаций использовали пакет Fastcov (<http://www.yanlilab.github.io/fastcov>), а также пакет Biostrings (<https://www.doi.org/10.7554/eLife.08932.001>).

Данные по локализации В- и Т-клеточных эпитопов белков E6 и E7 были получены из международной базы данных открытого доступа IEDB Analysis Resource (<http://www.tools.iedb.org>, даты выгрузки – 02.07.2019 и 12.10.2022). Эпитопные карты строили в программе JalView 2.11.14 (<https://www.jalview.org/>).

**Статистический анализ.** Для выявления статистически значимых различий между группами сравнения применяли двусторонний тест Фишера. Различия считались значимыми при  $p < 0.05$ . Анализ проводили с помощью языка R с визуализацией через Rstudio (<https://www.posit.co/download/rstudio-desktop/>).

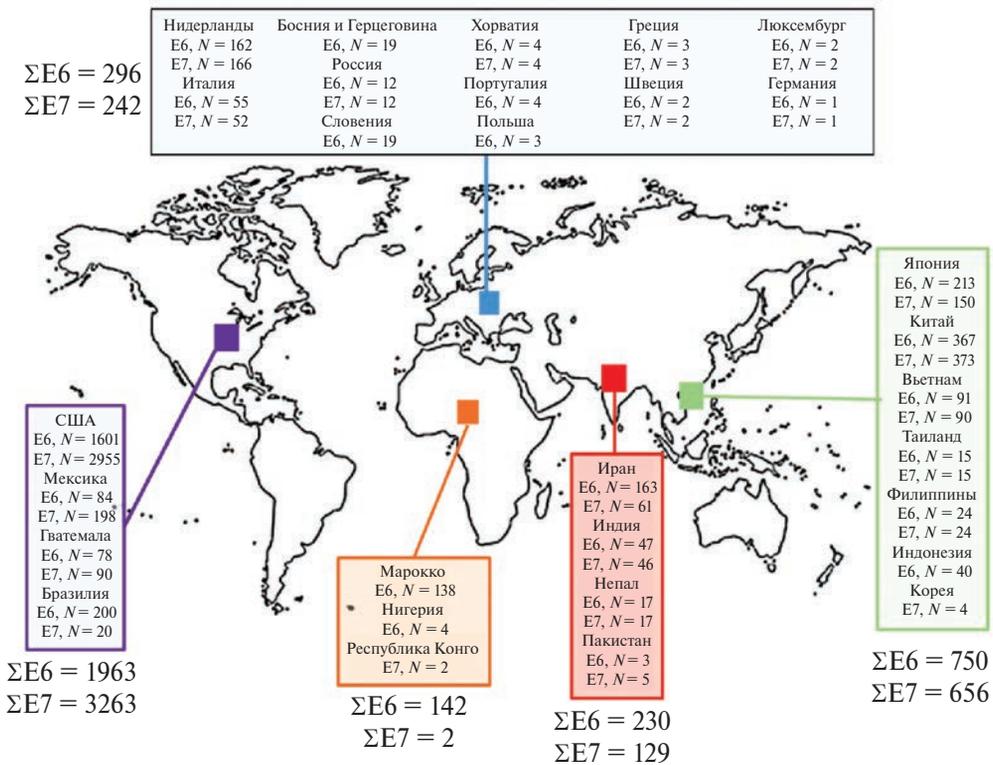
РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ

Общая характеристика вариабельности онкобелков E6 и E7

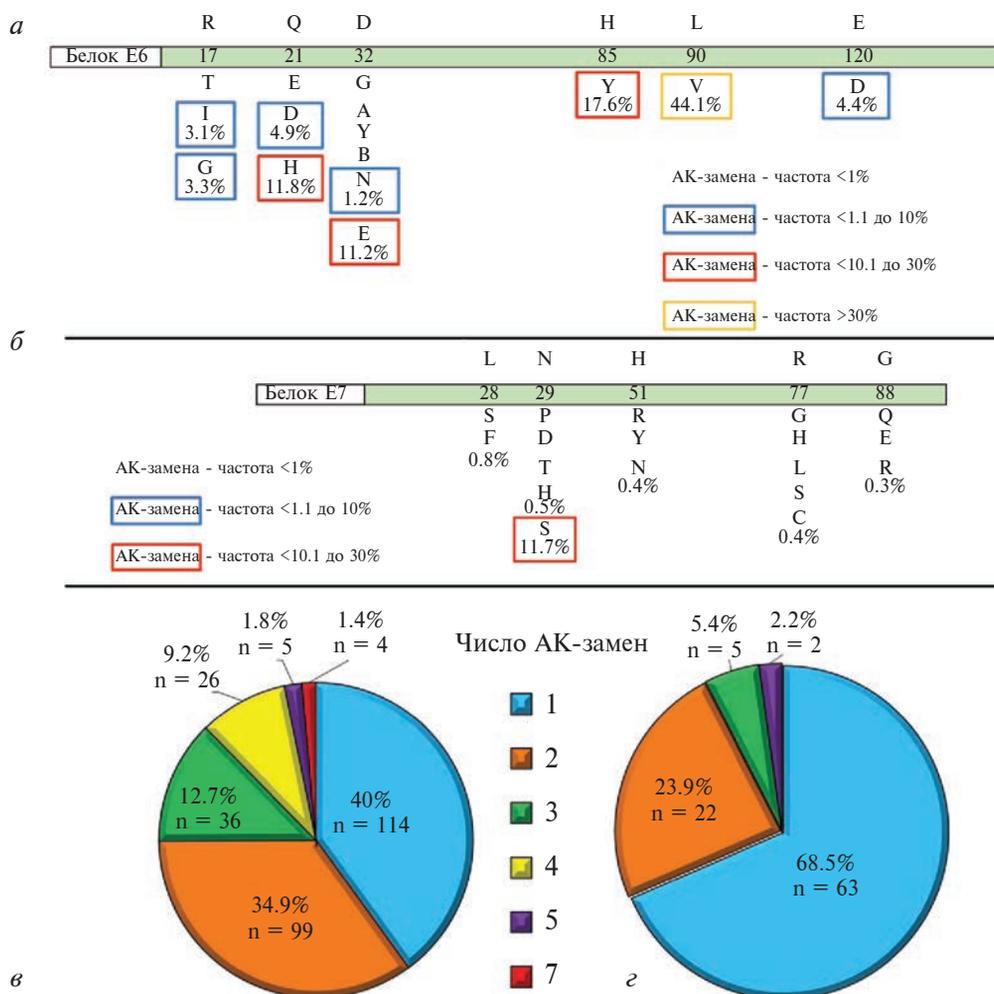
В общей сложности нами проанализирована 3651 нуклеотидная последовательность, кодирующая белок E6 ВПЧ16, из 29 стран и 4578 нуклеотидных последовательностей, кодирующих белок E7, из 23 стран. Все страны оказались распределены по 5 регионам мира: Европе, Африке, Америке, Юго-Западной и Южной Азии, а также Юго-Восточной Азии (рис. 1).

Из 158 аминокислотных остатков (а.о.) E6 39 оказались консервативными во всех анализируемых последовательностях, хотя в кодирующих их триплетах нуклеотидов встречались синонимичные замены. В остальных 119 а.о. белка E6 выявлены АК-замены.

Из 99 а.о. E7 43 оказались консервативными (100%) во всех анализируемых последовательностях, при этом в соответствующей им нуклеотидной последовательности встречались синонимичные замены. Остальные 56 кодонов содержали замены нуклеотидов, приводящие к АК-заменам. Частота наиболее распространенных АК-замен в белках E6 и E7 ВПЧ16 представлена на рис. 2а и 2б соответственно.



**Рис. 1.** Распределение отобранных образцов последовательностей генов белков E6 и E7 ВПЧ16 по географическим регионам: Европе (серая рамка), Африке (оранжевая рамка), Америке (фиолетовая рамка), Юго-Восточной Азии (зеленая рамка), Южной и Юго-Западной Азии (красная рамка). N – число последовательностей E6 и E7, депонированных в базе GenBank, с указанием представительства отдельных стран в каждом из регионов.



**Рис. 2.** Изменчивость онкобелков E6 и E7 ВПЧ16. Аминокислотные замены, встречающиеся в >1% последовательностей E6 (а) и E7 (б); распределение образцов по числу АК-замен в обогащенных выборках белка E6 (284 образца из 356 последовательностей) (в) и белка E7 (92 образца из 132 последовательностей) (г). Обогащенные выборки исключали последовательности, идентичность которых превышала 99.8%. Цветовое обозначение числа АК-замен общее для панелей (в) и (г).

В 170 из 356 (47.8%) анализируемых образцов E6 выявлено не менее двух АК-замен (рис. 2в). Заметим, что среди последовательностей одновременно с двумя АК-заменами наиболее частым было их сочетание в позициях 21 и 85, а также различные тандемы с аминокислотой валин в позиции 90. Стоит отметить, что сочетание замен в позициях 85 и 90 встречалось крайне редко. Среди трех и более замен чаще встречалась комбинация Q21H, H85Y и L90V.

Для последовательностей белка E7 подобных сочетаний не выявлено, в большинстве образцов АК-замены были единичные (рис. 2г).

#### *Вариабельность по отдельным географическим регионам*

На втором этапе был проведен анализ нуклеотидных последовательностей для выявления географических и/или популяционных особенностей

белков E6 и E7 ВПЧ16 на основании образцов, для которых была известна страна и год сбора. Распределение АК-замен в популяции указывало на кластеризацию последовательностей белков E6 и E7 ВПЧ16 по географическому признаку. Выявлено, что характер вариабельности различался в 5 географических группах: страны Европы (EU), Африки (AF), Америки (AM), Юго-Западной и Южной Азии (AS1) и Юго-Восточной Азии (AS2). Значимыми АК-заменами для региона считали те, частота встречаемости которых в данной позиции достоверно превышала таковую в других позициях этой выборки.

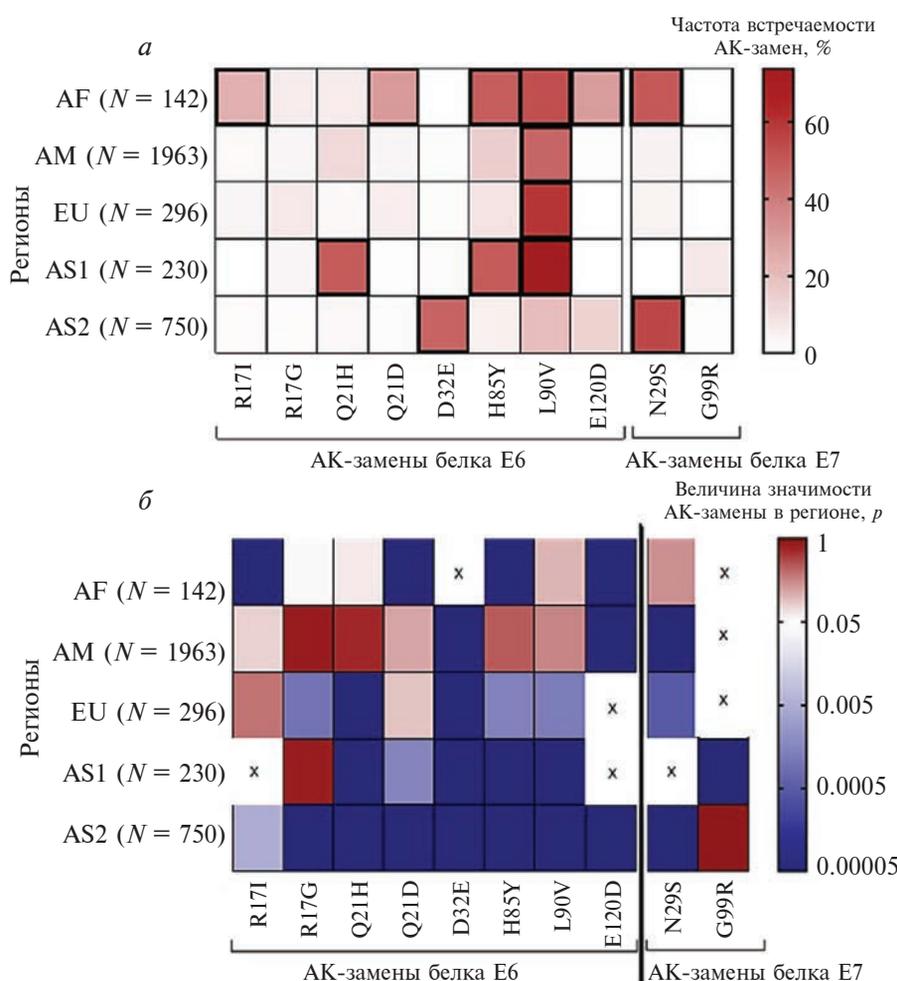
Так, в странах Африки (AF) в белке E6 в 52.8% случаев в позиции 90 преобладала замена L→V ( $p = 0.0306$ ). Частыми в этой географической группе были также замены H85Y (48.6%), E120D (30.3%), Q21D (30.9%) и R17I (23.9%) ( $p < 0.0001$  для всех вариантов; табл. 1, рис. 3).

**Таблица 1.** Частота встречаемости АК-замен в белке E6 изолятов ВПЧ16 из стран Африки (выборка AF,  $N = 142$ )

АК-позиция <sup>a</sup>	АК-замены <sup>b</sup>					
	Вариант 1		Вариант 2		Вариант 3	
	АК	n (%)	АК	n (%)	АК	n (%)
R17	G	9 (6.3%)	I	<b>34 (23.9%)</b>	T	2 (1.4%)
Q21	D	<b>44 (31%)</b>	E	2 (1.4%)	H	9 (6.3%)
A68	G	3 (2.1%)	–	–	–	–
D71	E	11 (7.7%)	–	–	–	–
H85	Y	<b>69 (48.6%)</b>	–	–	–	–
L90	V	<b>75 (52.8%)</b>	–	–	–	–
E120	D	<b>43 (30.3%)</b>	–	–	–	–

Здесь и далее: <sup>a</sup>В качестве референсной аминокислотной последовательности белка E6 выбрана NCBI Reference Sequence NP\_041325.1.

<sup>b</sup>Таблица не включает АК-замены, встречающиеся один раз. Приведены варианты АК-замен, их число (n) и частота встречаемости (%). АК-замены, встречающиеся в >10% последовательностей, выделены жирным шрифтом.



**Рис. 3.** Тепловые карты встречаемости АК-замен в пяти основных географических регионах мира. Карта распределения частот встречаемости: паттерн АК-замен в белках E6 и E7, цветовая шкала справа отражают частоту встречаемости замены для региона; обведены ячейки с частотой встречаемости >20% (а). Карта значимости АК-замены в белках E6 и E7 для региона при попарном сравнении с частотой встречаемости данной замены в общей популяции. Цветовая шкала справа отражает величину значимости,  $p$ . Сравнение проводили с использованием критерия Фишера (б). Тепловые карты сделаны с помощью языка R с визуализацией через Rstudio. Отметка "x" означает отсутствие АК-замены в регионе, поэтому сравнение не проводили.

**Таблица 2.** Частота встречаемости АК-замен в белке Е6 изолятов ВПЧ16 из США и стран Центральной и Южной Америки (выборка АМ,  $N = 1963$ )

АК-позиция	АК-замены									
	Вариант 1		Вариант 2		Вариант 3		Вариант 4		Вариант 5	
	АК	n (%)	АК	n (%)	АК	n (%)	АК	n (%)	АК	n (%)
M1	L	23 (1.17%)	-	-	-	-	-	-	-	-
K4	E	2 (0.1%)	N	1 (0.05%)	Q	1 (0.05%)	-	-	-	-
T6	P	4 (0.2%)	-	-	-	-	-	-	-	-
Q10	E	2 (0.1%)	-	-	-	-	-	-	-	-
R17	G	66 (3.36%)	I	43 (2.2%)	T	19 (0.97%)	-	-	-	-
K18	T	4 (0.2%)	-	-	-	-	-	-	-	-
Q21	H	<b>228 (11.6%)</b>	D	66 (3.36%)	-	-	-	-	-	-
Q27	R	3 (0.15%)	L	1 (0.05%)	-	-	-	-	-	-
H31	N	5 (0.25%)	Y	4 (0.2%)	-	-	-	-	-	-
D32	E	27 (1.4%)	N	22 (1.12%)	H	3 (0.15%)	A	1 (0.05%)	G	1 (0.05%)
I34	R	33 (1.68%)	L	10 (0.5%)	T	2 (0.1%)	K	1 (0.05%)	-	-
L35	V	3 (0.15%)	-	-	-	-	-	-	-	-
E36	Q	18 (0.9%)	K	1 (0.05%)	G	1 (0.05%)	A	1 (0.05%)	-	-
R47	H	2 (0.1%)	C	1 (0.05%)	-	-	-	-	-	-
E48	K	3 (0.15%)	-	-	-	-	-	-	-	-
D51	N	6 (0.3%)	-	-	-	-	-	-	-	-
I59	V	11 (0.56%)	-	-	-	-	-	-	-	-
D63	H	1 (0.05%)	Y	1 (0.05%)	-	-	-	-	-	-
N65	S	2 (0.1%)	-	-	-	-	-	-	-	-
D71	N	6 (0.3%)	E	1 (0.05%)	-	-	-	-	-	-
S78	C	5 (0.25%)	-	-	-	-	-	-	-	-
H85	Y	<b>298 (15.2%)</b>	-	-	-	-	-	-	-	-
L90	V	<b>904 (46.1%)</b>	-	-	-	-	-	-	-	-
E120	D	3 (0.15%)	-	-	-	-	-	-	-	-
K128	E	2 (0.1%)	-	-	-	-	-	-	-	-
R136	K	2 (0.1%)	T	1 (0.05%)	-	-	-	-	-	-
R138	L	2 (0.1%)	-	-	-	-	-	-	-	-
S145	C	1 (0.05%)	F	1 (0.05%)	-	-	-	-	-	-
C147	S	2 (0.1%)	-	-	-	-	-	-	-	-
R148	T	3 (0.15%)	I	1 (0.05%)	-	-	-	-	-	-
Q157	R	1 (0.05%)	H	1 (0.05%)	-	-	-	-	-	-

В США и странах Центральной и Южной Америки (АМ) наиболее распространенными в белке Е6 были замены L90V (46.1%,  $p = 0.0020$ ), H85Y (15.2%,  $p = 0.0582$ ), а также Q21H, хотя частота ее встречаемости достоверно не превышала этот показатель в других позициях (11.6%,  $p = 0.2425$ ; табл. 2, рис. 3).

В европейских странах (ЕU) наиболее распространенной была замена L90V, встречающаяся в 60.6% образцов ( $p < 0.0001$ ; табл. 3, рис. 3).

В странах Юго-Западной и Южной Азии (AS1) в онкобелке Е6 замены L90V, H85Y и Q21H встречались в 73.9, 49.1 и 49.1% образцов соответственно ( $p < 0.0001$  для всех вариантов; табл. 4, рис. 3).

В странах Юго-Восточной Азии (AS2) в белке Е6 довольно часто встречались замены D32E (46.7%), L90V (20.5%) и E120D (13.5%), а в позиции 29 белка Е7 преобладал серин (55%) ( $p < 0.0001$  для всех вариантов; табл. 5, 6; рис. 3).

При анализе полученных данных замечено, что в разных регионах в одних и тех же позициях

**Таблица 3.** Частота встречаемости АК-замен в белке Е6 ВПЧ16-изолятов из стран Европы (выборка EU, N = 296)

АК-позиция	АК-замены					
	Вариант 1		Вариант 2		Вариант 3	
	АК	n (%)	АК	n (%)	АК	n (%)
R5	K	1 (0.3%)	S	1 (0.3%)	T	1 (0.3%)
T6	P	1 (0.3%)	S	1 (0.3%)	–	–
Q10	H	1 (0.3%)	E	1 (0.3%)	–	–
E14	D	1 (0.3%)	K	1 (0.3%)	–	–
R15	P	2 (0.6%)	–	–	–	–
R17	G	22 (7%)	I	11 (3.7%)	T	2 (0.6%)
Q21	D	17 (6%)	H	8 (2.4%)	–	–
D32	E	4 (1.2%)	N	1 (0.3%)	–	–
I34	L	1 (0.3%)	R	1 (0.3%)	–	–
E36	Q	1 (0.3%)	K	1 (0.3%)	G	1 (0.3%)
H85	Y	26 (8.8%)	–	–	–	–
L90	V	<b>176 (59.5%)</b>	–	–	–	–
R131	T	2 (0.6%)	–	–	–	–

встречаются различные аминокислоты. Так, в белке Е6 в африканской группе, а также в странах Юго-Восточной Азии в позиции 17 преобладала АК-замена R17I, в то время как в американской и европейской группах – R17G. В странах Юго-Западной и Южной

Азии в этой позиции встречался только вариант R17G (табл. 1–5).

Аналогично для позиции 21 белка Е6 в африканской и европейской группе преобладала АК-замена Q21D, в то время как в Америке (AM) и Азии (AS1 и AS2) – Q21H (табл. 1–5).

**Таблица 4.** Частота встречаемости АК-замен в белке Е6 ВПЧ-изолятов из стран Юго-Западной и Южной Азии (выборка AS1, N = 230)

АК-позиция	АК-замены					
	Вариант 1		Вариант 2		Вариант 3	
	АК	n (%)	АК	n (%)	АК	n (%)
D11	H	2 (0.87%)	–	–	–	–
R17	G	8 (3.5%)	–	–	–	–
Q21	H	<b>113 (49.1%)</b>	D	1 (0.4%)	–	–
D32	E	3 (1.3%)	H	1 (0.4%)	N	4 (1.7%)
E48	G	2 (0.87%)	–	–	–	–
A53	G	2 (0.87%)	–	–	–	–
F54	V	2 (0.87%)	–	–	–	–
S78	C	5 (2.2%)	–	–	–	–
R84	S	2 (0.87%)	–	–	–	–
H85	Y	<b>113 (49.1%)</b>	–	–	–	–
L90	V	<b>170 (73.9%)</b>	–	–	–	–
P102	L	2 (0.87%)	A	1 (0.4%)	–	–
D105	N	2 (0.87%)	–	–	–	–
L106	V	2 (0.87%)	–	–	–	–
Q114	K	2 (0.87%)	–	–	–	–
R131	S	1 (0.4%)	T	1 (0.4%)	–	–
E155	K	2 (0.87%)	–	–	–	–

**Таблица 5.** Частота встречаемости АК-замен в белке E6 ВПЧ16-изолятов в странах Юго-Восточной Азии (выборка AS2, N = 750)

АК-позиция	АК-замены							
	Вариант 1		Вариант 2		Вариант 3		Вариант 4	
	АК	n (%)	АК	n (%)	АК	n (%)	АК	n (%)
M1	K	2 (0.26%)	-	-	-	-	-	-
R5	G	22 (2.93%)	-	-	-	-	-	-
M8	R	2 (0.26%)	K	3 (0.4%)	L	1 (0.13%)	-	-
Q10	E	2 (0.26%)	-	-	-	-	-	-
R17	G	5 (0.67%)	I	9 (1.2%)	T	1 (0.13%)	-	-
K18	N	1 (0.13%)	Q	1 (0.13%)	-	-	-	-
L19	F	1 (0.13%)	V	1 (0.13%)	-	-	-	-
Q21	H	22 (2.93%)	D	9 (1.2%)	-	-	-	-
T29	S	3 (0.4%)	-	-	-	-	-	-
D32	E	<b>350 (46.67%)</b>	H	1 (0.13%)	N	8 (1.07%)	Y	1 (0.13%)
I34	R	5 (0.67%)	L	2 (0.26%)	T	2 (0.26%)	M	2 (0.26%)
L35	V	9 (1.2%)	-	-	-	-	-	-
E36	Q	6 (0.8%)	-	-	-	-	-	-
R46	Q	2 (0.26%)	-	-	-	-	-	-
D51	E	2 (0.26%)	-	-	-	-	-	-
R62	K	4 (0.53%)	I	2 (0.26%)	-	-	-	-
N65	S	28 (3.7%)	-	-	-	-	-	-
P66	A	2 (0.26%)	S	1 (0.13%)	-	-	-	-
D71	N	1 (0.13%)	E	6 (0.8%)	-	-	-	-
F76	L	3 (0.4%)	-	-	-	-	-	-
H85	Y	41 (5.5%)	-	-	-	-	-	-
L90	V	<b>154 (20.5%)</b>	-	-	-	-	-	-
D105	H	2 (0.26%)	-	-	-	-	-	-
N112	Y	2 (0.26%)	-	-	-	-	-	-
C118	G	2 (0.26%)	-	-	-	-	-	-
E120	D	<b>101 (13.5%)</b>	-	-	-	-	-	-
R151	T	2 (0.26%)	-	-	-	-	-	-

Следует заметить, что при сравнении частот встречаемости наиболее значимых замен по регионам замена N29S в белке E7, широко представленная в AS2-регионе (55%), обнаружена также в американской и европейской группах (4.1 и 4.5% образцов соответственно; табл. 6), в то время как в регионе AS1, расположенном географически намного ближе к AS2, эта замена не встречалась вообще. Это позволяет предположить, что некоторые АК-замены в региональных образцах ВПЧ16 могут быть связаны с глобальными процессами миграции населения.

Частота встречаемости АК-замен значительно различалась внутри регионов (табл. 1–6). Региональное распределение замен проиллюстрировано на рис. 4, анализ частот распространения замен в регионах – тепловыми картами на рис. 3 (см. также табл. S1,

рис. S1, рис. S2, Дополнительные материалы на сайте [http://www.molecbio.ru/downloads/2024/4/supp\\_Zelenova\\_rus.pdf](http://www.molecbio.ru/downloads/2024/4/supp_Zelenova_rus.pdf), а также доступны по DOI: 10.31857/S0026898424040036). В белке E6 ВПЧ16 замена R17I достоверно чаще встречалась в AF-выборке ( $p < 0.0001$ ), Q21H – в AS1, Q21D – в AF (по сравнению с другими выборками  $p < 0.0001$ ; рис. 3, рис. 4). Замена D32E чаще всего встречалась в группе AS2, частоты ее встречаемости в других выборках не различались. Замена H85Y была одинаково распространена в группах AF и AS1, немного реже в выборке AM и достоверно реже в EU и AS2. Замена L90V была широко распространена во всех выборках, кроме AS2. Замена E120D была характерна только для изолятов AF и AS2 (рис. 3). Анализ региональных выборок последовательностей высококонсервативного белка E7 выявил высокую частоту встречаемости

**Таблица 6.** Частота встречаемости АК-замен в белке E7 ВПЧ-изолятов в различных регионах мира

АК-позиция <sup>a</sup>	АК-замены <sup>b</sup>							
	Вариант 1		Вариант 2		Вариант 3		Вариант 4	
	АК	n (%)	АК	n (%)	АК	n (%)	АК	n (%)
<b>Африканская группа (AF, N = 2)</b>								
N29	S	<b>1 (50%)</b>	-	-	-	-	-	-
<b>Американская группа (AM, N = 3263)</b>								
D4	N	2 (0.06%)	-	-	-	-	-	-
N29	S	133 (4.1%)	P	1 (0.03%)	H	1 (0.03%)	-	-
E33	K	3 (0.09%)	-	-	-	-	-	-
A42	T	2 (0.06%)	V	1 (0.03%)	-	-	-	-
H51	R	1 (0.03%)	N	13 (0.4%)	-	-	-	-
D62	H	1 (0.03%)	N	1 (0.03%)	-	-	-	-
R77	C	2 (0.06%)	G	1 (0.03%)	S	1 (0.03%)	-	-
P92	L	2 (0.06%)	S	1 (0.03%)	-	-	-	-
<b>Европейская группа (EU, N = 242)</b>								
N29	S	11 (4.5%)	-	-	-	-	-	-
D75	N	2 (0.83%)	-	-	-	-	-	-
<b>Азиатская группа 1 типа (AS1, N = 129)</b>								
L15	S	2 (1.6%)	-	-	-	-	-	-
T20	S	2 (1.6%)	-	-	-	-	-	-
G88	R	10 (7.8%)	-	-	-	-	-	-
I93	T	2 (1.6%)	-	-	-	-	-	-
C94	S	2 (1.6%)	-	-	-	-	-	-
<b>Азиатская группа 2 типа (AS2, N = 656)</b>								
D14	H	2 (0.3%)	-	-	-	-	-	-
L28	F	29 (4.4%)	-	-	-	-	-	-
N29	S	<b>361 (55%)</b>	H	20 (3%)	D	1 (0.15%)	T	1 (0.15%)
D39	N	1 (0.15%)	H	1 (0.15%)	-	-	-	-
H51	N	2 (0.3%)	Y	1 (0.15%)	-	-	-	-
S63	A	2 (0.3%)	F	6 (0.9%)	-	-	-	-
D75	Y	2 (0.3%)	-	-	-	-	-	-
R77	C	17 (2.6%)	L	1 (0.15%)	H	1 (0.15%)	-	-
G88	Q	1 (0.15%)	E	1 (0.15%)	R	1 (0.15%)	-	-
P92	S	2 (0.3%)	-	-	-	-	-	-
Q96	R	1 (0.15%)	H	1 (0.15%)	-	-	-	-

<sup>a</sup>В качестве референсной аминокислотной последовательности белка E7 выбрана NCBI Reference Sequence NP\_041326.1.

<sup>b</sup>Таблица не включает АК-замены, встречающиеся один раз. Приведены варианты АК-замен, их число (n) и частота встречаемости (%). АК-замены, встречающиеся в >10% последовательностей, выделены жирным шрифтом.

замены N29S в выборке AS2, причем достоверно чаще, чем в AM и EU; в то время как в группе AS1 эта замена не обнаружена (рис. 3, рис. 4). Замена G88R в белке E7 наблюдалась только в выборках AS2 и AS1, причем в AS1 чаще, чем в AS2 ( $p < 0.0001$ ; рис. 3, рис. 4).

Распределение образцов белка E6 ВПЧ16 внутри одного региона отличалось неравномерностью

(характерно для AM, AF, EU). Несмотря на это, сравнение спектров АК-замен в стране, лидирующей по числу образцов, и в других странах каждого из этих регионов достоверно значимых различий не выявило. Заметим, что для ряда стран, в том числе с немногочисленной выборкой, в дополнение к регионспецифичным АК-заменам находили 1–2 уникальные (табл. 7, рис. 3). Различия в частоте

распространения АК-замен обнаружены и между отдельными провинциями Китая. Так, в центральных и восточных преобладали замены, характерные для региона AS2 (N29S), с минимальной представленностью замены L90V. Для северо-западных провинций паттерн АК-замен оказался близким к европейскому: высокая частота L90V с низкой встречаемостью “азиатской” N29S (рис. 4, табл. 7). В целом, на основании полученных результатов можно сделать вывод о циркуляции в различных географических локациях регионспецифичных вариантов ВПЧ16, различающихся по паттернам АК-замен в онкобелках E6 и E7.

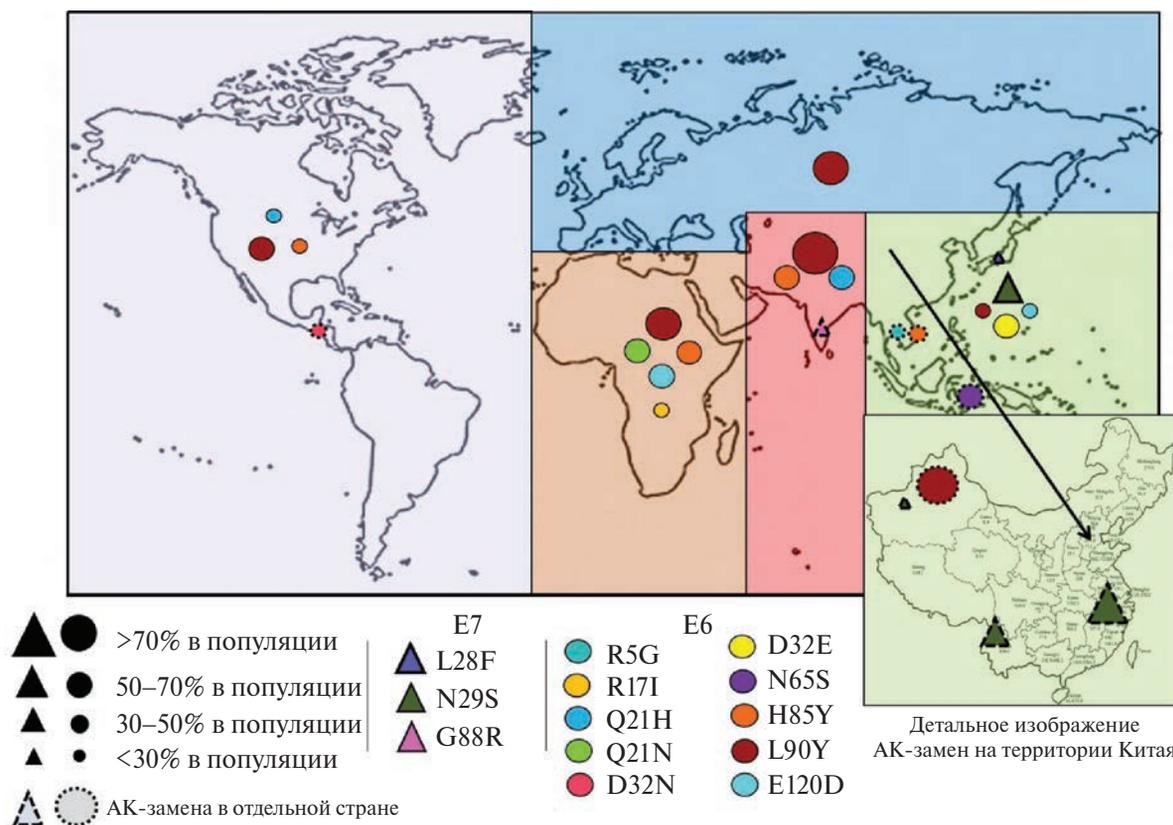
#### Ковариационный анализ нуклеотидных и аминокислотных замен

Присутствие АК-замен в определенных сочетаниях позволяет предположить наличие ковариационных связей (сцепления) между аминокислотными остатками и/или кодирующими их нуклеотидами. Для проверки этой гипотезы проведен ковариационный анализ нуклеотидных и аминокислотных последовательностей E6 и E7.

При анализе выборки из 356 нуклеотидных последовательностей, кодирующих белок E6 ВПЧ16, выявлено два паттерна сцепления четырех

нуклеотидов: 63T-204A-207G-253T с вероятностью 0.9312 и 63G-204T-207A-253C с вероятностью 0.9795 (этот паттерн входит в состав референсной последовательности). Замены нуклеотидов в позициях 204 и 207 были синонимичными. Замена нуклеотидов G63T определяла две самые встречаемые АК-замены в позиции 21: Q21H и Q21D, – а замена 253 нуклеотида – одну из распространенных АК-замен – H85Y. На основании вышеизложенного можно предположить сцепление аминокислотных остатков в позициях 21 и 85. Однако анализ ковариаций полного набора аминокислотных последовательностей белка E6 не подтвердил наличия групп сцепления по отдельным АК-заменам, в частности в положениях 21 и 85; сцепленность сохранилась только на уровне E6-кодирующей ДНК.

При проведении анализа ковариаций для полной выборки нуклеотидных последовательностей, кодирующих белок E6, помимо двух вышеописанных паттернов выявлены еще два дополнительных – со сцеплением трех нуклеотидов: 50G-61C-321A с вероятностью 0.9937 (этот паттерн входит в состав референсной последовательности) и 50T-61G-321G с вероятностью 0.8377. Важно заметить, что сцепленность зарегистрирована



**Рис. 4.** Распределение АК-замен, встречающихся в >10% последовательностей выборок по пяти географическим регионам, включая АК-замены в отдельных странах. Фиолетовый блок – США, страны Центральной и Южной Америки; синий блок – страны Европы; оранжевый блок – страны Африки; красный блок – страны Юго-Западной и Южной Азии; зеленый блок – страны Юго-Восточной Азии.

Таблица 7. Распределение АК-замен в белках E6 и E7 в отдельных странах\*

Страна	Варианты АК-замен в белке E6		Варианты АК-замен в белке E7	
	АК	n (%)	АК	n (%)
Регион AS1			G88R	10 (7.8%)
Индия (N = 46)			G88R	10 (21.7%) <i>p</i> < 0.0001
Регион AM	D32N	22 (1.12%)		
Гватемала (N = 78)	D32N	11 (14.1%) <i>p</i> < 0.0001		
Регион AS2			L28F	29 (4.4%)
Япония (N = 150)			L28F	26 (17.3%) <i>p</i> < 0.0001
Регион AS2	N65S	28 (3.7%)		
Индонезия (N = 40)	N65S	13 (32.5%) <i>p</i> < 0.0001		
Регион AS2	R5G H85Y	22 (2.93%) 41 (5.5%)		
Вьетнам (N = 91)	R5G H85Y	18 (19.7%) <i>p</i> < 0.0001 10 (10.9%) <i>p</i> = 0.0233		
Регион AS2	D32E L90V	350 (46.67%) 154 (20.5%)	N29S	361 (55%)
Китай, центральные и восточные регионы (N = 154)			N29S	115 (74.7%)
Китай, юго-западный регион (N = 47)			N29S	25 (53.6%)
Китай, северный регион (E6 N = 46) (E7 N = 50)	D32E L90V	2 (4.3%) 36 (78.2%)	N29S	3 (6%)

\*Указаны АК-замены, уникальные для страны в сравнении с регионом. По умолчанию все остальные АК-замены в указанной стране встречаются с той же частотой, что и в других странах данного региона (пустая строка).

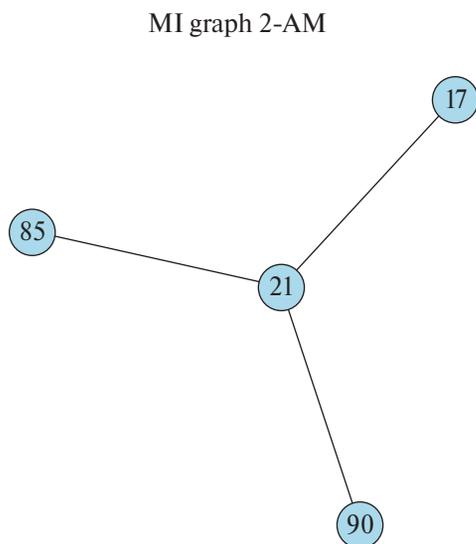
на уровне исключительно нуклеотидной последовательности (для аминокислот ковариаций не выявлено).

Анализ ковариаций выполнен также для нуклеотидных и аминокислотных последовательностей E7 ВПЧ16. В результате не выявлено групп сцепления ни по нуклеотидным, ни по отдельным АК-заменам.

Анализ проводили на всех образцах – без разделения по регионам мира. Отсутствие групп сцепления может быть обусловлено различиями в сцеплениях между отдельными региональными выборками, с потерей достоверности сцеплений при объединении выборок. Для проверки этой гипотезы мы провели анализ ковариации (сцеплений) отдельно для каждой из 5 выборок (страны Европы, Африки, Америки, Юго-Западной и Южной Азии

и Юго-Восточной Азии). Анализ проводили с использованием второго биоинформатического метода (пакет Biostrings) с калькулятором MI (<http://www.yanlilab.github.io/fastcov>). Для образцов E6 американского происхождения (выборка AM) выявлена ковариация аминокислотных остатков в позициях 17, 21, 85 и 90 (рис. 5). Интересно отметить, что сцепленными оказались 4 из 6 варибельных позиций E6 (рис. 2a, табл. 2).

Таким образом, хотя в целом по всей выборке позиции подверженных заменам аминокислотных остатков в составе как E6, так и E7 не сцеплены между собой, при ограничении выборки одним географическим регионом удалось выявить сцепление аминокислотных остатков в 4 из 6 варибельных позиций белка E6.



**Рис. 5.** Ковариационный анализ АК-замен белка Е6 для образцов американского региона (АМ). Анализ проведен с использованием второго биоинформатического метода с применением пакета Biostrings с калькулятором MI (<http://www.yanlilab.github.io/fastcov>). В окружностях указаны позиции аминокислотных остатков, линии между ними – сцепление двух АК-замен, вероятность сцепления между которыми  $p < 0.05$  (MI).

*Реконструкция истории возникновения и последующего распространения вариантов гена Е6, приводящих к АК-заменам в белке Е6*

Для реконструкции истории возникновения и последующего распространения АК-замен ВПЧ16 был проведен филогенетический анализ 356 нуклеотидных последовательностей гена *Е6* (рис. 6). Из анализа были исключены все последовательности, идентичность которых превышала 99.8%. Аналогичное исследование предполагалось выполнить и для гена *Е7*, однако из-за высокой гомологии последовательностей временной сигнал не был зарегистрирован. Принцип кластеризации заключался в выделении отдельных клад на основании близости нуклеотидных последовательностей (традиционный принцип сродства последовательностей) с добавлением принципа географического единства после завершения построения дерева (“разрезание дерева”) для интерпретации полученных результатов.

В ходе анализа установлено, что общий предок известных современных вариантов ВПЧ16 имеет американское происхождение и к настоящему моменту разделился на 9 кластеров (рис. 6, табл. 8), различающихся филогенетически и/или по региону циркуляции. Первое крупное разделение на две принципиально разные ветви (группа образцов, отходящая от узла на основании возникновения ключевой замены нуклеотидов) произошло только в 1840–1850 гг. (HPD 95%: 1702–1948 гг.).

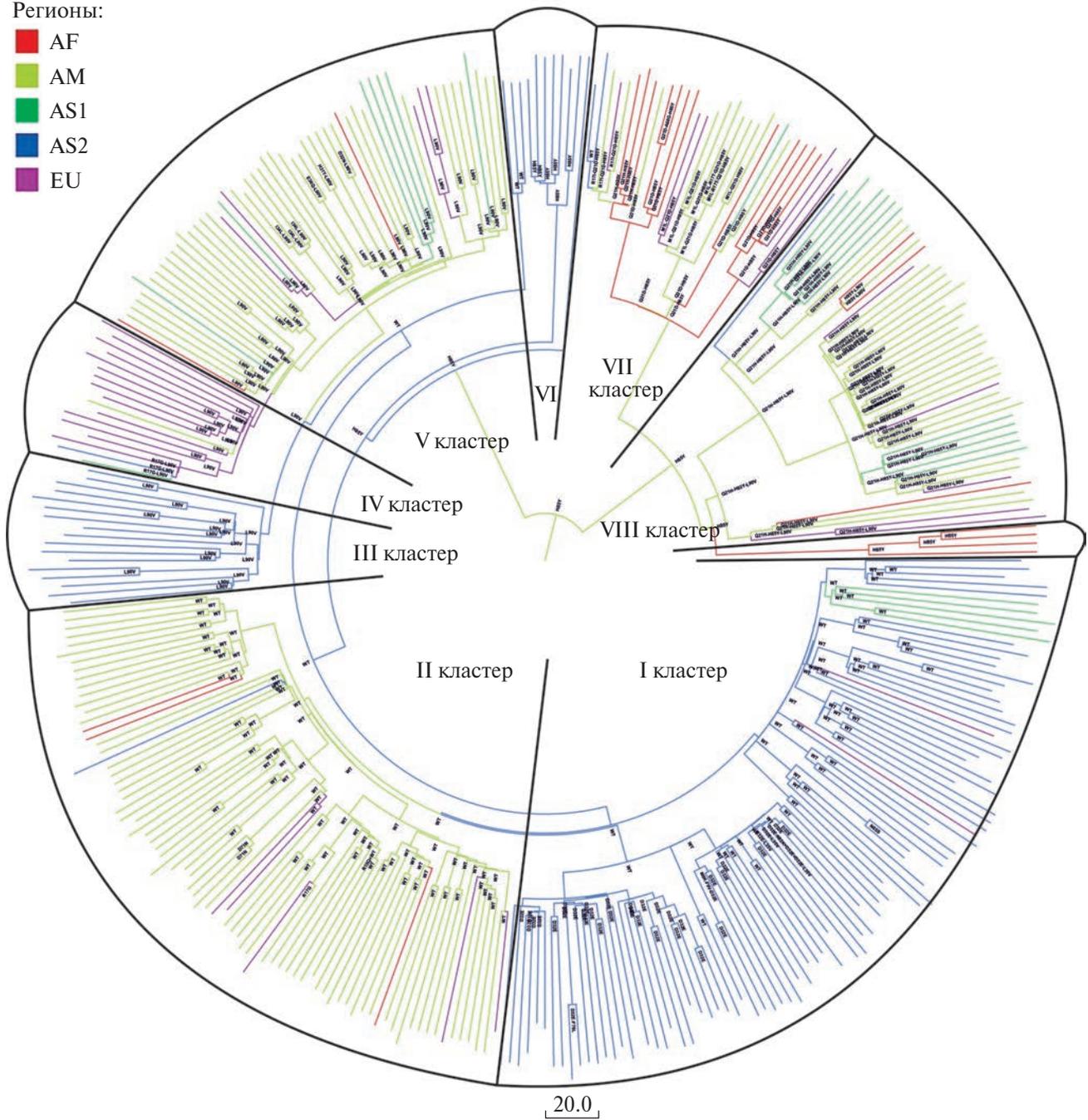
Первая ветвь, преимущественно американской выборки в 1880–1890 гг. (HPD 95%: 1796–1965 гг.), разделилась на 2 группы, при этом в каждой из них с интервалом примерно в 30 лет произошло независимое отделение африканских линий: в 1895 году (HPD 95%: 1852–1973 гг.) и в 1925 году (HPD 95%: 1873–1979 гг.) (рис. 6, кластеры VII и IX соответственно). Это разделение на клады (конечная группа образцов, не подлежащая дальнейшему делению, объединенная на основании максимального сходства нуклеотидной последовательности) подтвердили при оценке состава АК-замен в данных группах: для кластера IX это H85Y, в то время как изоляты кластера VII, кроме вышеуказанной, несут дополнительную – Q21D (табл. 8). Важно подчеркнуть, что изоляты этих кластеров сформировались в результате независимой эволюции ВПЧ16 разных линий. Интересны первичные структуры анализируемого фрагмента белка Е6 из кластера VIII, в котором большинство образцов содержит более одной АК-замены, при этом наиболее часто встречается сочетание трех замен: Q21H-H85Y-L90V (табл. 8; рис. 6, кластер VIII). Описывая эволюцию образцов ВПЧ16 первой ветви, стоит также заметить, что в дальнейшем из разных линий исходной американской группы со схожим набором АК-замен (Q21H, H85Y, L90V) примерно в 1950 году произошло одновременное выделение двух групп AS1 в кластере V и кластере VIII (HPD 95%: 1917–1984 гг. и 1930–1983 гг.) (рис. 6).

Для изолятов второй ветви, образованной от общего предка в 1840–1850 гг., характерна более четкая кластеризация по географическому признаку (рис. 6). Обращает внимание раннее выделение небольшой группы китайского происхождения из 7 образцов в 1905–1910 гг. (HPD 95%: 1822–1972 гг.) (рис. 6, кластер VI), что может быть обусловлено завозом штаммов ВПЧ16 из Китая и дальнейшим распространением в американской популяции. Эта гипотеза подтверждена при анализе аминокислотного состава этой группы изолятов, а именно сохранением замены H85Y; в то время как в большинстве образцов, происходящих от второй ветви, произошла реверсивная замена. В отличие от H85Y, замена L90V, возникшая еще у гипотетического предка, сохранилась как в американском, так и в азиатском кластере, что позволяет считать ее не связанной с этносом.

Также в 1920 году (HPD 95%: 1860–1975 гг.) произошло знаковое разделение изолятов второй ветви, что сопровождалось изменением аминокислотного состава – появлением L90V вне зависимости от географического распределения. Позднее, в 1930 году (HPD 95%: 1893–1978 гг.), произошло выделение группы AS2 (рис. 6). Примерно в это же время, в 1930 г. (HPD 95%: 1943–1964 гг.), среди обширной группы изолятов ВПЧ16 с белком Е6 без замены L90V четко выделилась группа AS2

Регионы:

- AF
- AM
- AS1
- AS2
- EU



**Рис. 6.** Филогенетическое дерево с использованием байесовского подхода с временной шкалой для последовательностей гена *E6*, длина фрагмента – 477 п.н. (координаты по эталонному образцу NC\_001526.4:7125–7601). Для каждой последовательности указан номер в базе данных GenBank, страна и год выделения. В узлах дерева приведены аминокислотные остатки, предсказанные для гипотетического предка каждой филогруппы. Обозначения: EU – страны Европы, AF – страны Африки, AM – страны Америки, AS1 – страны Юго-Западной и Южной Азии, AS2 – страны Юго-Восточной Азии.

Регионы:

- AF
- AM
- AS1
- AS2
- EU

Кластер I

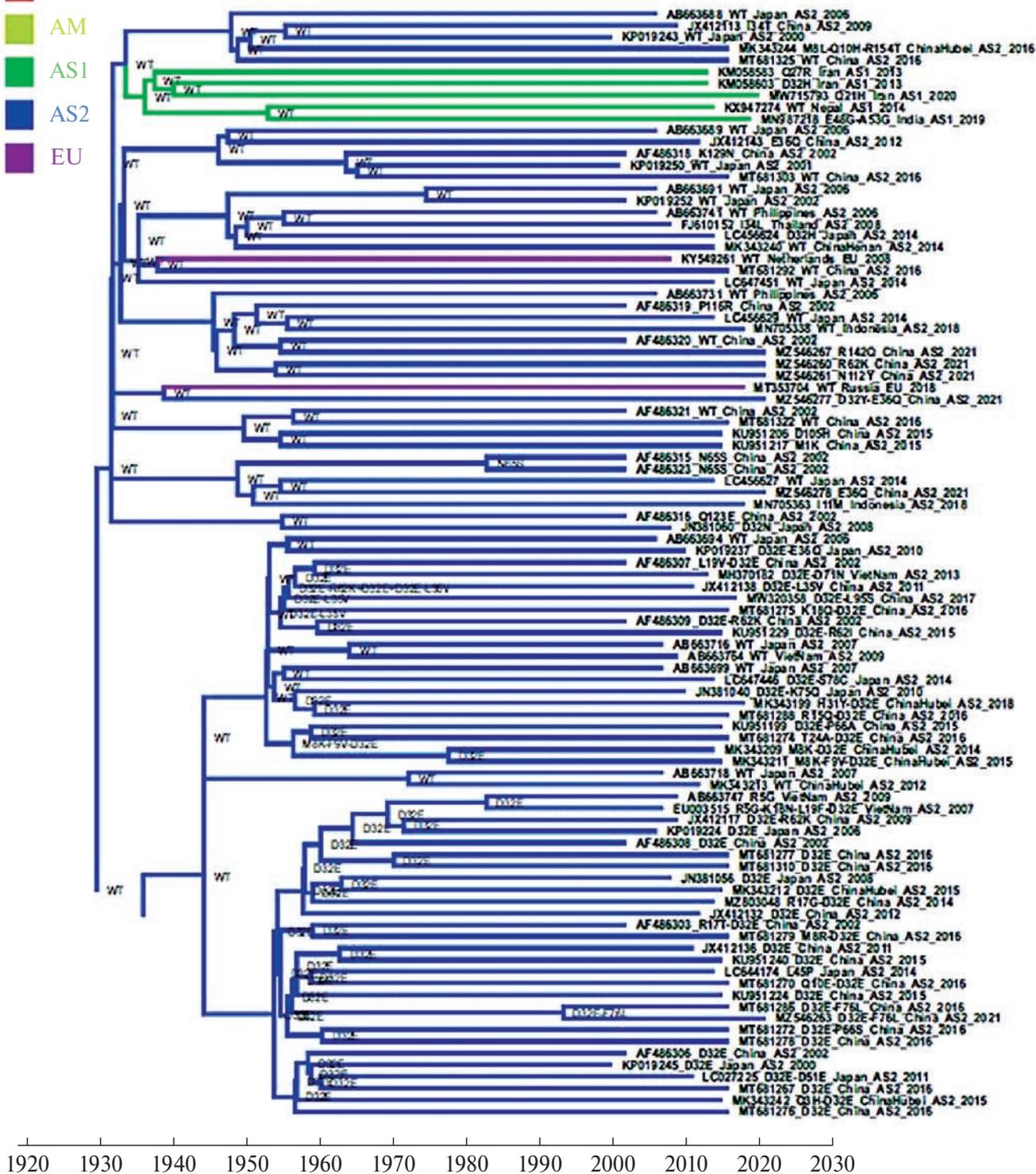


Рис. 6. Продолжение.

Регионы:

- AF
- AM
- AS1
- AS2
- EU

Кластер II

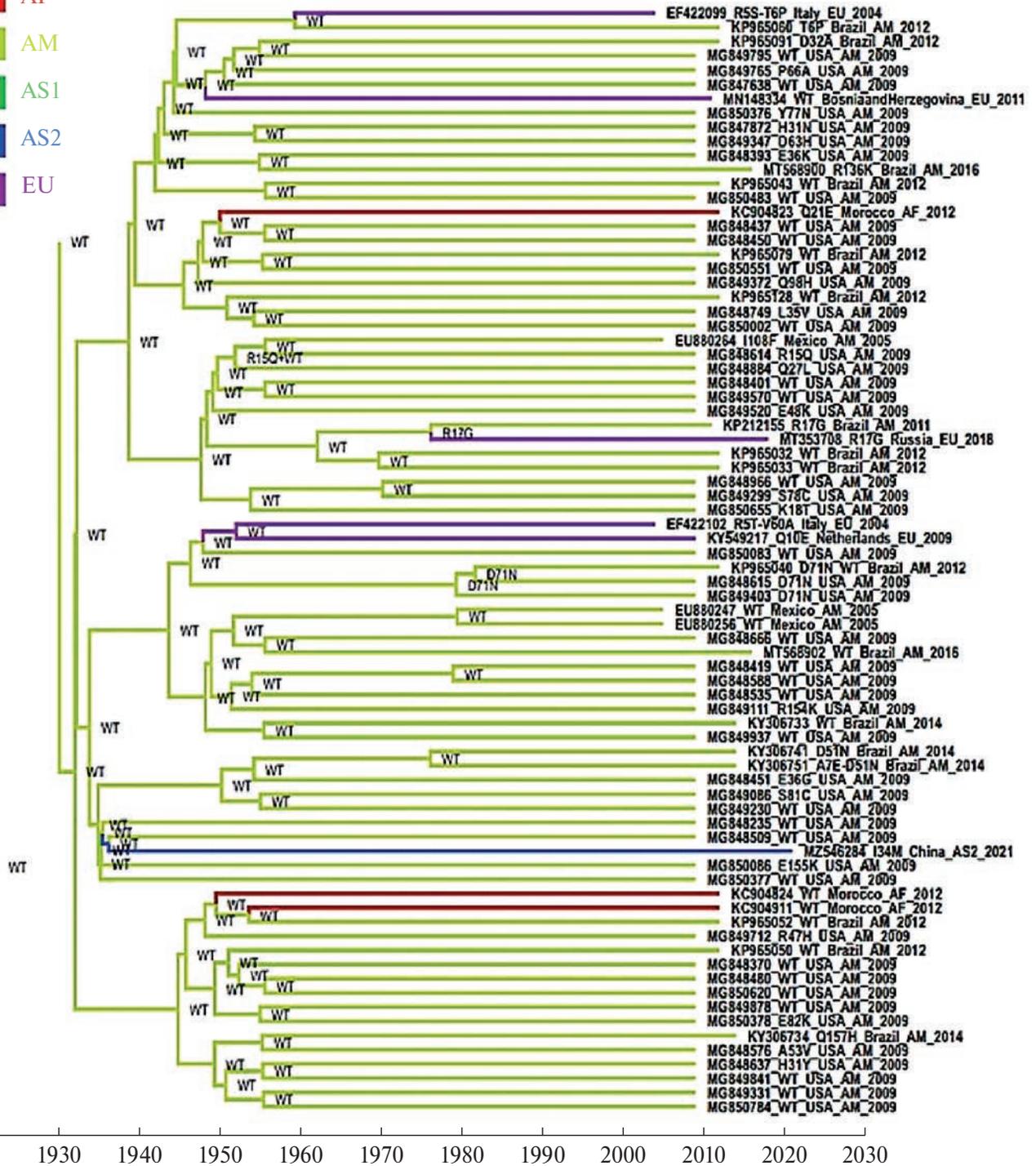
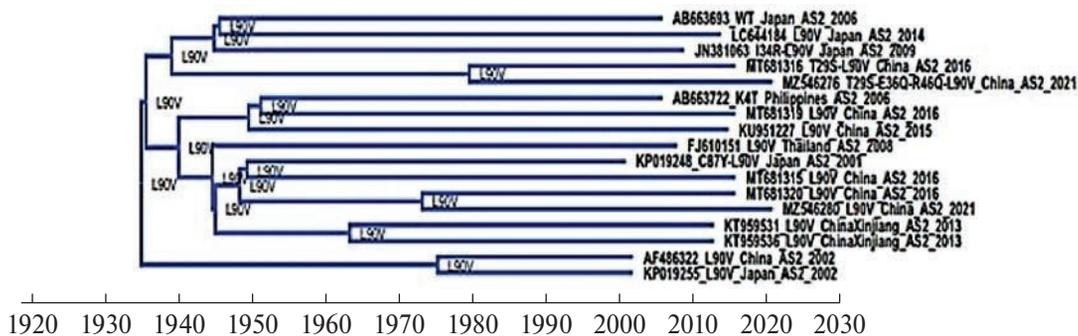


Рис. 6. Продолжение.

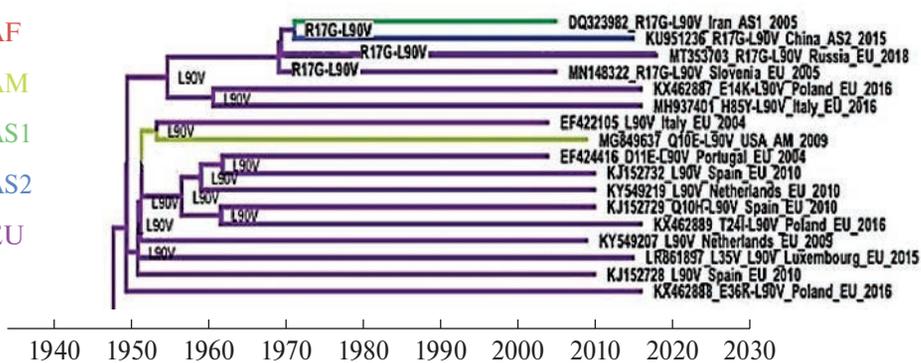
Кластер III



Кластер IV

Регионы:

- AF
- AM
- AS1
- AS2
- EU



Кластер V

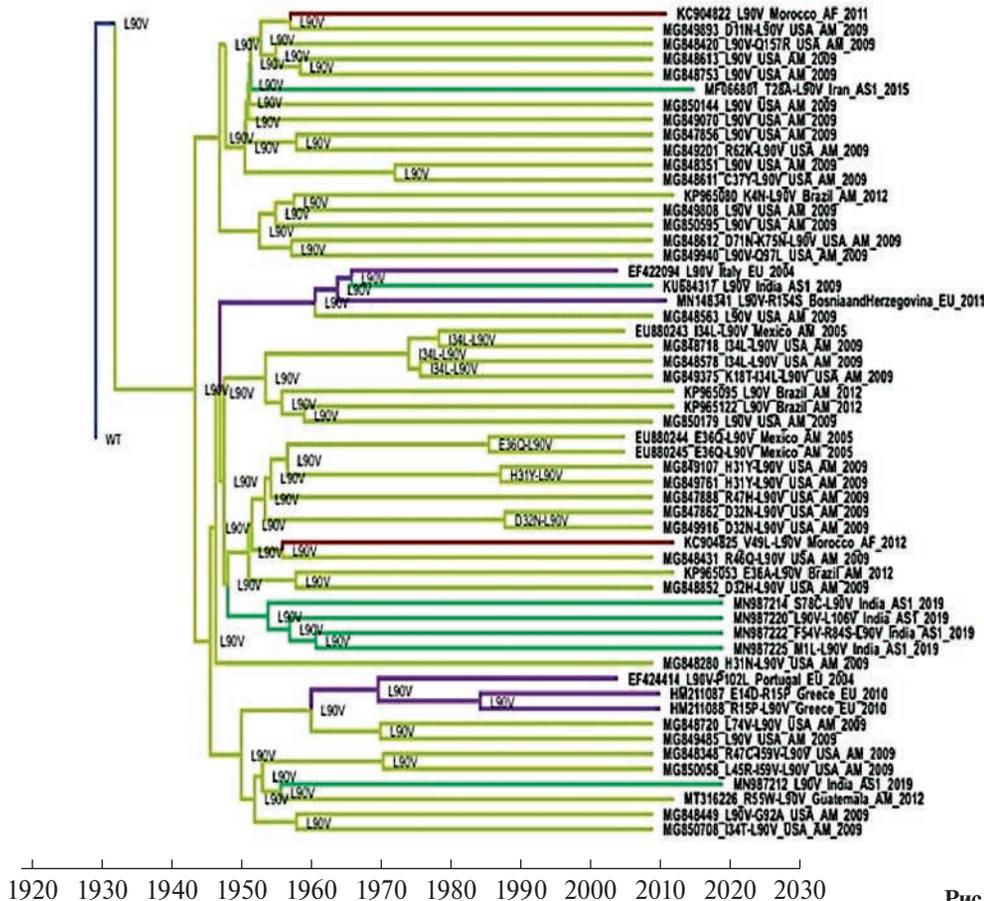


Рис. 6. Продолжение.



с группоспецифической заменой D32E и американская ветвь (рис. 6, кластеры I и II).

Параллельно с этим, среди изолятов, происходящих от общего предка с АК-заменой L90V, в 1930–1940 гг. (HPD 95%: 1894–1979 гг.) произошло четкое разделение на образцы американской группы и азиатской типа 2 (рис. 6, кластеры III и V). Как и среди изолятов первой ветви, из образцов американского происхождения второй ветви в 1950 году отделилась группа AS2 (рис. 6, кластер V). Примерно в это же время отделилась и европейская группа (HPD 95%: 1967–1969 гг.; рис. 6, кластер IV).

Описываемая нами эволюция ВПЧ16 на основе анализа фрагмента его генома, кодирующего белок E6, включает относительно небольшой временной интервал – продолжительностью всего 150 лет. По результатам исследования, АК-замена H85Y филогенетически оказалась самой древней, а скорость накопления АК-замен в исследованном фрагменте генома ВПЧ16 составляла  $10^{-5}$  замен на сайт в год. Основные АК-замены в белке E6 штаммов ВПЧ16, циркулирующих в человеческой популяции на современном этапе, возникли в течение последних 50 лет; причем самая “молодая” – F76L, обнаруженная в китайских изолятах конца XX века (1990 год, HPD 95%: 1966–2016 гг.).

#### *Аминокислотные замены в белках E6 и E7 ВПЧ16 и противовирусный иммунный ответ*

Различия в вариантах E6 и E7 ВПЧ16, циркулирующих в различных географических регионах (рис. 2), а также реверсии АК-замен при переходе ВПЧ16 в другой географический регион (рис. 6) могут быть связаны с этническими особенностями населения, в частности с различиями в экспрессии молекул главного комплекса гистосовместимости (HLA) и их аллелей. Последние ассоциированы с различной специфичностью В- и Т-клеточного противовирусного иммунного ответа. Это приводит к селекции различных вариантов вируса в той или иной человеческой популяции, которые успешно избегают иммунного ответа хозяина.

Для проверки этой гипотезы выполнено картирование эпитопов E6 и E7, взятых из международной базы данных открытого доступа. Для онкобелка E6 ВПЧ16 описано 126 эпитопов, для E7 – 142. В ходе анализа установлено, что практически все значимые АК-замены, описанные выше, входят в состав более 10 эпитопов (рис. 7, 8).

### ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Известно, что распространенность генотипов ВПЧ в разных регионах мира варьирует. В настоящее время условно выделяют “европейские” типы ВПЧ (16, 18, 31, 33, 45 и др.) и “азиатские” (52, 58, 59, 65 и др.) [11]. В результате секвенирования геномов ВПЧ выявлены популяционные различия

даже внутри одного генотипа вируса. В литературе описано 4 генотипические группы ВПЧ16: A, B, C и D [12], а также подгруппы A1–4, B1–2, C и D1–3, различающиеся по последовательности капсидного белка L1. Интересно, что эти подгруппы имеют четкую региональную представленность [13]. Вышеописанную классификацию активно используют в фундаментальных и прикладных научных исследованиях, так как именно на белках капсида основаны разработки профилактических ВПЧ-вакцин [14].

В проведенном нами исследовании акцент сделан на онкобелках E6 и E7 ВПЧ16. По сравнению с капсидными, белки E6 и E7 отличаются большей консервативностью. Именно эти вирусные белки реализуют канцерогенный потенциал вируса [7], который зависит, в том числе, от их первичной структуры и со временем появляющихся мутаций [8–10]. Именно с этими особенностями белков E6 и E7 связан тот факт, что секвенирование и филогенетический анализ этих фрагментов генома ВПЧ сфокусированы в основном на выявлении корреляций между полиморфизмами в генах E6 и E7, частотой случаев хронической персистенции вируса и последующей малигнизацией. И хотя полногеномные исследования ВПЧ проводят во многих странах, до сих пор не было сравнительного анализа выявленных вариаций в первичной структуре вирусных онкобелков по разным географическим регионам.

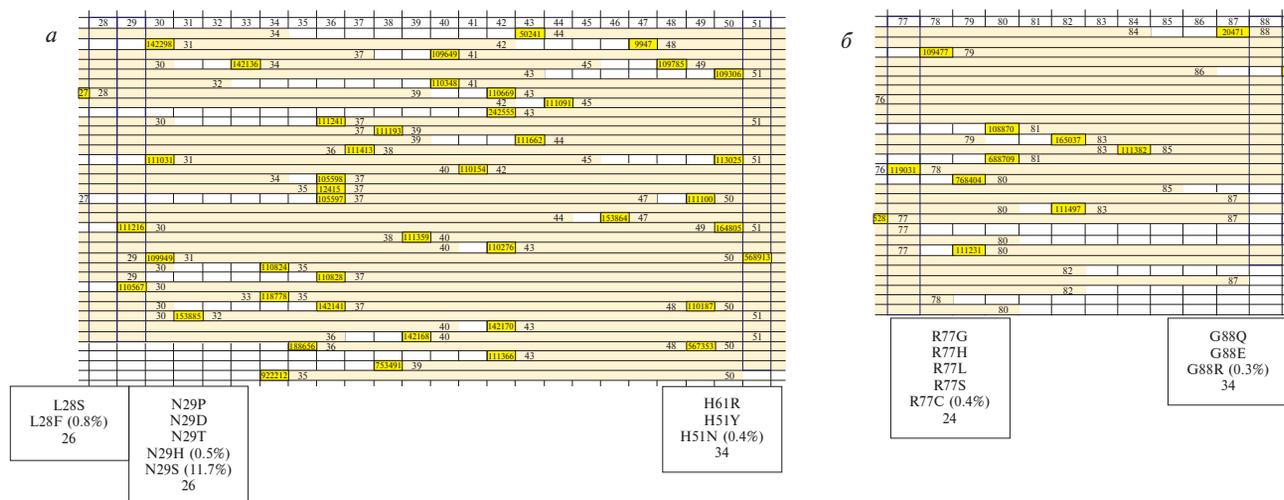
Нами проведен анализ результатов многочисленных исследований и идентифицированы АК-замены в белках ВПЧ16, характерные для различных регионов мира. Важно отметить, что ген E6 имеет два старт-кодона (AUG) и две открытые рамки считывания: 1) ORF-E6(1) между нуклеотидами 83 и 559, которая кодирует белок длиной 158 а.о.; и 2) ORF-E6(2) между нуклеотидами 104 и 559, которая кодирует белок, состоящий из 151 а.о. По сравнению с AUG в позиции 104–106 старт-кодон в позиции 83 слабее, так как в его последовательности Козак отсутствует пурин в положении –3 AUG и G в позиции +4 AUG, что ослабляет связывание мРНК E6 рибосомой. В силу этого старт-кодон AUG в позиции 83 реже используется для инициации трансляции, поэтому преимущественно синтезируется белок длиной 151 а.о. [15]. Тем не менее обе формы белка E6 активно изучают, в связи с чем нумерация остатков может различаться. Например, замена R17G может быть представлена как R20G, L90V – как L83V и т.д., что осложняет анализ последовательностей. В проведенном исследовании в качестве референсной мы использовали нуклеотидную последовательность гена E6 ВПЧ16 (NC\_001526, GeneID: 1489078) и белковый продукт, кодируемый ORF-E6(1) (83–559 н.), то есть длиной 158 а.о.

Таблица 8. Кластеризация изолятов Е6\*

Позиция	Кластер I N = 93		Кластер II N = 77		Кластер III N = 17		Кластер IV N = 17		Кластер V N = 53		Кластер VI N = 13		Кластер VII N = 28		Кластер VIII N = 47		Кластер IX N = 4	
	АК	n (%)	АК	n (%)	АК	n (%)	АК	n (%)	АК	n (%)	АК	n (%)	АК	n (%)	АК	n (%)	АК	n (%)
M1	-		-		-		-		-		-		L	7 (25%)	-		-	
R17	T G	1 (1.1%) 1 (1.1%)	G	2 (2.6%)	-		G	4 (23.5%)	-		I	1 (7.7%)	T I G	5 (17.9%) 3 (10.7%) 1 (3.6%)	-		-	
Q21	H	1 (1.1%)	E	1 (1.3%)	-		-		-		D H	1 (7.7%) 1 (7.7%)	D H E	24 (85.7%) 4 (14.3%) 1 (3.6%)	H D	40 (85.1%) 1 (2.1%)	H	1 (25%)
D32	E N H	42 (45.2%) 1 (1.1%) 2 (2.2%)	A	1 (1.3%)	-		-		N H	2 (3.8%) 1 (1.9%)	-		-		E N G	3 (6.4%) 1 (2.1%) 1 (2.1%)	-	
H85	-		-		-		Y	1 (5.9%)	-		Y	9 (69%)	Y	28 (100%)	Y	44 (93.6%)	Y	4 (100%)
L90	-		V	2 (2.6%)	V	15 (88.2%)	V	17 (100%)	V	52 (98%)	-	V	3 (10.7%)	V	43 (91.5%)	V	-	-
E120	-		-		-		-		-		D	1 (7.7%)	-	-	-	-	-	-

\*Указано число образцов (N) в каждом кластере, варианты замен (АК), их число (n) и частота встречаемости (%). АК-замены, встречающиеся в >10% последовательностей выделены жирным шрифтом.





**Рис. 8.** Эпитопная карта онкобелка E7. Представлены фрагменты 28–51 (а) и 52–99 (б) аминокислотной последовательности. Выборка включала последовательности, идентичность которых превышала 99.8%. Желтым цветом обозначены идентификационные номера (ID) эпитопов в базе данных IEDB, светло-желтым – последовательность картированного эпитопа, цифры внутри блока – позиции первых и последних аминокислотных остатков. В рамке внизу каждого блока указаны частоты встречаемости АК-замен в варибельной позиции белка E7 в составе эпитопного кластера. Цифры внизу – суммарное число эпитопов, затрагиваемых АК-заменами в указанных позициях.

образцах ВПЧ16, собранных в греческой популяции, выявили высокую частоту замены L90V в белке E6: 78.2% при высокодифференцированных опухолях шейки матки и несколько ниже, 64.2%, при низкодифференцированных. Корректных полногеномных данных секвенирования генов E6 и E7 из стран Западной Европы на сегодняшний день крайне мало, в связи с чем сравнить АК-замены в странах Восточной и Западной Европы на настоящий момент не представляется возможным.

В исследовании J. Zhao и др. [16], проведенном на образцах, собранных на территории Шанхая (Китай), описаны АК-замены H85Y и E120D в белке E6 и N29S в белке E7 и их связь с канцерогенезом (частота выявления данных замен была выше в группе пациенток с дисплазией шейки матки). Результаты проведенного нами анализа подтверждают эти данные: замена N29S в белке E7 широко распространена на территории Китая (57.2%, 210/367 образцов). Что касается замены E120D в белке E6, нами показано, что она встречается в Юго-Восточной Азии с частотой 13.5% (101/750 образцов), в то время как на территории Китая – всего в 7.9% образцов; при этом замена H85Y вообще не характерна для региона Юго-Восточной Азии – ее наибольшая частота приходится на вьетнамскую популяцию (10.9%) (рис. 4, табл. 7).

В то же время в Синьцзян-Уйгурском автономном районе Китая преобладает европейский вариант ВПЧ16 – с характерной заменой L90V в белке E6 и минимальной представленностью “азиатской” замены N29S в белке E7. Согласно нашим данным, частота “европейской” АК-замены L90V в этом

регионе составляет 78.2% (36/46 образцов), а “азиатская” (тип 2) D32E встречается редко (4.3%, 2/46 образцов) (рис. 4, табл. 7). Эти результаты согласуются с данными Н. Не и др. [21], полученными при анализе образцов ВПЧ16 из провинции Синьцзян. Преобладание определенных вариантов онкобелков E6 и E7 характерно не только для отдельных провинций Китая, но и для ряда других стран и/или их отдельных областей, в частности в Японии, в странах Африки (рис. 4, табл. 7). Интересно, что эти страны и их отдельные области некоторые исследователи считают моноэтническими [22, 23].

V. Pimenoff с соавт. [13] исследовали эволюции линий (субгенотипов) ВПЧ16, типированных на основании анализа полногеномных последовательностей вируса с последующим типированием на основе гена капсидного белка L1. Авторы составили карту географического распределения субгенотипов с выделением регионов с доминантным представителем субтипов А1–3 (Европа), А4 (Юго-Восточная Азия), В и С (Африка), D (Центральная и Южная Америка). Интересно, что эти регионы в целом совпадают с выявленной нами географической распространенностью вариантов белков E6 и E7. В дальнейшем мы планируем идентифицировать ВПЧ16-субгенотипспецифичные паттерны АК-замен в белках E6 и E7.

Анализ встречаемости АК-замен в онкобелках E6 и E7 представляется крайне важным ввиду предположения о связи определенных а.о. с канцерогенезом. Именно поэтому ряд исследований был нацелен на сравнение встречаемости АК-замен при РШМ и хронической персистенции ВПЧ

с частотой аналогичных замен в случаях спонтанной элиминации вируса. Таким образом, в настоящее время идет поиск корреляций между наборами АК-замен и течением, а также исходом хронической ВПЧ-инфекции. P. van der Weele с соавт. [24] не выявили в образцах ВПЧ16 из Нидерландов четкой корреляции между полиморфизмами генов *E6/E7* и персистенцией вируса и сделали вывод о том, что особенности генома ВПЧ не играют существенной роли в процессе элиминации/хронизации инфекции, а решающий фактор — особенности организма хозяина. Напротив, Noé Escobar-Escamilla и др. [9] продемонстрировали связь замены L90V в белке E6 ВПЧ16 с длительной персистенцией вируса. В Китае с высоким риском малигнизации и развитием РШМ ассоциированы полиморфизмы D32E и H85Y в белке E6 ВПЧ16 [8, 25].

Описанные АК-замены мало затрагивают ключевые структуры белков E6 и E7: два “цинковых пальца”, образованных мотивами CXXC (63–66, 106–109, 139–142 а.о.) и PBM (PDZ-binding motif; 148–151 а.о.), в белке E6 и ключевые участки в доменах CD2 (22–26, 31–32 а.о.) и CD3 (58–61, 91–94 а.о.) белка E7 [26]. Эти последовательности необходимы для корректного функционирования онкобелков и, как правило, высококонсервативны. Исключение составил участок 30–33 а.о. E6, несущий один из четырех мотивов CXXC, затронутый полиморфизмом D32E. Интересно, что в клетках, зараженных *in vitro* ВПЧ16 с заменой D32E в белке E6, экспрессия p53 была снижена по сравнению с клетками, инфицированными вирусом с E6 дикого типа [8]. T. Mesplède T и др. [10] показали, что замены R17G, L90V и их сочетание R17G/L90V в онкобелке E6 ВПЧ16 ассоциированы со значимым снижением уровня p53 по сравнению с его экспрессией в клетках, инфицированных ВПЧ16 с E6 дикого типа. Таким образом, хотя большинство полиморфизмов гена *E6* не приводит к изменениям в структурно-функциональных доменах кодируемого белка, некоторые несинонимичные замены все-таки влияют на функционирование E6, как правило, усиливая его онкогенный потенциал. Что касается онкобелка E7, то его неизменная первичная структура, по-видимому, важна для сохранения онкогенных свойств [27].

В этой связи имеет смысл отдельно рассмотреть коэволюционные изменения: когда возникновение АК-замены в определенной позиции стимулирует появление вторичных адаптивных замен в других положениях, что необходимо для сохранения функциональности белка. Коэволюцию выявляют методом ковариационного анализа. Такой анализ, проведенный нами для общей совокупности аминокислотных последовательностей, не позволил выявить наличие ковариационных связей ни в белке E6, ни в E7. Ковариационный паттерн,

охватывающий переменные позиции 17, 21, 85 и 90 для белка E6, обнаружен только для последовательностей американского происхождения. Возможно, наличие разных паттернов сцепления для генетически отдаленных наборов последовательностей маскирует существующие ковариационные связи. Учитывая это, в дальнейшем мы планируем исследовать наличие групп сцепления аминокислотных остатков на генетически однородных выборках образцов с использованием различных алгоритмов биоинформатической обработки. Успешность такого подхода была продемонстрирована нами ранее при проведении ковариационного анализа АК-замен в белке NS5A вируса гепатита С [28].

Влияние первичной структуры белка на его функции и их изменения в ходе эволюции вполне согласуются с появлением естественных мутаций в белках длительно циркулирующего в человеческой популяции ВПЧ, хотя трудно понять их филогеографию: географические различия распространности отдельных АК-замен и их сочетаний. Возможно, причина кроется в этнически обусловленных различиях хозяев, например при процессинге вирусных белков и ответе иммунной системы. Уже были сообщения о полиморфизмах системы процессинга антигенов и рецепторов иммунных клеток организма-хозяина (TLR, FAS, Tap1/2 и др.), об этнических различиях в распространности аллелей и гаплотипов HLA [29–31]. Совокупность этих факторов обуславливает этнические и, как следствие, географические различия в распознавании В- и Т-клеточных эпитопов. Однонуклеотидные замены в геноме ВПЧ, в том числе несинонимичные, происходят случайным образом. Однако, если АК-замена затрагивает структуру эпитопа, то это может нарушить распознавание соответствующего антигена иммунными клетками, что приводит к ускользанию вируса-мутанта от иммунного ответа хозяина. Конечно, такой вирус получает преимущества над родительским и продолжает циркуляцию в этой популяции, вытесняя другие штаммы. Так, по результатам биоинформатического анализа J. He и др. [25] выявили, что замены R17G, D32N, D32E, H85Y и L90V в белке E6 затрагивают Т-клеточные эпитопы, а это означает появление новых эпитопов или изменение аффинности связывания молекулами HLA [25]. К высокоиммуногенной относится область E14–I34 а.о. онкобелка E6 [32]. Экспериментально показано, что замены в позиции 17 белка E6 влияют на распознавание Т-клеточного эпитопа молекулой HLA-B7 [33]. Достоверно большее число АК-замен в иммуногенных областях E14–I34 и L45–A68 у пациентов с инвазивной цервикальной карциномой по сравнению с пациентами с раком *in situ* считают отражением факта ухода вируса от иммунного ответа/эпитопного дрейфа, вызывающего потерю организмом контроля за его репликацией [32]. Такого

рода замены могут носить сочетанный характер. Так, для гемагглютинина вируса гриппа А выявлено сочетанное изменение аминокислотных остатков на поверхности белка в В-клеточных эпитопах или вблизи, что определяет ускользание вируса от иммунного ответа хозяина [34].

Учитывая различную частоту встречаемости АК-замен в географических популяциях, выявленную филогенетическим анализом их реверсию, а также тот факт, что наиболее частые АК-замены Е6 ВПЧ затрагивали структуры эпитопов, мы пришли к выводу, что распознавание вирусного антигена клетками иммунной системы хозяина – один из ключевых факторов в закреплении новой мутации или, наоборот, элиминации мутировавшего варианта вируса из данного региона.

Изменчивость вирусов и ее направленность, особенно при хронических инфекциях, – результат действия многих факторов. Случайные нуклеотидные замены в геноме вируса, вызванные ошибками в работе вирусной полимеразы, приводят к мутациям в белках, влияя на их функционирование и, в итоге, на преобладание определенных вариантов вируса [35]. Однако не только репликативная активность и эффективность сборки вирионов определяют успешное распространение определенных вариантов вирусов. Давление со стороны иммунной системы хозяина относится к важнейшим факторам, определяющим эволюцию вирусов, так как АК-замены в В- и Т-клеточных эпитопах вирусных белков позволяют мутантному варианту ускользнуть от иммунного ответа. Так, А. Китаг и др. [36] продемонстрировали важность 16–60 и 76–94 а.о. белка Е6 ввиду их влияния на иммуногенность вируса, так как эти области содержат большинство известных эпитопов. Иммунное давление зависит от представленности в популяции организма-хозяина некоторых иммунных гаплотипов, которые и определяют различия в генетике человека, в том числе из разных этносов. Именно поэтому в результате коэволюции вирусного патогена и человека в определенных этнических группах преобладают разные варианты одного и того же вируса, несущие те или иные АК-замены в иммуногенных эпитопах. Ранее мы выдвигали гипотезу разнонаправленной эволюции вируса гепатита С в 7 этнически разных популяциях под действием различного иммунного давления [28, 37]. Адаптация вирусов к HLA-ассоциированному иммунному давлению описана для других вирусов, включая ВИЧ-1 [38] и вирус гепатита В [39].

В результате проведенного филогенетического анализа гена *Е6* продемонстрировано четкое распределение АК-замен в кодируемом белке по популяциям, что подтверждает кластеризацию нуклеотидных последовательностей по географическому признаку. Ускользание ВПЧ16 от иммунного давления, приводящее к разнонаправленной

эволюции вируса в разных этносах (то есть направление эволюции определяется преобладающими в популяции иммунными гаплотипами), до сих пор описано не было. Этноспецифические особенности иммунного ответа на ВПЧ и его адаптивную изменчивость – тема дальнейших исследований.

В результате проведенного анализа нами идентифицирован общий предок ВПЧ16 – вирус американского происхождения. Однако остался нерешенным ряд вопросов:

– образовавшиеся варианты ВПЧ16 перешли из Америки в другие регионы и сформировали там очаги или американские штаммы приобрели те же “региональные” нуклеотидные замены, потому что попали в те же этнические группы, что в Африке и в Азии;

– связано ли появление в Америке этноспецифических вариантов ВПЧ16 (по белкам Е6 и Е7) с массовой миграцией в конце XIX–начале XX вв.

Ограниченность выборки нуклеотидных последовательностей по времени (самые ранние образцы были собраны в конце 1980 гг.) не позволяет ответить на эти вопросы, что определяет необходимость дальнейшего изучения этой темы.

К существенным ограничениям для экстраполяции полученных данных на популяции всех описанных географических регионов относится недостаточный охват секвенированием генома ВПЧ16, поэтому спектр идентифицированных АК-замен и реконструкция их географического распространения носит мозаичный характер. В дальнейших исследованиях мы планируем охватить секвенированием гены *Е6* и *Е7* изолятов ВПЧ16, циркулирующих на европейской и азиатской территориях Российской Федерации, а также провести полногеномный анализ этих штаммов, что позволит подтвердить или опровергнуть результаты, полученные на фрагментах генома (гены *Е6* и *Е7*).

Нами продемонстрировано, что циркулирующие штаммы ВПЧ16 имеют общего предка, который, попав в разные географические регионы, эволюционировал по-разному, что подтверждено различным спектром АК-замен в белке Е6 в разных популяциях. Мы считаем, что различия в паттернах мутаций обусловлены этноспецифическими особенностями иммунного ответа. Полученные результаты могут быть использованы при надзоре за распространением ВПЧ-инфекции, разработке диагностических тест-систем, а также вакцин против ВПЧ-ассоциированных онкологических заболеваний.

Авторы декларируют отсутствие внешнего финансирования для проведения исследования и публикации статьи.

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

Авторы сообщают об отсутствии конфликта интересов.

### СПИСОК ЛИТЕРАТУРЫ

1. Тихомиров А.Л., Сарсания С.И., Филатова Г.А. (2018) Вирус папилломы человека: от понимания иммунопатогенеза к рациональной тактике ведения. *Гинекология*. **3**, 5–11. doi: 10.26442/2079-5696-2018.3.5-11
2. Волгарева Г.М. (2020) Папилломавирусный канцерогенез. Основные достижения и некоторые проблемы Часть I. Общие представления о папилломавирусах. Формы рака, ассоциированные с вирусами папилломы человека. *Российский биотерапевтический журнал*. **19**(1), 6–12. doi: 10.17650/1726-9784-2019-19-1-6-12
3. Huibregtse J.M., Scheffner M., Howley P.M. (1991) A cellular protein mediates association of p53 with the E6 oncoprotein of human papillomavirus types 16 or 18. *EMBO J.* **10**(13), 4129–4135. doi: 10.1002/j.1460-2075.1991.tb04990.x
4. Hebner C., Beglin M., Laimins L.A. (2007) Human papillomavirus E6 proteins mediate resistance to interferon-induced growth arrest through inhibition of p53 acetylation. *J. Virol.* **81**(23), 12740–12747. doi: 10.1128/JVI.00987-07
5. Um S.J., Rhyu J.W., Kim E.J., Jeon K.C., Hwang E.S., Park J.S. (2002) Abrogation of IRF-1 response by high-risk HPV E7 protein *in vivo*. *Cancer Lett.* **179**(2), 205–212. doi: 10.1016/s0304-3835(01)00871-0.
6. Курмышкина О.В., Волкова Т.О., Ковчур П.И., Бахлаев И.Е., Немова Н.Н. (2011) Гены раннего ответа в патогенезе рака шейки матки: обзор. *Опухоли женской репродуктивной системы*. **1**, 96–105. doi: 10.17650/1994-4098-2011-0-1-96-105
7. Duensing A., Spardy N., Chatterjee P., Zheng L., Parry J., Cuevas R., Korzeniewski N., Duensing S. (2009) Centrosome overduplication, chromosomal instability, and human papillomavirus oncoproteins. *Environ. Mol. Mutagen.* **50**(8), 741–747. doi: 10.1002/em.20478
8. Ai W., Wu C., Jia L., Xiao X., Xu X., Ren M., Xue T., Zhou X., Wang Y., Gao C. (2022) Deep sequencing of HPV16 E6 region reveals unique mutation pattern of HPV16 and predicts cervical cancer. *Microbiol. Spectr.* **10**(4), e0140122. doi: 10.1128/spectrum.01401-22
9. Escobar-Escamilla N., González-Martínez B.E., Araiza-Rodríguez A., Fragoso-Fonseca D.E., Pedroza-Torres A., Landa-Flores M.G., Garcés-Ayala F., Mendieta-Condado E., Díaz-Quiñonez J.A., Castro-Escarpulli G., Ramírez-González J.E. (2019) Mutational landscape and intra-host diversity of human papillomavirus type 16 long control region and E6 variants in cervical samples. *Arch. Virol.* **164**(12), 2953–2961. doi: 10.1007/s00705-019-04407-6
10. Mesplède T., Gagnon D., Bergeron-Labrecque F., Azar I., Sénéchal H., Coutlée F., Archambault J. (2012) p53 degradation activity, expression, and subcellular localization of E6 proteins from 29 human papillomavirus genotypes. *J. Virol.* **86**(1), 94–107. doi: 10.1128/JVI.00751-11
11. Бестаева Н.В., Назарова Н.М., Прилепская В.Н., Трофимов Д.Ю., Бурменская О.В., Суламанидзе Л.А. (2013) Папилломавирусная инфекция: новые взгляды на диагностику и лечение (обзор литературы). *Гинекология*. **3**, 4–7. <https://gynecology.orscience.ru/2079-5831/article/view/28190>
12. Burk R.D., Harari A., Chen Z. (2013) Human papillomavirus genome variants. *Virology*. **445**(1–2), 232–243. doi: 10.1016/j.virol.2013.07.018
13. Pimenoff V.N., de Oliveira C.M., Bravo I.G. (2017) Transmission between archaic and modern human ancestors during the evolution of the oncogenic human papillomavirus 16. *Mol. Biol. Evol.* **34**(1), 4–19. doi: 10.1093/molbev/msw214
14. Kirnbauer R., Hubbert N.L., Wheeler C.M., Becker T.M., Lowy D.R., Schiller J.T. (1994) A virus-like particle enzyme-linked immunosorbent assay detects serum antibodies in a majority of women infected with human papillomavirus type 16. *J. Natl. Cancer Inst.* **86**(7), 494–499. doi: 10.1093/jnci/86.7.494
15. Shally M., Alloul N., Jackman A., Muller M., Gissmann L., Sherman L. (1996) The E6 variant proteins E6I–E6IV of human papillomavirus 16: expression in cell free systems and bacteria and study of their interaction with p53. *Virus Res.* **42**(1–2), 81–96. doi: 10.1016/0168-1702(96)01301-9
16. Zhao J., Zhu J., Guo J., Zhu T., Zhong J., Liu M., Ruan Y., Liao S., Li F. (2020) Genetic variability and functional implication of HPV16 from cervical intraepithelial neoplasia in Shanghai women. *J. Med. Virol.* **92**(3), 372–381. doi: 10.1002/jmv.25618
17. Ortiz-Ortiz J., Alarcón-Romero L.D.C., Jiménez-López M.A., Garzón-Barrientos V.H., Calleja-Macias I., Barrera-Saldaña H.A., Leyva-Vázquez M.A., Illades-Aguilar B. (2015) Association of human papillomavirus 16 E6 variants with cervical carcinoma and precursor lesions in women from Southern Mexico. *Virol. J.* **12**, 29. doi: 10.1186/s12985-015-0242-3
18. Ramas V., Mirazo S., Bonilla S., Ruchansky D., Arbiza J. (2018) Analysis of human papillomavirus 16 E6, E7 genes and long control region in cervical samples from Uruguayan women. *Gene*. **654**, 103–109. doi: 10.1016/j.gene.2018.02.023
19. Szostek S., Zawilinska B., Klimek M., Kosz-Vnenchak M. (2017) HPV16 E6 polymorphism and physical state of viral genome in relation to the risk of cervical cancer in women from the south of Poland. *Acta Biochim. Pol.* **64**(1), 143–149. doi: 10.18388/abp.2016\_1364
20. Tsakogiannis D., Papadopoulou A., Kontostathi G., Rutherford I.G.A., Kyriakopoulou Z., Dimitriou T.G., Orfanoudakis G., Markoulatos P. (2013) Molecular and evolutionary analysis of HPV16 E6 and E7 genes in Greek women. *J. Med. Microbiol.* **62**, 1688–1696. doi: 10.1099/jmm.0.055491-0
21. He H., Li H., Fan P., Zhu J., Pan Z., Pan H., Wu D., Ren X., Guo X., Li D., Pan Z., Shao R. (2016) Variants of human papillomaviruses 16 (HPV16) in Uigur women in Xinjiang, China. *Infect. Agent. Cancer*. **11**, 44. doi: 10.1186/s13027-016-0089-2
22. Pan Z., Xu S. (2020) Population genomics of East Asian ethnic groups. *Hereditas*. **157**(1), 49. doi: 10.1186/s41065-020-00162-w

23. Hudson M.J., Nakagome S., Whitman J.B. (2020) The evolving Japanese: the dual structure hypothesis at 30. *Evol. Hum. Sci.* **2**, e6. doi: 10.1017/ehs.2020.6
24. van der Weele P., Meijer C.J.L.M., King A.J. (2017) Whole-genome sequencing and variant analysis of human papillomavirus 16 infections. *J. Virol.* **91**(19), e00844-17. doi: 10.1128/JVI.00844-17
25. He J., Li Q., Ma S., Li T., Chen Y., Liu Y., Cui Y., Peng J., Shi Y., Wei X., Ding X. (2022) The polymorphism analysis and epitope predicted of Alphapapillomavirus 9 E6 in Sichuan, China. *Viol. J.* **19**, 14. <https://doi.org/10.1186/s12985-021-01728-4>
26. Tomaić V. (2016) Functional roles of E6 and E7 oncoproteins in HPV-induced malignancies at diverse anatomical sites. *Cancers* (Basel). **8**, 95. doi.org/10.3390/cancers8100095
27. Brauburger K, Hume AJ, Mühlberger E, Olejnik J. (2012) Forty-five years of Marburg virus research. *Viruses.* **4**(10), 1878–927. doi: 10.3390/v4101878
28. Kyuregyan K.K., Kichatova V.S., Karlsen A.A., Isaeva O.V., Solonin S.A., Petkov S., Nielsen M., Isagulians M.G., Mikhailov M.I. (2020) Factors influencing the prevalence of resistance-associated substitutions in NS5A protein in treatment-naive patients with chronic hepatitis C. *Biomedicines.* **8**(4), 80. doi: 10.3390/biomedicines8040080
29. Mehta A.M., Mooij M., Branković I., Ouburg S., Morré S.A., Jordanova E.S. (2017) Cervical carcinogenesis and immune response gene polymorphisms: a review. *J. Immunol. Res.* **2017**, 8913860. doi: 10.1155/2017/8913860
30. Pandey N.O., Chauhan A.V., Raithatha N.S., Patel P.K., Khandelwal R., Desai A.N., Choksi Y., Kapadia R.S., Jain N.D. (2019) Association of *TLR4* and *TLR9* polymorphisms and haplotypes with cervical cancer susceptibility. *Sci. Rep.* **9**(1), 9729. doi: 10.1038/s41598-019-46077-z
31. Tan S.C., Ismail M.P., Duski D.R., Othman N.H., Ankathil R. (2017) FAS c.-671A>G polymorphism and cervical cancer risk: a case-control study and meta-analysis. *Cancer Genet.* **211**, 18–25. doi: 10.1016/j.cancergen.2017.01.004
32. Rader J.S., Tsaih S.W., Fullin D., Murray M.W., Iden M., Zimmermann M.T., Flister M.J. (2019) Genetic variations in human papillomavirus and cervical cancer outcomes. *Int. J. Cancer.* **144**(9), 2206–2214. doi: 10.1002/ijc.32038
33. Ellis J.R., Keating P.J., Baird J., Hounsell E.F., Renouf D.V., Rowe M., Hopkins D., Duggan-Keen M.F., Bartholomew J.S., Young L.S., Stern P.L. (1995) The association of an HPV16 oncogene variant with HLA-B\*7 has implications for vaccine design in cervical cancer. *Nat. Med.* **1**(5), 464–470. doi: 10.1038/nm0595-464
34. Duvvuri V.R., Duvvuri B., Cuff W.R., Wu G.E., Wu J. (2009) Role of positive selection pressure on the evolution of H5N1 hemagglutinin. *Genomics Proteomics Bioinformatics.* **7**(1–2), 47–56. doi: 10.1016/S1672-0229(08)60032-7
35. Stern A., Andino R. (2016) Viral evolution: it is all about mutations. In: *Viral Pathogenesis* (3<sup>rd</sup> edition). Eds Katz M.G., Lynn Law G., Korth M.J., Nathanson N. Elsevier, pp. 233–240. doi: 10.1016/B978-0-12-800964-2.00017-3
36. Kumar A., Hussain S., Yadav I.S., Gissmann L., Nataraajan K., Das B.C., Bharadwaj M. (2015) Identification of human papillomavirus-16 E6 variation in cervical cancer and their impact on T and B cell epitopes. *J. Virol. Methods.* **218**, 51–58. doi: 10.1016/j.jviromet.2015.03.008
37. Kichatova V.S., Kyuregyan K.K., Soboleva N.V., Karlsen A.A., Isaeva O.V., Isagulians M.G., Mikhailov M.I. (2018) Frequency of interferon-resistance conferring substitutions in amino acid positions 70 and 91 of core protein of the Russian HCV 1b isolates analyzed in the T-cell epitopic context. *J. Immunol. Res.* **7**, 7685371. doi: 10.1155/2018/7685371
38. Avila-Rios S., Carlson J.M., John M., Mallal S., Brumme Z.L. (2019) Clinical and evolutionary consequences of HIV adaptation to HLA: implications for vaccine and cure. *Curr. Opin. HIV AIDS.* **14**(3), 194–204. doi: 10.1097/COH.0000000000000541
39. Lumley S.F., McNaughton A.L., Klenerman P., Lythgoe K.A., Matthews P.C. (2018) Hepatitis B virus adaptation to the CD8<sup>+</sup> T cell response: consequences for host and pathogen. *Front. Immunol.* **9**, 1561. doi: 10.3389/fimmu.2018.01561

## Amino Acid Substitution Patterns in the E6 and E7 Proteins of HPV type 16: Phylogeography and Evolution

E. E. Zelenova<sup>1, 2, \*</sup>, A. A. Karlsen<sup>3, 4</sup>, D. V. Avdoshina<sup>5</sup>, K. K. Kyuregyan<sup>3, 4</sup>, M. G. Belikova<sup>3, 5, 6</sup>, I. D. Trotsenko<sup>3, 7</sup>

<sup>1</sup>Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Moscow, 119991 Russia

<sup>2</sup>Blokhin National Medical Research Center of Oncology, Moscow, 115478 Russia

<sup>3</sup>Peoples' Friendship University of Russia, Moscow, 117198 Russia

<sup>4</sup>Mechnikov Research Institute of Vaccines and Sera, Moscow, 105064 Russia

<sup>5</sup>Chumakov Federal Scientific Center for Research and Development of Immunobiological Drugs, Russian Academy of Sciences, Moscow, 117218 Russia

<sup>6</sup>National Research Center for Epidemiology and Microbiology named after Honorary Academician N. F. Gamaleya, Ministry of Health of the Russian Federation, Moscow, 123098 Russia

<sup>7</sup>Moscow Oncological Society, Moscow, 111123 Russia

\*e-mail: zelenovayeye@gmail.com

The E6 and E7 proteins of human papillomavirus (HPV) play a key role in the oncogenesis of papillomavirus infection. Data on the variability of these proteins are limited, and the factors affecting their variability are poorly understood. We analyzed the variability of the currently known sequences of HPV type 16 (HPV16) E6 and E7 proteins, taking into account their geographic origin and year of sample collection, as well as the direction of their evolution in major geographic regions of the world. All sequences belonging to HPV16 genome fragments encoding E6 and E7 oncoproteins were downloaded from the NCBI GenBank database on October 6, 2022. Samples were filtered according to the following parameters: the sequence includes at least one of the two whole open reading frames, the collection date and the country of origin are known. A total of 3,651 full-genome nucleotide sequences encoding the E6 protein and 4,578 full-genome nucleotide sequences encoding the E7 protein were sampled. The nucleotide sequences obtained after sampling and alignment were converted to amino acid sequences and analyzed using MEGA11, R, RStudio, Jmodeltest 2.1.20, BEAST v1.10.4, Fastcov, and Biostrings software. The highest variability in E6 protein structure was recorded at positions 17, 21, 32, 85, and 90, while in E7, positions 28, 29, 51, and 77 were the most variable. The samples were divided geographically into 5 heterogeneous groups: African, European, American, Southwest and South Asia and Southeast Asia. Unique amino acid substitutions (AA-substitutions) in the E6/E7 proteins of HPV16, presumably characteristic of certain ethnic groups, were identified for a number of countries. They are mainly localized in sites of known B- and T-cell epitopes and relatively rarely in structural and functional domains. The revealed differences in AA-substitutions in different ethnic groups and their colocalization with clusters of B- and T-cell epitopes suggest their possible relationship with the geographical distribution of alleles and haplotypes of the major histocompatibility complex (HLA). This may lead to the recognition of a different set of B- and T-cell epitopes of the virus, resulting in regional differences in the direction of epitope drift. Phylogenetic analysis of the nucleotide sequences encoding the E6 protein of HPV16 revealed a common ancestor, confirmed regional clustering of the E6 protein gene sequences by the set of the most common AA-substitutions, and identified cases of reversion of individual AA-substitutions when the virus distribution region changed. For the E7 protein, a similar analysis was not possible due to high sequence homology. Covariance analysis of the pooled sample revealed that there was no relationship between amino acid residues in the E6 protein, in the E7 protein, and between E6 and E7. Data obtained are important for the development of therapeutic vaccines against HPV of high carcinogenic risk.

**Keywords:** human papillomavirus type 16, E6 protein, E7 protein, primary structure, evolution, phylogenetic analysis, covariance analysis, B-cell epitopes, T-cell epitopes

УДК 577.29

## СПОСОБЫ ПОВЫШЕНИЯ УРОВНЯ НОКИНА КОНСТРУКЦИИ, КОДИРУЮЩЕЙ ПЕПТИДНЫЙ ИНГИБИТОР СЛИЯНИЯ ВИЧ-1 МТ-С34, В ЛОКУС *CXCR4* В Т-КЛЕТОЧНОЙ ЛИНИИ СЕМ/R5

© 2024 г. Д. С. Голубев<sup>а</sup>, Д. С. Комков<sup>а, б</sup>, М. В. Шепелев<sup>а</sup>, Д. В. Мазуров<sup>а, с</sup>, Н. А. Круглова<sup>а, \*</sup><sup>а</sup>Центр высокоточного редактирования и генетических технологий для биомедицины, Институт биологии гена Российской академии наук, Москва, 119334 Россия<sup>б</sup>Department of Physiology and Cell Biology, Faculty of Health Sciences, Ben-Gurion University of the Negev, Beer-Sheva, 8410501 Israel<sup>с</sup>Division of Infectious Diseases and International Medicine, Department of Medicine, University of Minnesota, Minneapolis, 55455 USA

\*e-mail: natalya.a.kruglova@yandex.ru

Поступила в редакцию 31.10.2023 г.

После доработки 25.11.2023 г.

Принята к публикации 04.12.2023 г.

Низкая эффективность нокина генетических конструкций, особенно в первичных клетках человека, ограничивает применение технологии редактирования генома для терапевтических целей. Таким образом, остается актуальным поиск способов повышения нокина. В представленной работе на модели нокина конструкции, кодирующей пептидный ингибитор слияния ВИЧ-1 МТ-С34, в локус *CXCR4* человека в Т-клеточной линии СЕМ/R5 изучены возможности нескольких подходов к повышению эффективности этой технологии. В первую очередь оценена модификация донорной ДНК как способа повышения эффективности транспорта плазмид в ядро, а именно: введение в донорную плазмиду последовательностей DTS (DNA transporting sequence) вируса-40 обезьян (SV40) или сайтов связывания транскрипционного фактора NF-κB, влияние которых на уровень нокина было неизвестно. На использованной нами модели нокина в локус *CXCR4* такая модификация оказалась неэффективной. Второй подход, заключающийся в модификации нуклеазы Cas9 путем введения двух дополнительных сигналов ядерной локализации (NLS), позволил повысить уровень нокина на 30%. Наконец, блокировка репарации ДНК по пути негомологичного соединения концов с помощью ингибиторов ДНК-зависимой протеинкиназы вызвала повышение нокина в 1.8 раз. Комбинация двух последних подходов дала аддитивный эффект. Таким образом, с помощью увеличения числа NLS в белке Cas9 и ингибирования репарации ДНК по пути негомологичного соединения концов нам удалось значительно повысить уровень нокина конструкции, кодирующей пептидный ингибитор слияния ВИЧ-1, в клинически релевантный локус *CXCR4*, что может быть использовано для разработки эффективных генотерапевтических подходов к лечению ВИЧ-инфекции.

**Ключевые слова:** CRISPR/Cas9, редактирование генома, гомологичная рекомбинация, Т-лимфоциты, нокин, ингибиторы NHEJ, DTS, сигнал ядерной локализации

DOI: 10.31857/S0026898424040044, EDN: INCWAY

Технологии геномного редактирования с помощью системы CRISPR/Cas9 позволяют вводить направленные изменения в геном клеток млекопитающих [1]. Наиболее часто используемая эндонуклеаза Cas9 из *Streptococcus pyogenes* с помощью

короткого РНК-гида (гРНК) распознает последовательность-мишень в геноме и вносит двухцепочечный разрыв (double-strand break, DSB) в строго определенном месте [2]. В клетках млекопитающих репарация DSB происходит наиболее эффективно по механизму негомологичного соединения концов (non-homologous end joining, NHEJ), что приводит к формированию инделов, вызывающих сдвиг рамки считывания и нокаут гена. NHEJ – самый быстрый способ репарации и функционирует в любой фазе клеточного цикла [3]. Другой путь репарации DSB – по пути гомологичной рекомбинации (homology-directed repair, HDR) – происходит в присутствии донорной ДНК и позволяет

Сокращения: гРНК – РНК-гид; DNA-PK (DNA-dependent protein kinase) – ДНК-зависимая протеинкиназа; DSB (double-strand break) – двухцепочечный разрыв; DTS (DNA transporting sequence) – сигнал транспорта ДНК; HDR (homology-directed repair) – репарация по пути гомологичной рекомбинации; NHEJ (non-homologous end joining) – негомологичное соединение концов; SV40 (Simian virus 40) – вирус-40 обезьян.

проводить нокин целевых последовательностей в заданный локус генома. Однако эффективность репарации DSB по механизму HDR существенно ниже по сравнению с NHEJ, а активность HDR ограничена только S- и G2-фазами клеточного цикла [3].

Система CRISPR/Cas9 широко применяется для разработки новых методов генной терапии и уже используется в клинических испытаниях для лечения ряда наследственных болезней, включая серповидноклеточную анемию и бета-талассемию, а также онкологические и инфекционные заболевания [4]. Большое внимание уделяется редактированию T-клеток *ex vivo* для придания им противоопухолевых свойств или устойчивости к заражению ВИЧ [5, 6]. В настоящее время гораздо большие успехи достигнуты в случае, когда требуется провести нокаут гена по сравнению с введением новых последовательностей с использованием технологии нокина. Низкий уровень нокина, в особенности в первичных клетках человека, по-прежнему остается одним из главных ограничений для применения CRISPR/Cas9 в терапии. Для решения этой проблемы предложены различные подходы к повышению эффективности HDR и, как следствие, нокина [7].

Ранее нами разработана CRISPR/Cas9-платформа для нокина коротких нуклеотидных последовательностей, кодирующих пептидные ингибиторы слияния ВИЧ-1 с клеткой [8]. Эта технология перспективна с точки зрения использования в клинике, однако, безусловно, требует дальнейшего усовершенствования. В частности, необходим эффективный нокин конструкций, кодирующих пептидные ингибиторы, в локусы генов-корцепторов ВИЧ. Нами был продолжен поиск способов повышения уровня нокина конструкции, кодирующей пептидный ингибитор слияния ВИЧ-1 MT-C34, в локус *CXCR4* (далее: нокин *MTC34*). Во-первых, предложено модифицировать донорную плазмиду для увеличения эффективности доставки целевой плазмиды в ядро. Для этого использовали сигналы DTS (DNA transporting sequence), которые при введении в плазмиду способствуют направленному транспорту плазмид в ядро [9]. Один из таких сигналов – энхансер SV40, с которым связываются клеточные транскрипционные факторы, включая AP1, AP2 и NF- $\kappa$ B, и благодаря своим NLS-последовательностям переносят плазмиду из цитоплазмы в ядро [9, 10]. Например, плазмиды с DTS-сигналами обеспечивали более высокий уровень экспрессии репортера при введении в различные типы клеток [11, 12]. Кроме того, показано, что сайты связывания транскрипционного фактора NF- $\kappa$ B также повышают уровень экспрессии репортерного гена люциферазы в клетках HeLa, HEK293, Hep G2 и U373 [13]; при этом максимальный эффект получали при стимуляции

клеток под действием TNF $\alpha$ , активирующего ядерную транслокацию NF- $\kappa$ B [13]. Мы решили проверить, может ли модификация донорной ДНК сигналами DTS или сайтами связывания NF- $\kappa$ B повысить уровень CRISPR/Cas9-опосредованного нокина. Второй подход был направлен на усиление активности Cas9 за счет увеличения числа NLS [14, 15]. Наконец, третий способ повышения эффективности нокина состоял в воздействии на клеточные пути репарации DSB с помощью низкомолекулярных соединений, ингибирующих NHEJ и стимулирующих HDR [16].

## ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

**Клеточные линии.** Клетки линии SEM/R5, полученные на основе T-клеточной линии CCRF-SEM (ATCC) [8], культивировали на среде DMEM/F12 (“ПанЭко”, Россия) с добавлением 10% фетальной сыворотки крупного рогатого скота (“HyClone”, США), 4 mM L-глутамин и 40 мг/л гентамицин (“ПанЭко”) при 37°C в увлажненной атмосфере 5% CO<sub>2</sub>.

**Конструирование плазмид.** Плазмиды для эукариотической экспрессии Cas9 с 0×, 1× и 3× NLS конструировали на основе плазмиды pcDNA3.3-hCas9 (#41815; “Addgene”, США). Для этого амплифицировали C-концевой фрагмент Cas9 на матрице плазмиды pcDNA3.3-hCas9 с помощью праймеров 5'-EcoRI-Cas9 и 3'-C\_end-Cas9 (последовательности праймеров, использованных в работе, приведены в табл. 1; плазмидные конструкции, использованные и полученные в данной работе, указаны в табл. 2) и фрагмент, кодирующий эпитоп 3×HA, на матрице плазмиды pCMV-Cas12a-3×HA с праймерами 5'-3×HA и 3'-3×HA. Эпитоп 3×HA представлял собой трижды повторенную последовательность YPYDVPDYA, соответствующую участку 98–106 аминокислотной последовательности гемагглютинаина (HA) вируса гриппа человека. Далее амплифицированные фрагменты объединяли с помощью ПЦР по методу overlap extension и продукт реакции клонировали в вектор pJET1.2 (“Thermo Fisher Scientific”, США). Фрагмент EcoRI-AgeI из pJET1.2 клонировали по соответствующим сайтам в плазмиду pcDNA3.3-hCas9, получая pcDNA3.3-hCas9-0×NLS-3×HA. Для создания конструкций с одним и тремя NLS фрагменты, кодирующие одну или три последовательности NLS, амплифицировали на матрице плазмиды SP-Cas9-3×NLS с праймерами 5'-linker-NLS\_SV40 и 3'-linker-NLS и полученные фрагменты клонировали в вектор pBluescript KS(+) (“Stratagene”, США), из которого затем клонировали фрагменты Acc65I-NheI по соответствующим сайтам в плазмиду pcDNA3.3-hCas9-0×NLS-3×HA, получая pcDNA3.3-hCas9-1×NLS-3×HA, кодирующую Cas9 с одним NLS из SV40, и pcDNA3.3-hCas9-3×NLS-3×HA, кодирующую Cas9 с двумя NLS из SV40 и одним

**Таблица 1.** Олигонуклеотиды, использованные в работе

Название	Нуклеотидная последовательность, 5'→3'
5'-EcoRI-Cas9	GAGCAAATAAGCGAATTCTCC
3'-C_end-Cas9	CTAGCACCAGCGGTACCGTCTCCACCGAGCTGAG
5'-3×HA	GGTACCGCTGGTGCTAGCTACCCATACGATGTTCCAG
3'-3×HA	ACTAACCGGTCAGGCATAGTCGGG
5'-linker-NLS_SV40	ATACGGTACCGGCTCCGGCACCCCGT
3'-linker-NLS_SV40	GGTAGCTAGCCGAGCCACCGCCCAC
5'-Xba-SV40dts-Nco	CTAGATGCTTTGCATACTTCTGCCTGCTGGGGAGCCTGGGGACTTTCCACACC CTAACTGACACACATTCCACAG
3'-Xba-SV40dts-Nco	CATGCTGTGGAATGTGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAG CAGGCAGAAGTATGCAAAGCAT
5'-Not-DTS-Psp	GGCCATGCTTTGCATACTTCTGCCTGCTGGGGAGCCTGGGGACTTTCCACACC CTAACTGACACACATTCCACAG
3'-Not-DTS-Psp	TGCGACTGTGGAATGTGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAG CAGGCAGAAGTATGCAAAGCAT
5'-PstI-DTS-DraIII	GATGCTTTGCATACTTCTGCCTGCTGGGGAGCCTGGGGACTTTCCACACCCTA ACTGACACACATTCCACAGCT
3'-PstI-DTS-DraIII	TGTGGAATGTGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGC AGAAGTATGCAAAGCATCTGCA
5'-NF-κB_1_BbsI	GGCCGCTCTAGAGGGACTTTCCAGCTGGGGACTTTCCAGCTGGTCTTCAC
3'-NF-κB_1_BbsI	TCGAGTGAAGACCAGCTGGAAAGTCCCCAGCTGGAAAGTCCCTCTAGAGC
5'-NF-κB_2	CAGCTGGGGACTTTCCAGCTGGGGACTTTCCCCATGGGC
3'-NF-κB_2	TCGAGCCCATGGGGAAAGTCCCCAGCTGGAAAGTCCCCA

NLS из нуклеоплазмина (второй основной мотив PAAKKKK) [17].

Донорные плазмиды получали на основе плазмиды pJet-donor\_MT-C34, описанной ранее [8]. Для создания плазмиды pJet-donor\_MT-C34-3'DTS с одним DTS-сигналом одноцепочечные олигонуклеотиды 5'-Xba-SV40dts-Nco и 3'-Xba-SV40dts-Nco гибридизовали и полученный двухцепочечный олигонуклеотид клонировали в плазмиду pJet-donor\_MT-C34 по сайтам XbaI-NcoI.

Для создания плазмиды pJet-donor\_MT-C34-5'-3'DTS с двумя DTS-сигналами одноцепочечные олигонуклеотиды 5'-Not-DTS-Psp и 3'-Not-DTS-Psp гибридизовали и полученный двухцепочечный олигонуклеотид клонировали в плазмиду pJet-donor\_MT-C34-3'DTS по сайтам NotI-PspXI.

Для создания плазмиды pJet-donor\_MT-C34-2×5'-3'DTS с тремя DTS-сигналами одноцепочечные олигонуклеотиды 5'-PstI-DTS-DraIII и 3'-PstI-DTS-DraIII гибридизовали и полученный двухцепочечный олигонуклеотид клонировали в плазмиду pJet-donor\_MT-C34-5'-3'DTS по сайтам PstI-DraIII.

Для получения донорных плазмид с сайтами связывания NF-κB сначала олигонуклеотиды 5'-NF-κB\_1\_BbsI, 3'-NF-κB\_1\_BbsI, 5'-NF-κB\_2 и 3'-NF-κB\_2 гибридизовали и клонировали в плазмиду pBluescript KS(+) по сайтам NotI-PspXI. Полученную плазмиду обрабатывали рестриктазами

NotI-PspXI или NcoI-XbaI для клонирования последовательности с 4 сайтами связывания NF-κB на 5'- или 3'-конец донорной ДНК в плазмиду pJet-donor\_MT-C34 по соответствующим сайтам рестрикции. Далее последовательности всех сконструированных на основе вектора pJET1.2 донорных молекул клонировали из вектора pJET в вектор pBluescript KS(+) по сайтам PstI-ClaI.

Нуклеотидные последовательности всех сконструированных плазмид верифицировали с помощью секвенирования по Сэнгеру в компании “Евроген” (Россия). Всю плазмидную ДНК, кроме донорной, выделяли с помощью набора Plasmid Midiprep 2.0 (#BC124; “Евроген”), а плазмидные доноры – с использованием набора Plasmid Midi Kit (100) (#12145; “Qiagen”, Германия).

**Электропорация.** Для нокина конструкции, кодирующей пептид MT-C34, в локус *CXCR4* 10<sup>6</sup> клеток СЕМ/R5 электропорировали, используя 3 мкг плазмиды pcDNA3.3-hCas9 (#41815; “Addgene”), 1 мкг описанной ранее [8] плазмиды pKS-gRNA-X4ex2, кодирующей гРНК с мишенью в экзоне-2 гена *CXCR4* (спейсер: 5'-CACTTCAGATAACTACACCG-3', РАМ: AGG), а также 1 пмоль (~2.9–3.0 мкг) одной из донорных плазмид (все плазмиды приведены в табл. 2). Для трансфекции использовали прибор Neon electroporation system со 100-микролитровыми наконечниками (“Invitrogen”, США) и следующими

**Таблица 2.** Плазмидные конструкции, использованные в работе

Название	Описание	Источник/ссылка
pKSgRNA-X4ex2	Вектор для экспрессии гРНК против экзона-2 локуса <i>CXCR4</i>	[8]
pBluescript KS(+)	–	“Stratagene”
pJET1.2	–	“Thermo Fisher Scientific”
pCR-Blunt	–	“Invitrogen”
SP-Cas9-3×NLS	Плаزمида для бактериальной экспрессии Cas9 с тремя NLS получена на основе плазмиды SP-Cas9 (#62731, “Addgene”)	Не опубликована
pJet-donor_MT-C34	Немодифицированный плазмидный донор для нокина конструкции, кодирующей пептид MT-C34, в экзон-2 локуса <i>CXCR4</i>	[8]
pKS-donor_MT-C34	Немодифицированный плазмидный донор для нокина конструкции, кодирующей пептид MT-C34, в экзон-2 локуса <i>CXCR4</i>	Эта работа
pKS-donor_MT-C34-3'DTS	Плазмидный донор для нокина конструкции, кодирующей пептид MT-C34, в экзон-2 локуса <i>CXCR4</i> , модифицированный сигналом DTS на 3'-конце	Эта работа
pKS-donor_MT-C34-5'-3'DTS	Плазмидный донор конструкции, кодирующей пептид MT-C34, в экзон-2 локуса <i>CXCR4</i> , модифицированный двумя сигналами DTS: на 5'- и на 3'-конце	Эта работа
pKS-donor_MT-C34-2×5'-3'DTS	Плазмидный донор для нокина конструкции, кодирующей пептид MT-C34, в экзон-2 локуса <i>CXCR4</i> , модифицированный тремя сигналами DTS: двумя на 5'- и одним на 3'-конце	Эта работа
pKS-donor_MT-C34-3'-NF-κB	Плазмидный донор для нокина конструкции, кодирующей пептид MT-C34, в экзон-2 локуса <i>CXCR4</i> , модифицированный 4 сайтами связывания NF-κB на 3'-конце	Эта работа
pKS-donor_MT-C34-5'-NF-κB	Плазмидный донор для нокина конструкции, кодирующей пептид MT-C34, в экзон-2 локуса <i>CXCR4</i> , модифицированный 4 сайтами связывания NF-κB на 5'-конце	Эта работа
pKS-donor_MT-C34-5'-3'-NF-κB	Плазмидный донор для нокина конструкции, кодирующей пептид MT-C34, в экзон-2 локуса <i>CXCR4</i> , модифицированный 4 сайтами связывания NF-κB на 5'- и на 3'-конце	Эта работа
pcDNA3.3-hCas9	Плазмида для эукариотической экспрессии Cas9 с одним NLS без эпитопных меток	“Addgene”
pcDNA3.3-hCas9-0×NLS-3×HA	Плазмида для эукариотической экспрессии Cas9 без NLS	Эта работа
pcDNA3.3-hCas9-1×NLS-3×HA	Плазмида для эукариотической экспрессии Cas9 с одним NLS	Эта работа
pcDNA3.3-hCas9-3×NLS-3×HA	Плазмида для эукариотической экспрессии Cas9 с тремя NLS	Эта работа
pCMV-Cas12a-3×HA	Плазмида для эукариотической экспрессии AsCas12a <sup>a</sup> с тремя HA-эпитопами	Не опубликована

<sup>a</sup>Нуклеаза Cas12a из *Acidaminococcus* sp.

настройками: 1230 В, 40 мс, один импульс. Уровень нокаута и нокина анализировали с помощью проточной цитофлуориметрии на 5 сутки после электропорации.

**Низкомолекулярные соединения.** Сразу после электропорации к клеткам добавляли низкомолекулярные соединения в следующих конечных концентрациях: 1 мкМ SCR7 (SML1546; “Sigma”, США), 1 мкМ RS1 (R9782; “Sigma”), 5 мкМ L755507 (SML1362; “Sigma”), 2 мкМ KU-0060648 (S8045; “Selleckchem”, США), 0.5 мкМ NU7441 (S2638; “Selleckchem”), 0.05 мкМ трихостатин А (“Cayman Chemical”, США) и 2 мкМ M3814 (S8586; “Selleckchem”). В качестве контроля использовали растворитель диметилсульфоксид (DMSO). Среду меняли через 24 ч после добавления этих соединений.

**Проточная цитофлуориметрия.** Для иммунофлуоресцентного окрашивания клетки СЕМ/R5 инкубировали в течение 30 мин при 4°C в фосфатно-солевом буфере (PBS) с описанными ранее [8] мышинными моноклональными антителами С24 против пептида МТ-С34 или с мышинными моноклональными антителами против CXCR4 (клон 12G5, “Santa Cruz Biotechnology”, США). Клетки дважды промывали в PBS и инкубировали с козыми антителами, конъюгированными с Alexa488 или Alexa546, против IgG мыши (“Thermo Fisher Scientific”) в течение 30 мин при 4°C. Далее клетки два раза промывали в PBS и анализировали на проточном цитофлуориметре CytoFLEX S (“BeckmanCoulter”, США). В качестве отрицательного контроля использовали клетки, обработанные только вторичными антителами.

**Иммуноблотинг.** Для анализа уровня Cas9 в электропорированных клетках через 1 сут после электропорации клетки СЕМ/R5 лизировали в буфере RIPA (20 мМ Трис-НСl, рН 8.0, 150 мМ NaCl, 5 мМ ЭДТА, 1% (w/v) Triton X-100, 0.5% (w/v) дезоксихолат натрия, 0.1% (w/v) додецилсульфат натрия, 1 мМ фенилметилсульфонилфторид) и инкубировали при 4°C 15 мин. Затем лизаты центрифугировали при 12000 g и 4°C в течение 10 мин, после чего образцы супернатантов смешивали с 4× буфером для SDS-PAGE: 250 мМ Трис-НСl, рН 6.8, 40% глицерин, 8% SDS, 4% 2-меркаптоэтанол и 0.2% бромфеноловый синий – и инкубировали в течение 5 мин при 80°C. Лизаты анализировали электрофорезом в 10%-ном полиакриламидном геле в присутствии SDS (SDS-PAGE) в буферной системе Лэммли. После SDS-PAGE белки переносили на PVDF-мембрану полусухим методом в приборе Trans-Blot Turbo (“Bio-Rad”, США). Мембрану блокировали 5%-ным сухим обезжиренным молоком в PBS с 0.1% Tween-20. Для детекции Cas9 с 3×НА-эпитопом использовали моноклональные кроличьи антитела против эпитопа НА (клон С29F4, #3724; “Cell Signaling”, США),

для детекции  $\alpha$ -тубулина использовали моноклональные мышинные антитела клона 12G10 [18]. В качестве вторичных антител использовали поликлональные козы антитела против IgG кролика (#7074; “Cell Signaling”) или мыши (#7076; “Cell Signaling”), конъюгированные с пероксидазой хрена. Хемилюминесцентный сигнал детектировали на приборе ChemiDoc MP (“Bio-Rad”) с использованием реагента Immobilon (“Millipore”, США).

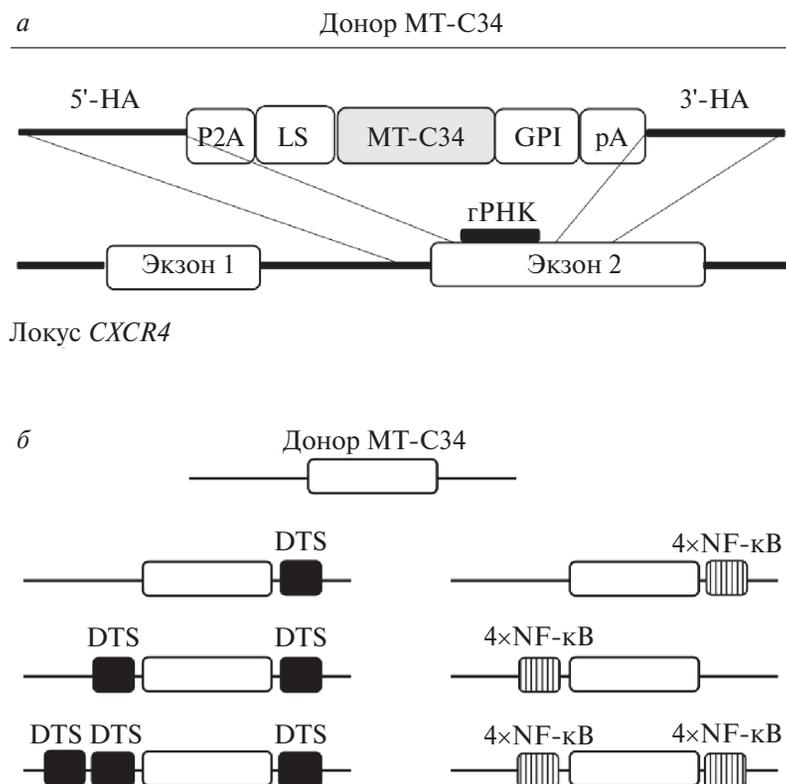
**Статистический анализ.** Обработку результатов проводили с использованием программы GraphPad Prism 8.0 (США). Для сравнения средних использовали однофакторный дисперсионный анализ (ANOVA) для связанных выборок с последующим применением теста Тьюки для множественных сравнений или с последующим применением теста Даннета для сравнения с контрольной группой (эксперимент с низкомолекулярными соединениями). Данные эксперимента с комбинацией подходов анализировали с использованием двухфакторного дисперсионного анализа с двумя факторами (число NLS, добавление соединения M3814) для связанных выборок с последующим применением теста Сидака для множественных сравнений.

## РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ

### *Модификация донорной плазмиды с помощью сигналов DTS и сайтов связывания NF- $\kappa$ B*

В качестве модели для оценки влияния различных модификаций донорной плазмиды на уровень нокина *MTC34* использовали полученную ранее Т-клеточную линию СЕМ/R5, стабильно экспрессирующую корецептор ВИЧ-1 CCR5 [8, 18]. Дизайн донорной конструкции для нокина, разработанный ранее [8], показан на рис. 1а. В результате корректного нокина в начало второго экзона гена *CXCR4* под контролем эндогенного промотора в клетках экспрессируется короткий фрагмент белка *CXCR4* и пептидный ингибитор слияния МТ-С34, разделенные Р2А-пептидом – сигналом пропуска рибосомы. В результате ингибиторный пептид МТ-С34 транслируется отдельно от фрагмента белка *CXCR4* и локализуется на плазматической мембране благодаря наличию лидерной последовательности и GPI-якоря (рис. 1а). Таким образом, все клетки с нокином *MTC34* также нокаутированы по *CXCR4* (нокаут вместе с нокином), в то время как клетки, потерявшие способность к экспрессии *CXCR4*, могут иметь как нокаут, так и нокаут вместе с нокином.

Мы предположили, что повысить уровень нокина можно за счет увеличения эффективности доставки донорной конструкции в ядро клетки. Для этого донорные плазмиды модифицировали путем введения DTS-сигналов из энхансера SV40 [9] или 4 сайтов связывания транскрипционного



**Рис. 1.** Дизайн конструкции для нокина *MTC34* в экзон-2 гена *CXCR4* (а) и схемы донорных конструкций с различными модификациями (б). 5'-НА – 5'-плечо гомологии; P2A – сигнал пропуска рибосомы; LS – лидерная последовательность; GPI – последовательность для модификации пептида GPI-якорем; pA – сигнал полиаденилирования; 3'-НА – 3'-плечо гомологии; гРНК – РНК-гид; DTS – ДНК-транспортирующая последовательность; 4×NF-κB – 4 сайта связывания NF-κB, разделенные короткими линкерами.

фактора NF-κB [19]. На основе донорной плазмиды pBluescript\_KS-donor\_MTC34 (далее: pKS-don\_MTC34) были получены плазмиды с одним, двумя или тремя сигналами DTS, а также плазмиды с четырьмя сайтами связывания NF-κB на 5'-, 3'-конце донорной последовательности или с обеих сторон от нее (см. рис. 1б).

Для оценки эффективности нокина клетки СЕМ/R5 электропорировали донорными плазмидами вместе с плазмидами, экспрессирующими Cas9 и гРНК, и на 5 сутки оценивали уровни нокина и нокаута с помощью проточной цитофлуориметрии с антителами соответственно против пептида МТ-С34 и корецептора CXCR4. Сначала определяли уровень нокина при различных дозах донорной ДНК и неизменной концентрации плазмид для экспрессии Cas9 и гРНК. Обнаружено, что при увеличении количества донорной ДНК от 0.5 до 4 пмоль возрастал и уровень нокина (рис. 2а), а уровень нокаута значимо не изменялся (рис. 2б). В результате соотношение нокина к нокауту увеличилось в 2 раза и достигало 60% (рис. 2в). Важно отметить, что при дозе донора в 4 пмоль наблюдалось снижение доли живых клеток (рис. S1, дополнительные материалы размещены в электронном виде по DOI: 10.31857/S0026898424040044 статьи,

а также размещены на сайте [http://www.molecbio.ru/downloads/2024/4/supp\\_Golubev\\_rus.pdf](http://www.molecbio.ru/downloads/2024/4/supp_Golubev_rus.pdf)). Кроме того, в работах, описывающих модификации донорной ДНК для повышения уровня нокина в Т-клетках, было показано, что как при низких, так и при высоких дозах донорной ДНК эффект от модификаций выражен слабо [20, 21], поэтому для дальнейших экспериментов мы выбрали промежуточную дозу: 1 пмоль донорной плазмиды.

Далее оценили влияние модификаций донорной плазмиды на уровни нокина и нокаута при выбранных условиях. Оказалось, что добавление сигналов DTS или сайтов связывания NF-κB не влияет на уровень как нокина *MTC34* (рис. 3а, б), так и нокаута *CXCR4* (рис. 3в, г).

#### Повышение нокина *MTC34* за счет увеличения числа NLS в Cas9

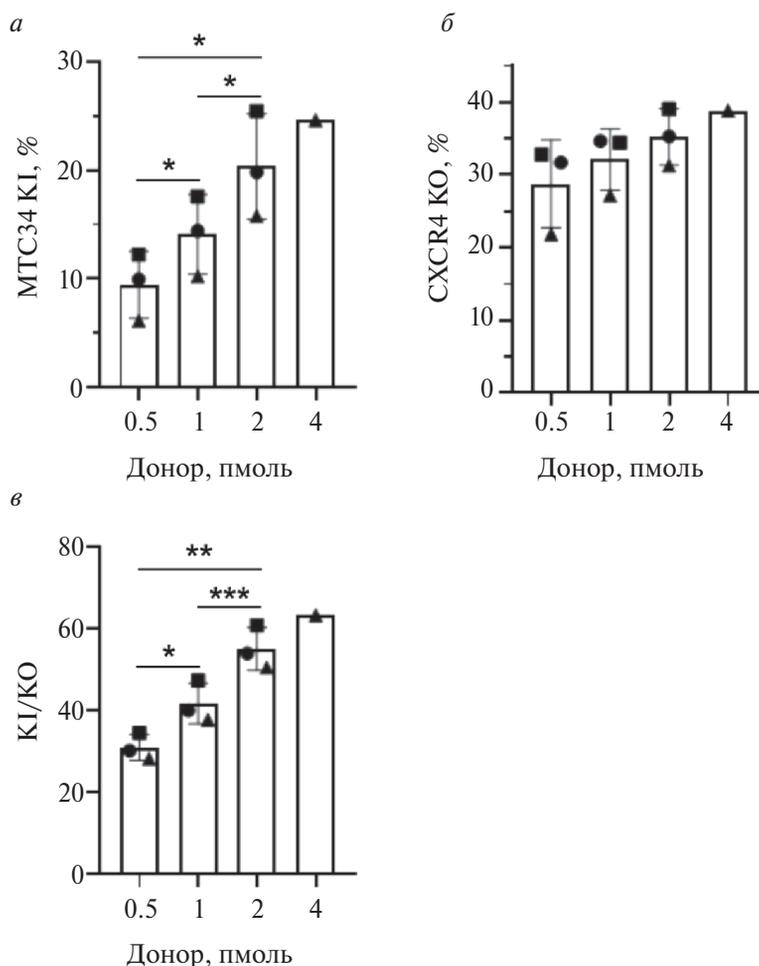
Известно, что количество и тип NLS в белке Cas9 влияют на эффективность геномного редактирования [14, 15]. Мы решили увеличить число NLS в белке Cas9 и сравнить активность нуклеаз с одним и тремя NLS, в дополнение использовали конструкцию без NLS (рис. 4а). Для облегчения детекции белков после иммуноблотинга на С-конец Cas9 после NLS добавили 3×НА-эпитоп

(рис. 4а), введение которого не влияло на уровень редактирования локуса *CXCR4* (рис. S2а, б, см. Дополнительные материалы). Плазмиды, кодирующие Cas9 и гРНК, вместе с донорной плазмидой без модификаций электропорировали в клетки СЕМ/Р5 и на 5 сутки оценивали уровень пептида МТ-С34 и белка *CXCR4* на поверхности клеток с помощью проточной цитофлуориметрии. Максимальный уровень нокина и нокаута достигался при экспрессии Cas9 с 3×NLS (рис. 4б, в). Обнаружено, что даже белок без сигналов NLS вызывал заметный уровень нокаута, однако соотношение нокина к нокауту при этом было снижено (рис. 4в, г). Интересно отметить, что введение последовательности NLS приводило к повышению уровня белка в электропорированных клетках СЕМ/Р5 (рис. 4г, д), а также в трансфицированных клетках НЕК293Т (рис. S2в, см. Дополнительные материалы). Это значит, что последовательности NLS,

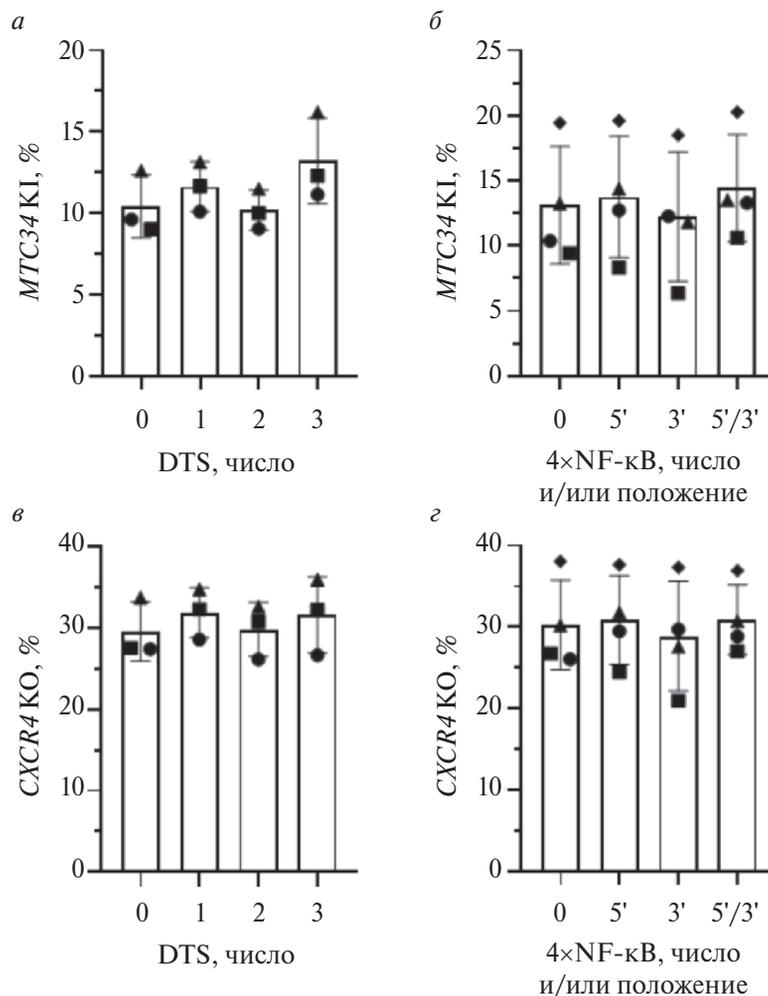
кроме обеспечения транслокации в ядро, могут стабилизировать белок Cas9.

#### Повышение уровня нокина за счет ингибирования NHEJ и стимуляции HDR

Для повышения уровня нокина часто используют низкомолекулярные соединения, влияющие на пути репарации двухцепочечных разрывов ДНК [16]. Учитывая это, для повышения уровня нокина *MTC34* в локус *CXCR4* мы проанализировали ряд соединений, которые, согласно данным литературы, повышают уровень репарации по механизму HDR или блокируют NHEJ. В качестве активаторов HDR использовали соединение RS-1, стабилизирующее Rad51-нуклеофиламенты [22]; агонист β-3 адренергического рецептора L755507 (механизм его влияния на HDR неизвестен) [16] и ингибитор гистондеацетилаз трихостатин А (Trichostatin A,



**Рис. 2.** Уровень нокина *MTC34* (а), нокаута *CXCR4* (б) и соотношение нокин/нокаут (в) в клетках СЕМ/Р5 при различном количестве донорной ДНК. Клетки электропорировали плазмидами рсDNA3.3-hCas9, рKS-gRNA-X4ex2 и донорной плазмидой рKS-don\_MТ-С34 и на 5 сутки после электропорации оценивали уровень нокаута *CXCR4* и нокина *MTC34* на поверхности клеток с помощью проточной цитофлуориметрии. KI/KO – соотношение нокина к нокауту. Результаты 3 независимых экспериментов представлены как среднее ± стандартное отклонение (SD) и индивидуальные значения; значками разной формы обозначены независимые эксперименты. \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ .



**Рис. 3.** Влияние DTS-сигналов в плазмидном доноре на уровень нокина. Уровень нокина *MTC34* (а, б) и нокаута *CXCR4* (в, г) в клетках SEM/R5. Клетки электропорировали плазмидами pсDNA3.3-hCas9, pKS-gRNA-X4ex2 и 1 пмоль донорной плазмиды pKS-don\_MT-C34 без модификаций, с различным числом сигналов DTS (а, в) или различным числом и/или положением сайтов связывания NF-κB (б, г). На 5 сутки после электропорации оценивали уровень нокаута *CXCR4* и нокина *MTC34* на поверхности клеток с помощью проточной цитофлуориметрии. Результаты 3–4 независимых экспериментов представлены как среднее ± SD и индивидуальные значения; значками разной формы обозначены независимые эксперименты. \* $p < 0.05$ , \*\* $p < 0.01$ .

TSA) [23]. Для блокировки NHEJ использовали ингибиторы ДНК-зависимой протеинкиназы (DNA-PK) Ku0060648, Nu7441 и M3814 и ингибитор ДНК-лигазы IV SCR7 [23] (рис. 5а).

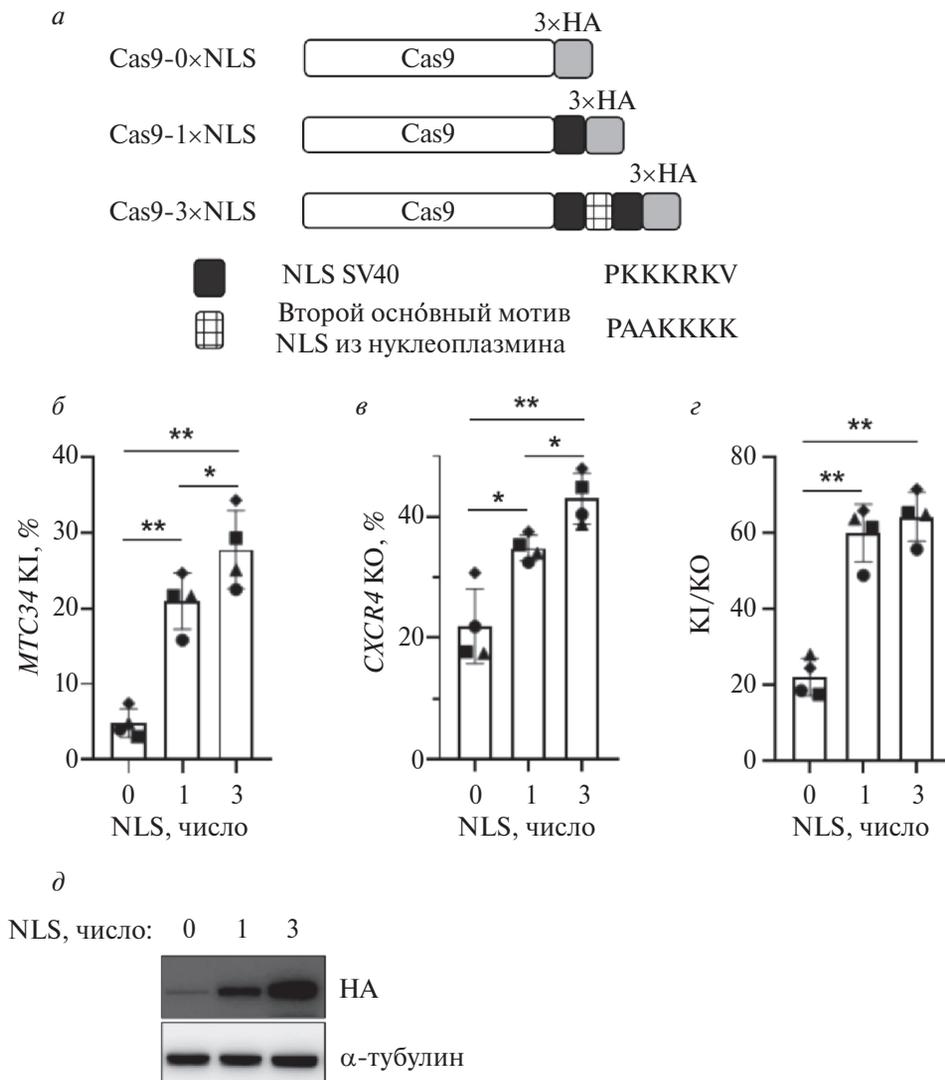
Для оценки эффективности низкомолекулярных соединений клетки SEM/R5 электропорировали плазмидами, кодирующими Cas9 и гРНК, вместе с 1 пмоль донорной плазмиды и сразу после электропорации добавляли в культуральную среду низкомолекулярные соединения, через 24 ч среду меняли на свежую и еще через 72 ч измеряли уровень нокаута *CXCR4* и нокина *MTC34* по экспрессии кодируемых ими белка *CXCR4* и пептида MT-C34 на поверхности клеток с помощью проточной цитофлуориметрии. Обнаружено, что уровень нокина достоверно повышался лишь в присутствии ингибиторов DNA-PK, среди которых соединения

Ku0060648 и M3814 обладали максимальным эффектом и увеличивали нокин в 1.8 раза (рис. 5б). Эффективность нокаута в присутствии низкомолекулярных соединений оставалась неизменной (рис. 5в). Это объясняется тем, что в используемой нами модельной системе нокаут *CXCR4* может быть вызван репарацией локуса-мишени как по пути NHEJ, так и по пути HDR. Во втором случае происходит нокин *MTC34* с одновременным нокаутом *CXCR4*. Более длительная инкубация клеток в присутствии низкомолекулярных соединений не приводила к повышению уровня нокина (рис. S3а–в, см. Дополнительные материалы). Стоит отметить, что введение в среду соединения Ku0060648 существенно замедляло рост клеток, в особенности при длительной инкубации (рис. S3г, см. Дополнительные материалы).

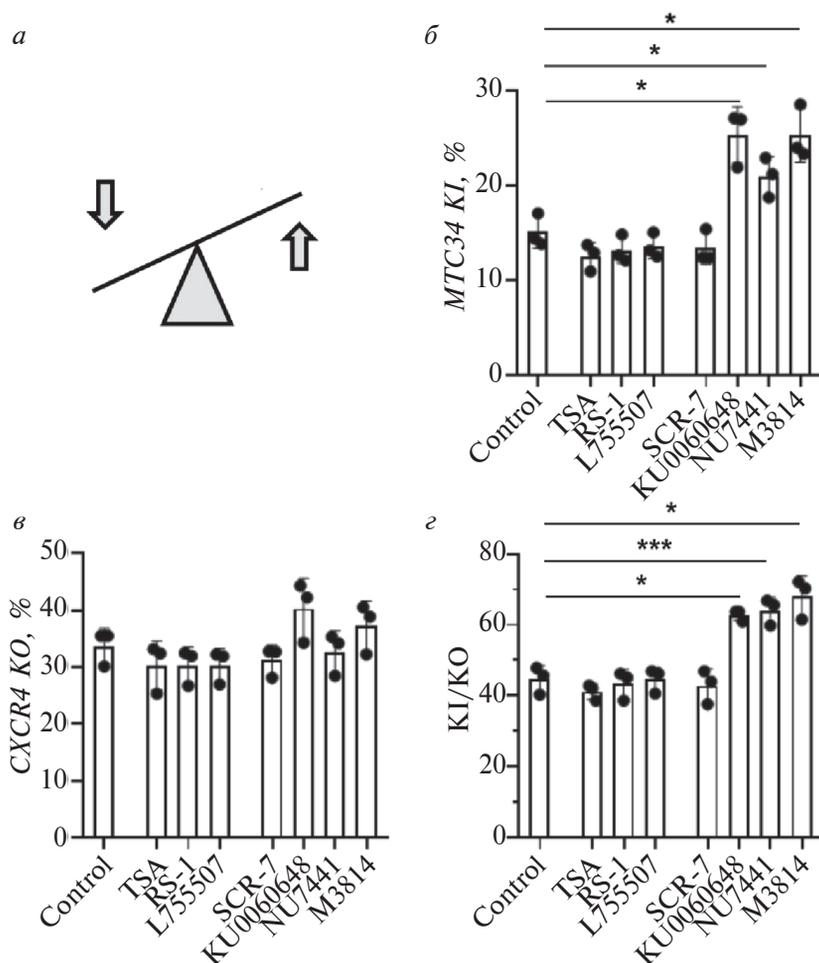
Комбинация ингибиторов DNA-PK  
с Cas9-3×NLS максимально повышала  
эффективность нокина MTC34

В результате проведенных экспериментов установлено, что использование Cas9-3×NLS и обработки клеток ингибиторами DNA-PK: Ku0060648 и M3814 — приводило к значимому повышению уровня нокина MTC34 в локус CXCR4. Далее оценили эффективность комбинирования этих подходов. Для этого клетки электропорировали смесью плазмид, кодирующих гРНК, донорную

последовательность без модификаций и Cas9 с 1×NLS или 3×NLS. После электропорации клетки в течение 24 ч культивировали в среде с DMSO или M3814, переводили в свежую среду и на 5 сутки после электропорации анализировали уровень CXCR4 и МТ-С34 на поверхности клеток с помощью проточной цитофлуориметрии. В образцах с Cas9-3×NLS по сравнению с образцами Cas9-1×NLS уровень нокаута был выше в среднем на 20%, а уровень нокина — на 30% (рис. 6а, б). При введении в среду соединения M3814 уровень нокина повышался на 30%, а уровень нокаута при



**Рис. 4.** Влияние числа NLS в белке Cas9 на эффективность редактирования локуса CXCR4. *а* — Схемы конструкций на основе плазмиды pcDNA3.3-hCas9. Приведены аминокислотные последовательности NLS SV40 и последовательность второго основного мотива из NLS нуклеоплазмина. Уровень нокина MTC34 (*б*), нокаута CXCR4 (*в*) и соотношение нокин/нокаут (*г*) в клетках СЕМ/Р5, электропорированных одной из конструкций 1–3 (*а*) вместе с плазмидами pKS-gRNA-X4ex2 и pKS-don\_MТ-С34. Уровень нокаута CXCR4 и нокина MTC34 оценивали методом проточной цитофлуориметрии на 5 сутки после электропорации. KI/KO — соотношение нокина к нокауту. Результаты 4 независимых экспериментов представлены как среднее ± SD и индивидуальные значения; значками разной формы обозначены независимые эксперименты. \* $p < 0.05$ , \*\* $p < 0.01$ . *д* — Анализ методом иммуноблоттинга экспрессии белка Cas9 с разным числом NLS (0, 1 и 3) в клетках СЕМ/Р5. Лизаты окрашивали антителами к HA-эпитопу (для детекции Cas9) и к α-тубулину (для контроля уровня общего белка в лизатах).



**Рис. 5.** Повышение уровня нокина за счет ингибирования механизма NHEJ и стимуляции HDR. *a* – Низкомолекулярные соединения, использованные в работе. Уровень нокина *MTC34* (*б*), нокаута *CXCR4* (*в*) и соотношения нокин/нокаут (*г*) в клетках СЕМ/Р5. Клетки электропорировали плазмидами рсDNA3.3-hCas9, рKS-gRNA-X4ex2 и рKS-don\_MТ-С34, через 24 ч среду меняли на свежую, еще через 72 ч оценивали уровень нокаута *CXCR4* и нокина *MTC34* методом проточной цитофлуориметрии. KI/KO – соотношение нокина к нокауту. Результаты 3 независимых экспериментов приведены как среднее  $\pm$  SD и индивидуальные значения; значками разной формы обозначены независимые эксперименты. \* $p < 0.05$ , \*\*\* $p < 0.001$ .

этом не изменялся. При совместном использовании обоих подходов наблюдался аддитивный эффект, при котором уровень нокина возрастал на 70% по сравнению с контрольным образцом.

## ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Изменение генома клеток с помощью CRISPR/Cas9-опосредованного нокина открывает новые возможности для моделирования и лечения заболеваний человека. Однако эффективность репарации DSB по механизму HDR зачастую невелика, поэтому усилия многих исследователей направлены на поиск условий для повышения уровня нокина. В частности, для этой цели применяли разные способы модификации нуклеазы Cas9, гРНК и донорной ДНК, а также воздействовали на пути репарации DSB [16]. Ранее нам удалось существенно повысить уровень нокина *MTC34* в локус *CXCR4*

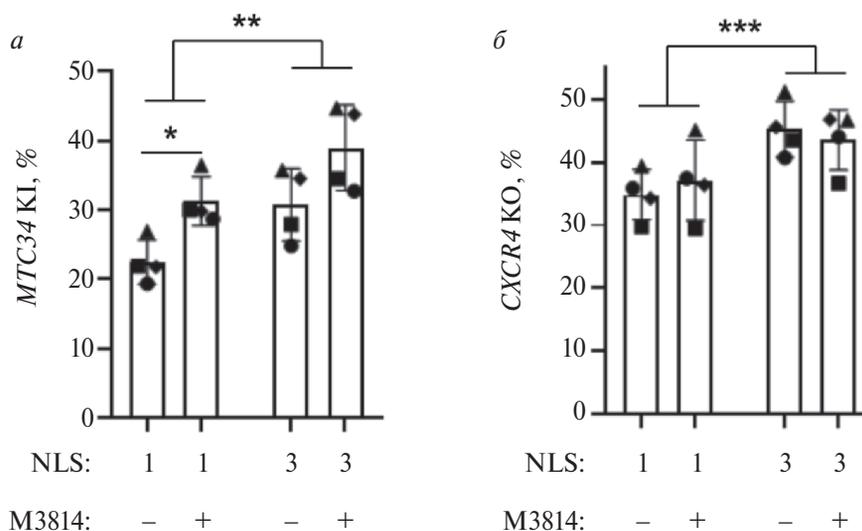
на Т-клеточной линии и первичных CD4<sup>+</sup> Т-лимфоцитах человека путем увеличения плеч гомологии донора со 100 до 500 п.н. и замены ПЦР-донора плазмидой [8]. В данной работе для повышения уровня нокина конструкции, кодирующей пептидный ингибитор слияния ВИЧ-1 МТ-С34, в локус гена *CXCR4* человека в Т-клеточной линии СЕМ/Р5 мы модифицировали донорную ДНК, нуклеазу Cas9 и воздействовали на активность путей репарации DSB с помощью низкомолекулярных соединений. Для повышения уровня нокина за счет более эффективной доставки донорной плазмидной ДНК в ядро клеток плазмиды модифицировали путем введения сигналов DTS, представляющих собой участки энхансера SV40, или сайтов связывания транскрипционного фактора NF- $\kappa$ B. Подобная модификация донорной ДНК легко выполняема и масштабируема, для ее применения не требуется вводить в клетки дополнительные белки

и, наконец, в отличие от воздействия на пути репарации ДНК с помощью низкомолекулярных соединений, такая модификация донорной ДНК потенциально менее токсична. Однако выбранные модификации донорной ДНК DTS-сигналами или сайтами связывания NF- $\kappa$ B не влияли на уровень нокина в используемой нами модельной системе.

Сигнал DTS из SV40 содержит несколько сайтов связывания различных транскрипционных факторов, экспрессируемых практически всеми типами клеток (AP1, AP2, NF- $\kappa$ B, Oct1, TEF-1) [10]. Как показано еще в 1999 году D. Dean и др. [9], введение в плазмиду последовательности, кодирующей DTS, повышает экспрессию репортера в трансфицированных клетках, что обусловлено направленным переносом плазмидной ДНК в ядро благодаря связыванию DTS-сигналов с транскрипционными факторами, имеющими NLS [9]. Впоследствии обнаружили, что индивидуальные сайты связывания некоторых транскрипционных факторов, в частности NF- $\kappa$ B, могут выступать в роли DTS [13, 24, 25]. Учитывая, что NF- $\kappa$ B – один из ключевых транскрипционных факторов, участвующих в процессе активации Т-клеток [26], мы оценили его вклад в повышение уровня нокина на Т-клеточной линии СЕМ/R5, вводя в донорную ДНК последовательности с сайтами связывания NF- $\kappa$ B, как описано S. Shin и др. [19]. Однако эта модификация не повлияла на уровень нокина *MTC34*, несмотря на то что белки семейства NF- $\kappa$ B содержатся в клетках линии СЕМ [27, 28]. Возможно, требуется дополнительная стимуляция клеток, приводящая

к усиленной транслокации NF- $\kappa$ B в ядро [13]. Можно предположить, что эффект от введения сигналов DTS и сайтов связывания NF- $\kappa$ B в донорную ДНК будет более выражен на первичных активированных Т-клетках, что предстоит проверить экспериментально.

Насколько нам известно, в литературе отсутствуют данные о том, что модификация донорной ДНК DTS-сигналами может повышать уровень CRISPR/Cas9-опосредованного нокина. Ранее лишь в одной работе, выполненной на зиготах крыс, применяли DTS-сигналы из SV40 для повышения CRISPR/Cas9-опосредованного нокина [29]. Однако авторам не удалось добиться детектируемого уровня сигнала, поэтому и вывода о роли DTS в этом процессе сделано не было. При выборе последовательности 4 сайтов связывания NF- $\kappa$ B, разделенных небольшими линкерами, мы ориентировались на работу S. Shin и др. [19]. В этой работе DTS-сигналы из SV40 и сайты связывания NF- $\kappa$ B использовали для повышения эффективности встраивания длинных конструкций по методу RMCE (Recombinase-Mediated Cassette Exchange) с участием Cre-рекомбиназы в локусы *AAVS1* или *ROSA26* [19]. Авторы наблюдали усиление сигнала репортера в 1.5–2 раза в присутствии вставки DTS SV40 или сайтов связывания NF- $\kappa$ B. В обоих случаях модификация требовалась по обоим концам встраиваемой конструкции, в то время как одного сигнала было недостаточно. В представленной нами работе даже введение трех DTS-сигналов SV40 в донорную ДНК не вызывало



**Рис. 6.** Повышение уровня нокина *MTC34* за счет экспрессии Cas9-3×NLS и обработки клеток ингибитором DNA-ПК. Уровень нокина *MTC34* (а) и нокаута *CXCR4* (б) в клетках СЕМ/R5, электропорированных плазмидами pKS-gRNA-X4ex2 и pKS-don\_MT-C34, а также одной из плазмид pсDNA3.3-hCas9-1×NLS (1) или pсDNA3.3-hCas9-3×NLS (3). В течение 24 ч после электропорации клетки культивировали в присутствии M3814 (+) или в отсутствие (-). На 5 сутки оценивали уровень нокаута *CXCR4* и нокина *MTC34* по экспрессии соответствующих белка *CXCR4* и пептида МТ-С34 на поверхности клеток методом проточной цитофлуориметрии. Результаты 4 независимых экспериментов представлены как среднее  $\pm$  SD и индивидуальные значения; значками разной формы обозначены независимые эксперименты. \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ .

заметного повышения уровня нокина *MTC34*. Различия могут быть связаны, во-первых, с экспрессией и активацией транскрипционных факторов, ответственных за связывание с сигналами DTS в различных клеточных линиях. Во-вторых, мы использовали Т-клеточную линию СЕМ/R5, а S. Shin с соавт. [19] работали с клетками НЕК293 и СНО. Клетки НЕК293 использовали и в работе, где впервые было показано положительное влияние сайтов связывания NF-κB в плазмиде на экспрессию люциферазы [13].

Второй подход был направлен на модификацию белка Cas9. Ранее показано, что увеличение числа NLS в молекуле Cas9 повышает эффективность доставки фермента в ядро и позволяет повысить уровень редактирования [14, 15]. Нами получена конструкция, кодирующая белок Cas9 с 3×NLS, которая позволила увеличить уровень нокаута *CXCR4* на 20%, а уровень нокина *MTC34* на 30% по сравнению с конструкцией, содержащей один сигнал NLS. Стоит отметить, что даже белок без сигналов NLS вызывал заметный уровень нокаута (20% по сравнению с 40% для белка Cas9-3×NLS); а уровень нокина был непропорционально снижен. Так, соотношение нокина к нокауту для белков Cas9-3×NLS и Cas9-1×NLS составляло около 60%, а для белка Cas9-0×NLS только 20%. В литературе описания похожего эффекта мы не нашли. Возможно, такая разница обусловлена различной кинетикой накопления Cas9-3×NLS и Cas9-0×NLS в ядре, что в последнем случае накладывает на снижение количества донорной ДНК, приводя в результате к падению уровня нокина. Кроме того, замечено, что уровень Cas9 в клетке повышался при увеличении числа NLS в молекуле (рис. 4д). В литературе мы не нашли такого типа данных ни для Cas9, ни для других белков. Недавно S. Shui и др. [30] сообщили о повышении эффективности редактирования локуса *CCR5* в Т-клеточной линии Jurkat при добавлении второго NLS SV40 на С-конец Cas9 [30], однако общий уровень белка Cas9 в электропорированных клетках авторы не оценивали. Стоит отметить, что S. Shui с соавт. [30] использовали рибонуклеопротеиновые (РНП) комплексы и добились повышения уровня редактирования в 1.5–2 раза, тогда как мы работали с плазмидными конструкциями и при введении двух сигналов NLS достигли повышения эффективности нокаута только на 20%. Можно предположить, что если Cas9 доставляется в клетки в форме РНП, то эффект от различных модификаций белка, в том числе добавления сигналов NLS, будет выражен сильнее по сравнению с доставкой Cas9 в виде экспрессионных конструкций. В последнем случае уровень нуклеазы в клетке оказывается гораздо выше, чем при использовании РНП, что может сгладить различия, существенные при невысокой концентрации нуклеазы в клетке. Кроме того, различия в наблюдаемом повышении эффективности

нокаута могут быть связаны с тем, что для оценки этого показателя мы использовали проточную цитофлуориметрию, тогда как S. Shui с соавт. [30] анализировали уровень инделов менее точным методом – T7E1.

Воздействие на пути репарации ДНК с помощью низкомолекулярных соединений с целью повышения уровня нокина описано во многих работах (см. обзор [16]). Однако из-за различий в клеточных моделях и в подобранных концентрациях исследуемых соединений сравнение наблюдаемых эффектов затруднено. При выборе низкомолекулярных соединений мы ориентировались на данные, полученные для Т-клеточных культур и первичных Т-клеток [21, 23, 31]. В результате из 7 протестированных веществ повышение нокина *MTC34* обнаружено лишь при обработке клеток ингибиторами DNA-ПК: M3814 и Ku0060648, – что согласуется с ранее опубликованными данными [23, 31]. Интересно, что для SCR7 и L755507 ранее также не было выявлено влияния на уровень нокина в клетках Jurkat [23], а для RS-1 эффект был зарегистрирован на клеточных линиях НЕК293А [22] и MCF-7 [32], но не на K562 [33]. Стоит отметить, что низкомолекулярные соединения, например M3814, могут быть полезным инструментом фундаментальных исследований, но для их применения в клинической практике возникнут затруднения. Так, показано, что при экспансии отредактированных CAR-T клеток в присутствии M3814 уровень нокина повышался, но существенно снижался суммарный выход клеток [21].

Наконец, при использовании комбинации двух подходов: экспрессия Cas9-3×NLS и обработка клеток M3814 – нами зарегистрирован аддитивный эффект на уровень нокина – повышение на 70%. Заметим, что при введении только дополнительных сигналов NLS в молекулу Cas9 увеличивался как уровень нокина, так и нокаута, а обработка M3814 повышала только уровень нокина. В результате при использовании комбинации Cas9-3×NLS + M3814 уровень нокина *MTC34* стабильно достигал 30–40%. В дальнейшем оптимизированные конструкции будут использованы для внесения нокина в первичные CD4<sup>+</sup> Т-клетки, в которых ранее уровень нокина для конструкций без модификаций составлял менее 1% [8]. Клетки с нокином можно будет дополнительно отсортировать с помощью полученных нами ранее мышинных моноклональных антител против МТ-С34 [8].

Таким образом, нам удалось повысить уровень CRISPR/Cas9-опосредованного нокина конструкции, кодирующей пептидный ингибитор слияния ВИЧ-1, в локус *CXCR4* человека с помощью модификации нуклеазы Cas9 и применения низкомолекулярных соединений – ингибиторов DNA-ПК. Полученные результаты могут быть использованы в разработке генотерапевтических подходов

к лечению ВИЧ-1, а также служить отправной точкой для поиска эффективных способов повышения нокина в релевантной с терапевтической точки зрения модели для экспрессии пептидных ингибиторов слияния ВИЧ.

Работа выполнена при финансовой поддержке Российского научного фонда (грант № 22-25-00310).

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

Авторы заявляют об отсутствии конфликта интересов.

### СПИСОК ЛИТЕРАТУРЫ

- Jinek M., East A., Cheng A., Lin S., Ma E., Doudna J. (2013) RNA-programmed genome editing in human cells. *Elife*. **2**, e00471. <https://doi.org/10.7554/ELIFE.00471>
- Jiang F., Doudna J.A. (2017) CRISPR–Cas9 structures and mechanisms. *Annu. Rev. Biophys.* **46**, 505–529. <https://doi.org/10.1146/annurev-biophys-062215-010822>
- Nambiar T.S., Baudrier L., Billon P., Ciccia A. (2022) CRISPR-based genome editing through the lens of DNA repair. *Mol. Cell*. **82**, 348–388. <https://doi.org/10.1016/j.molcel.2021.12.026>
- Li T., Yang Y., Qi H., Cui W., Zhang L., Fu X., He X., Liu M., Li P.F., Yu T. (2023) CRISPR/Cas9 therapeutics: progress and prospects. *Signal. Transduct. Target Ther.* **8**, 36. <https://doi.org/10.1038/s41392-023-01309-7>
- Pavlovic K., Tristán-Manzano M., Maldonado-Pérez N., Cortijo-Gutierrez M., Sánchez-Hernández S., Justicia-Lirio P., Carmona M.D., Herrera C., Martín F., Benabdelah K. (2020) Using gene editing approaches to fine-tune the immune system. *Front. Immunol.* **11**, 570672. <https://doi.org/10.3389/fimmu.2020.570672>
- Cornu T.I., Mussolino C., Müller M.C., Wehr C., Kern W.V., Cathomen T. (2021) HIV gene therapy: an update. *Hum. Gene Ther.* **32**, 52–65. <https://doi.org/10.1089/HUM.2020.159>
- Liu M., Rehman S., Tang X., Gu K., Fan Q., Chen D., Ma W. (2019) Methodologies for improving HDR efficiency. *Front. Genet.* **9**, 691. <https://doi.org/10.3389/fgene.2018.00691>
- Maslennikova A., Kruglova N., Kalinichenko S., Komkov D., Shepelev M., Golubev D., Siniavin A., Vzorov A., Filatov A., Mazurov D. (2022) Engineering T-cell resistance to HIV-1 infection via knock-in of peptides from the heptad repeat 2 domain of gp41. *mBio*. **13**, e0358921. <https://doi.org/10.1128/mbio.03589-21>
- Dean D.A., Dean B.S., Muller S., Smith L.C. (1999) Sequence requirements for plasmid nuclear import. *Exp. Cell Res.* **253**, 713–722. <https://doi.org/10.1006/EXCR.1999.4716>
- Bai H., Lester G.M.S., Petishnok L.C., Dean D.A. (2017) Cytoplasmic transport and nuclear import of plasmid DNA. *Biosci. Rep.* **37**, BSR20160616. <https://doi.org/10.1042/BSR20160616>
- Vacik J., Dean B.S., Zimmer W.E., Dean D.A. (1999) Cell-specific nuclear import of plasmid DNA. *Gene Ther.* **6**, 1006–1014. <https://doi.org/10.1038/sj.gt.3300924>
- Young J.L., Benoit J.N., Dean D.A. (2003) Effect of a DNA nuclear targeting sequence on gene transfer and expression of plasmids in the intact vasculature. *Gene Ther.* **10**, 1465–1470. <https://doi.org/10.1038/sj.gt.3302021>
- Mesika A., Grigoreva I., Zohar M., Reich Z. (2001) A regulated, NFκB-assisted import of plasmid DNA into mammalian cell nuclei. *Mol. Ther.* **3**, 653–657. <https://doi.org/10.1006/mthe.2001.0312>
- 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. <https://doi.org/10.1126/science.1231143>
- Maggio I., Zittersteijn H.A., Wang Q., Liu J., Janssen J.M., Ojeda I.T., van der Maarel S.M., Lankester A.C., Hoeben R.C., Gonçalves M.A.F.V. (2020) Integrating gene delivery and gene-editing technologies by adenoviral vector transfer of optimized CRISPR–Cas9 components. *Gene Ther.* **27**, 209–225. <https://doi.org/10.1038/s41434-019-0119-y>
- Shams F., Bayat H., Mohammadian O., Mahboudi S., Vahidnezhad H., Soosanabadi M., Rahimpour A. (2022) Advance trends in targeting homology-directed repair for accurate gene editing: an inclusive review of small molecules and modified CRISPR–Cas9 systems. *BioImpacts*. **12**, 371–391. <https://doi.org/10.34172/bi.2022.23871>
- Makkerh J.P.S., Dingwall C., Laskey R.A. (1996) Comparative mutagenesis of nuclear localization signals reveals the importance of neutral and acidic amino acids. *Curr. Biol.* **6**, 1025–1027. [https://doi.org/10.1016/S0960-9822\(02\)00648-6](https://doi.org/10.1016/S0960-9822(02)00648-6)
- Zotova A., Pichugin A., Atemasova A., Knyazhanskaya E., Lopatukhina E., Mitkin N., Holmuhamedov E., Gotikh M., Kuprash D., Filatov A., Mazurov D. (2019) Isolation of gene-edited cells via knock-in of short glycoposphatidylinositol-anchored epitope tags. *Sci. Rep.* **9**, 3132. <https://doi.org/10.1038/S41598-019-40219-Z>
- Shin S., Kim S.H., Lee J.S., Lee G.M. (2021) Streamlined human cell-based recombinase-mediated cassette exchange platform enables multigene expression for the production of therapeutic proteins. *ACS Synth. Biol.* **10**, 1715–1727. <https://doi.org/10.1021/acssynbio.1c00113>
- Nguyen D.N., Roth T.L., Li P.J., Chen P.A., Apathy R., Mamedov M.R., Vo L.T., Tobin V.R., Goodman D., Shifrut E., Bluestone J.A., Puck J.M., Szoka F.C., Marson A. (2020) Polymer-stabilized Cas9 nanoparticles and modified repair templates increase genome editing efficiency. *Nat. Biotechnol.* **38**, 44–49. <https://doi.org/10.1038/s41587-019-0325-6>
- Shy B.R., Vykunta V.S., Ha A., Talbot A., Roth T.L., Nguyen D.N., Pfeifer W.G., Chen Y.Y., Blaeschke F., Shifrut E., Vedova S., Mamedov M.R., Chung J.J., Li H., Yu R., Wu D., Wolf J., Martin T.G., Castro C.E., Ye L., Esensten J.H., Eyquem J., Marson A. (2023) High-yield genome engineering in primary cells using a hybrid ssDNA repair template and small-molecule cocktails. *Nat. Biotechnol.* **41**, 521–531. <https://doi.org/10.1038/s41587-022-01418-8>

22. Pinder J., Salsman J., Delleire G. (2015) Nuclear domain 'knock-in' screen for the evaluation and identification of small molecule enhancers of CRISPR-based genome editing. *Nucleic Acids Res.* **43**, 9379–9392. <https://doi.org/10.1093/nar/gkv993>
23. Kath J., Du W., Pruene A., Braun T., Thommandru B., Turk R., Sturgeon M.L., Kurgan G.L., Amini L., Stein M., Zittel T., Martini S., Ostendorf L., Wilhelm A., Akyüz L., Rehm A., Höpken U.E., Pruß A., Künkele A., Jacobi A.M., Volk H.D., Schmueck-Henneresse M., Stripecke R., Reinke P., Wagner D.L. (2022) Pharmacological interventions enhance virus-free generation of TRAC-replaced CAR T cells. *Mol. Ther. Methods Clin. Dev.* **25**, 311–330. <https://doi.org/10.1016/j.omtm.2022.03.018>
24. Young J.L., Zimmer W.E., Dean D.A. (2008) Smooth muscle-specific gene delivery in the vasculature based on restriction of DNA nuclear import. *Exp. Biol. Med.* **233**, 840–848. <https://doi.org/10.3181/0712-RM-331>
25. Degiulio J.V., Kaufman C.D., Dean D.A. (2010) The SP-C promoter facilitates alveolar type II epithelial cell-specific plasmid nuclear import and gene expression. *Gene Ther.* **17**, 541–549. <https://doi.org/10.1038/gt.2009.166>
26. Schulze-Luehrmann J., Ghosh S. (2006) Antigen-receptor signaling to nuclear factor  $\kappa$ B. *Immunity.* **25**, 701–715. <https://doi.org/10.1016/j.immuni.2006.10.010>
27. Wu W., Nie L., Zhang L., Li Y. (2018) The notch pathway promotes NF- $\kappa$ B activation through Asb2 in T cell acute lymphoblastic leukemia cells. *Cell. Mol. Biol. Lett.* **23**, 37. <https://doi.org/10.1186/s11658-018-0102-4>
28. Castro-caldas M., Mendes A.F., Carvalho A.P., Duarte C.B., Lopes M.C. (2003) Dexamethasone prevents interleukin-1 $\beta$ -induced nuclear factor- $\kappa$ B activation by up-regulating I $\kappa$ B- $\alpha$  synthesis, in lymphoblastic cells. *Mediators Inflamm.* **12**, 37–46. <https://doi.org/10.1080/0962935031000096953>
29. Remy S., Chenouard V., Tesson L., Usal C., Ménoret S., Brusselle L., Heslan J.M., Nguyen T.H., Bellien J., Merot J., De Cian A., Giovannangeli C., Concordet J.P., Aneon I. (2017) Generation of gene-edited rats by delivery of CRISPR/Cas9 protein and donor DNA into intact zygotes using electroporation. *Sci. Rep.* **7**, 16554. <https://doi.org/10.1038/s41598-017-16328-y>
30. Shui S., Wang S., Liu J. (2022) Systematic investigation of the effects of multiple SV40 nuclear localization signal fusion on the genome editing activity of purified SpCas9. *Bioengineering.* **9**, 83. <https://doi.org/10.3390/bioengineering9020083>
31. Fu Y.-W., Dai X.Y., Wang W.T., Yang Z.X., Zhao J.J., Zhang J.P., Wen W., Zhang F., Oberg K.C., Zhang L., Cheng T., Zhang X.B. (2021) Dynamics and competition of CRISPR-Cas9 ribonucleoproteins and AAV donor-mediated NHEJ, MMEJ and HDR editing. *Nucleic Acids Res.* **49**, 969–985. <https://doi.org/10.1093/nar/gkaa1251>
32. Killian T., Dickopf S., Haas A.K., Kirstenpfad C., Mayer K., Brinkmann U. (2017) Disruption of diphthamide synthesis genes and resulting toxin resistance as a robust technology for quantifying and optimizing CRISPR/Cas9-mediated gene editing. *Sci. Rep.* **7**, 15480. <https://doi.org/10.1038/s41598-017-15206-x>
33. Wienert B., Nguyen D.N., Guenther A., Feng S.J., Locke M.N., Wyman S.K., Shin J., Kazane K.R., Gregory G.L., Carter M.A.M., Wright F., Conklin B.R., Marson A., Richardson C.D., Corn J.E. (2020) Timed inhibition of CDC7 increases CRISPR-Cas9 mediated templated repair. *Nat. Commun.* **11**, 2109. <https://doi.org/10.1038/s41467-020-15845-1>

## Increasing the Level of Knock-in of a Construct Encoding the HIV-1 Fusion Inhibitor, MT-C34 Peptide, into the *CXCR4* Locus in the CEM/R5 T Cell Line

D. S. Golubev<sup>1</sup>, D. S. Komkov<sup>1,2</sup>, M. V. Shepelev<sup>1</sup>, D. V. Mazurov<sup>1,3</sup>, N. A. Kruglova<sup>1</sup>, \*

<sup>1</sup>*Center for Precision Genome Editing and Genetic Technologies for Biomedicine, Institute of Gene Biology, Russian Academy of Sciences, Moscow, 119334 Russia*

<sup>2</sup>*Department of Physiology and Cell Biology, Faculty of Health Sciences, Ben-Gurion University of the Negev, Beer-Sheva, 8410501 Israel*

<sup>3</sup>*Division of Infectious Diseases and International Medicine, Department of Medicine, University of Minnesota, Minneapolis, 55455 USA*

\*e-mail: natalya.a.kruglova@yandex.ru

The low efficiency of knock-in, especially in primary human cells, limits the use of genome editing technology for therapeutic purposes, which makes it important to develop approaches for increasing knock-in levels. In this work, using a knock-in model of the peptide fusion inhibitor of HIV MT-C34 into the human *CXCR4* locus in the CEM/R5 T cell line, we analyzed the effectiveness of several approaches to increasing knock-in levels. First, donor DNA modification aimed at improving the efficiency of plasmid transport into the nucleus was evaluated, namely the introduction into the donor plasmid of the SV40 DNA transport sequence (DTS) or the binding sites for the transcription factor NF- $\kappa$ B, whose effects on knock-in levels have not been described. In the MT-C34 knock-in model into the *CXCR4* locus, this modification was ineffective. The second approach, modifying the Cas9 nuclease by introducing two additional nuclear localization signals (NLS), increased the knock-in level by 30%. Finally, blocking DNA repair via the nonhomologous end joining pathway using DNA-dependent protein kinase inhibitors caused a 1.8-fold increase in knock-in. The combination of the last two approaches caused an additive effect. Thus, increasing the number of NLSs in the Cas9 protein and inhibiting DNA repair via the nonhomologous end joining pathway significantly increased the level of knock-in of the HIV-1 peptide fusion inhibitor into the clinically relevant locus *CXCR4*, which can be used to develop effective gene therapy approaches for the treatment of HIV infection.

**Keywords:** CRISPR/Cas9, genome editing, homology-directed repair, T lymphocytes, knock-in, NHEJ inhibitors, DTS, NLS

## ПОВЫШЕНИЕ УРОВНЯ НОКИНА МТ-С34-КОДИРУЮЩЕЙ КОНСТРУКЦИИ В ЛОКУС *CXCR4* С ПОМОЩЬЮ МОДИФИКАЦИИ ДОНОРНОЙ ДНК САЙТАМИ-МИШЕНЯМИ Cas9

© 2024 г. М. В. Шепелев<sup>а</sup>, Д. С. Комков<sup>а, б</sup>, Д. С. Голубев<sup>а</sup>, С. Е. Боровикова<sup>с</sup>,  
Д. В. Мазуров<sup>а, д</sup>, Н. А. Круглова<sup>а, \*</sup>

<sup>а</sup>Центр высокоточного редактирования и генетических технологий для биомедицины,  
Институт биологии гена Российской академии наук, Москва, 119334 Россия

<sup>б</sup>Department of Physiology and Cell Biology, Faculty of Health Sciences, Ben-Gurion University of the Negev,  
Beer-Sheva, 8410501 Israel

<sup>с</sup>Институт биологии гена Российской академии наук, Москва, 119334 Россия

<sup>д</sup>Division of Infectious Diseases and International Medicine, Department of Medicine, University of Minnesota,  
Minneapolis, 55455 USA

\*e-mail: natalya.a.kruglova@yandex.ru

Поступила в редакцию 03.11.2023 г.

После доработки 04.12.2023 г.

Принята к публикации 04.12.2023 г.

Для успешного применения в клинической практике технологии редактирования генома с помощью системы CRISPR/Cas9 необходимо добиться высокой эффективности нокина, то есть встраивания генетической конструкции в заданное место генома клетки-мишени. Один из подходов к повышению эффективности нокина заключается в модификации донорной ДНК путем введения в нее тех же мишеней для Cas9 (Cas9 targeting sequence, CTS), которые используются для внесения двухцепочечного разрыва в геном клетки (метод “double-cut donor”). Другой подход основан на введении в донорную ДНК укороченных мишеней для Cas9 (truncated, tCTS), включающих сайт PAM и 16 проксимальных к нему нуклеотидов, которые предположительно не должны вызывать расщепление донорной плазмиды, но способны стимулировать ее транспорт в ядро за счет Cas9. Однако точные механизмы повышения уровня нокина под действием обоих типов модификаций донорной ДНК неизвестны. Мы оценили влияние этих модификаций донорной ДНК на эффективность нокина конструкции, кодирующей пептидный ингибитор слияния ВИЧ-1 МТ-С34, в локус *CXCR4* в Т-клеточной линии СЕМ/R5. При введении в донорную плазмидную ДНК полноразмерных CTS уровень нокина повышался в 2 раза вне зависимости от числа CTS или их положения относительно донорной последовательности. Модификации tCTS не влияли на уровень нокина. Установлено, что *in vitro* оба типа сайтов эффективно расщеплялись под действием Cas9. Для детального изучения механизма действия этих модификаций необходимо оценивать их процессинг *in vitro* и *in vivo*.

**Ключевые слова:** CRISPR/Cas9, редактирование генома, Т-лимфоциты, knock-in, донорная ДНК, мишени Cas9, ядерный транспорт ДНК

DOI: 10.31857/S0026898424040058, EDN: INCOYT

Появление технологий редактирования генома с помощью программируемых нуклеаз открыло новые возможности для лечения различных заболеваний человека [1]. Наиболее широко используемая нуклеаза Cas9 с помощью короткой молекулы

РНК-гида (гРНК) распознает ДНК-мишень и вносит двухцепочечный разрыв (DSB) в геном [2]. Репарация DSB в клетках происходит преимущественно по механизму негомологичного соединения концов (non-homologous end joining, NHEJ), который вызывает появление небольших делеций или инсерций (инделов), что можно использовать для нокаута гена. Репарация по механизму гомологичной рекомбинации (homology-directed repair, HDR) менее эффективна по сравнению с NHEJ и требует присутствия ДНК-матрицы (донорной ДНК). Доставка в клетки экзогенной донорной ДНК позволяет вносить в геном целевые нуклеотидные последовательности, что называется

Сокращения: гРНК – РНК-гид; РНП – рибонуклеопротеин; DSB (double-strand breaks) – двухцепочечные разрывы; CTS (Cas9 targeting sequence) – последовательность-мишень Cas9; HDR (homology-directed repair) – репарация по пути гомологичной рекомбинации; NHEJ (non-homologous end joining) – негомологичное соединение концов; tCTS (truncated Cas9 targeting sequence) – укороченная последовательность-мишень Cas9.

knock-in (нокин). Для генной терапии ряда заболеваний человека с помощью системы CRISPR/Cas9 можно использовать нокаут генов [3, 4]. Однако наиболее интересна возможность внесения направленных изменений в геном с помощью нокина, например, для исправления точечных мутаций при лечении моногенных заболеваний [5] или для создания CAR-T клеток [4]. К сожалению, эффективность CRISPR/Cas9-опосредованного нокина в большинстве случаев остается низкой, что ограничивает использование этой технологии в клинической практике. В связи с этим разработана новых подходов к повышению эффективности нокина, в первую очередь на клинически релевантных моделях в первичных клетках человека, не теряет своей актуальности.

Описаны различные способы модификации Cas9, гРНК и донорной ДНК, приводящие к повышению уровня нокина [6]. Кроме того, можно воздействовать на клеточный цикл или механизмы репарации ДНК с помощью низкомолекулярных соединений, чтобы сместить равновесие с механизма NHEJ в сторону HDR [7, 8]. Ряд подходов, направленных на модификацию донорной ДНК, включает использование различных типов ДНК-матриц (одноцепочечные олигонуклеотиды, оцДНК, дцДНК и т.д.), варьирование длины и симметрии плеч гомологии, химические модификации одноцепочечных олигонуклеотидов и др. [9]. Например, для конструкции длиной более 100 п.н. уровень нокина коррелирует с длиной плеч гомологии, однако достигает плато после некоторого значения (800–1500 п.н.) и затем падает из-за возрастающей токсичности ДНК [10]. Показано, что модификации концов линейной молекулы ДНК триэтиленгликолем и/или 2'-О-метил-рибонуклеотидами поддерживают стабильность донора и тем самым способствуют процессу HDR [11].

Для успешного нокина донорная ДНК должна находиться в ядре в момент образования DSB. При невирусном способе доставки донора какие-либо специфические механизмы его транспорта в ядро отсутствуют. В делящихся клетках трансфицированная ДНК оказывается в ядре в результате циклического разрушения и формирования ядерной мембраны [12]. Для повышения эффективности доставки донора к месту DSB и его пространственного сближения с комплексом Cas9/гРНК предложены различные способы ковалентной сшивки донора с гРНК или Cas9 [13–16]. Однако все эти способы требуют дополнительных этапов подготовки *in vitro* и чаще всего снижают способность комплекса расщеплять мишень. В результате повышение уровня HDR, описываемое для таких модификаций, на самом деле представляет собой увеличение соотношения HDR/NHEJ, в то время как абсолютный уровень HDR падает [13–16].

Кроме таких технически сложных методов описано два способа модификации донорной ДНК путем введения на 5'- и 3'-концы последовательностей-мишеней для Cas9, которые представляют собой протоспейсер и следующий за ним сайт РАМ (protospacer adjacent motif) [10, 17]. Первый тип модификации донора, включающий полноразмерный протоспейсер длиной 20 нуклеотидов, был назван "double-cut" [10], потому что Cas9 расщепляет такую донорную ДНК в двух местах, фланкирующих плечи гомологии. Показано, что одновременное внесение разрыва в геномную мишень и в донорную ДНК приводит к повышению уровня нокина более чем в 10 раз [10]. Однако механизм такого эффекта детально не исследован.

Авторы второго способа предположили, что, если Cas9 будет связываться с донорной ДНК в цитоплазме, но не будет вносить в нее разрыв, то донор в комплексе с Cas9 может эффективно транспортироваться в ядро за счет белка Cas9, имеющего сигналы ядерной локализации (NLS) [17]. Для этого донорные молекулы модифицировали путем добавления укороченных последовательностей-мишеней Cas9 (truncated Cas9 targeting sequence, tCTS) с протоспейсером длиной 16 нуклеотидов, которые, как показано J. Zhang с соавт. [18], не расщепляются под действием Cas9. Влияние модификации донорной ДНК сайтами tCTS на уровень нокина было изучено в ряде работ [19–22], при этом сообщалось как о повышении нокина в 2–4 [17, 19] или в 1.5 раза [20, 22], так и об отсутствии влияния [21]. Важно отметить, что в перечисленных работах не проверяли способность Cas9 расщеплять tCTS-сайты *in vitro* или в клетках. Кроме того, не было исследовано, действительно ли донорная ДНК с такими модификациями в большей степени накапливается в ядре по сравнению с немодифицированной.

В представленной работе на клинически значимой модели нокина конструкции, кодирующей пептидный ингибитор слияния ВИЧ-1 МТ-С34 (далее: нокин МТС34), в локус *CXCR4* в Т-клеточной линии СЕМ/Р5 [23] была оценена эффективность двух способов повышения уровня нокина: модификация донорной плазмиды по методу "double-cut" [10] и с помощью введения tCTS-сайтов, которые, предположительно, способствуют переносу донорной плазмиды в ядро за счет белка Cas9 [17].

## ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

**Клеточные линии.** Клетки линии СЕМ/Р5, полученные на основе Т-клеточной линии ССRF-СЕМ (АТСС) [23], культивировали в среде DMEM/F12 ("ПанЭко", Россия) с добавлением 10% фетальной сыворотки крупного рогатого скота ("HyClone", США), 4 мМ L-глутамин и 40 мг/л гентамицина

Таблица 1. Праймеры, использованные в работе

Клонирование полноразмерных сигналов CTS в плазмидный донор	
на 5'-конец	
5'-Not-X4ex2-Psp	<b>GGCCGAATTCACTTCAGATAACTACACCGAGG</b>
3'-Not-X4ex2-Psp	TCGACCTCGGTGTAGTTATCTGAAGTGAATTC
на 3'-конец	
5'-Xba-X4ex2-Nco	CTAGCCTCGGTGTAGTTATCTGAAGTGAATTC
3'-Xba-X4ex2-Nco	<b>CATGGAATTCACCTTCAGATAACTACACCGAGG</b>
Клонирование tCTS с 4 некомплементами основаниями в плазмидный донор	
на 5'-конец	
5'-tCTS(4MM)-Not-X4ex2-Psp	<b>GGCCGAattTCAGATAACTACACCGAGG</b>
3'-tCTS(4MM)-Not-X4ex2-Psp	TCGACCTCGGTGTAGTTATCTGAGAATTC
на 3'-конец	
5'-tCTS(4MM)-Xba-X4ex2-Nco	CTAGCCTCGGTGTAGTTATCTGAGAATTC
3'-tCTS(4MM)-Xba-X4ex2-Nco	<b>CATGGAattTCAGATAACTACACCGAGG</b>
Клонирование tCTS с 6 некомплементами основаниями в плазмидный донор	
на 5'-конец	
5'-tCTS(6MM)-Not-X4ex2-Psp	<b>GGCCGAattcagAGATAACTACACCGAGG</b>
3'-tCTS(6MM)-Not-X4ex2-Psp	TCGACCTCGGTGTAGTTATCTCTGAATTC
на 3'-конец	
5'-tCTS(6MM)-Xba-X4ex2-Nco	CTAGCCTCGGTGTAGTTATCTCTGAATTC
3'-tCTS(6MM)-Xba-X4ex2-Nco	<b>CATGGAattcagAGATAACTACACCGAGG</b>
Клонирование tCTS с 8 некомплементами основаниями в плазмидный донор	
на 5'-конец	
5'-tCTS(8MM)-Not-X4ex2-Psp	<b>GGCCGAattcagtcATAACTACACCGAGG</b>
3'-tCTS(8MM)-Not-X4ex2-Psp	TCGACCTCGGTGTAGTTATGACTGAATTC
на 3'-конец	
5'-tCTS(8MM)-Xba-X4ex2-Nco	CTAGCCTCGGTGTAGTTATGACTGAATTC
3'-tCTS(8MM)-Xba-X4ex2-Nco	<b>CATGGAattcagtcATAACTACACCGAGG</b>

Примечание. Для каждого праймера, содержащего целевую последовательность, жирным шрифтом выделен сайт PAM, протоспейсер подчеркнут, некомплементами основаниями в его составе показаны строчными символами.

(“ПанЭко”) при 37°C в увлажненной атмосфере 5% CO<sub>2</sub>.

**Конструирование плазмид.** Донорные плазмиды получали на основе плазмиды pJet-donor<sub>MT-C34</sub>, описанной ранее [23]. Для создания донорной плазмиды pJet-donor<sub>MT-C34</sub> с сайтом связывания Cas9 (Cas9 targeting sequence, CTS) на 5'-конце донорной последовательности, несущим 0, 4, 6 или 8 некомплементами оснований, соответствующие одноцепочечные олигонуклеотиды: 5'-Not-X4ex2-Psp и 3'-Not-X4ex2-Psp – гибридизовали и клонировали в плазмиду pJet-donor<sub>MT-C34</sub> по сайтам NotI-PspXI (все последовательности праймеров приведены в табл. 1, а список использованных и полученных в работе плазмид – в табл. 2).

Для создания донорной плазмиды pJet-donor<sub>MT-C34</sub> с CTS-сайтом на 3'-конце донорной последовательности, несущим 0, 4, 6 или 8 некомплементами оснований, соответствующие олигонуклеотиды: 5'-Xba-X4ex2-Nco и 3'-Xba-X4ex2-Nco – гибридизовали и клонировали в плазмиду pJet-donor<sub>MT-C34</sub> по сайтам XbaI-NcoI. Для получения донора с двумя CTS плазмиду pJet-donor<sub>MT-C34</sub>-5'CTS обрабатывали рестриктазами XmaI-Bsu36I, фрагмент размером 566 п.н. клонировали в плазмиду pJet-donor<sub>MT-C34</sub>-3'CTS по соответствующим сайтам. Полученные последовательности донорной ДНК в векторе pJet переклонировали в вектор pBluescript KS(+) (“Stratagene”),

Таблица 2. Плазмидные конструкции, использованные в работе

Название	Описание	Источник/ссылка
pKSgRNA-X4ex2	Вектор для экспрессии гРНК против экзона-2 локуса <i>CXCR4</i>	[23]
pJET1.2		“Thermo Fisher Scientific”
pBluescript KS (+)		“Stratagene”
pJet-donor_MT-C34	Немодифицированный плазмидный донор для нокина конструкции, кодирующей пептид MT-C34, в экзон-2 локуса <i>CXCR4</i>	[23]
pKS-donor_MT-C34	Немодифицированный плазмидный донор для нокина конструкции, кодирующей пептид MT-C34, в экзон-2 локуса <i>CXCR4</i>	Не опубликована
pKS-donor_MT-C34-CTS-5'	Плазмидный донор для нокина конструкции, кодирующей пептид MT-C34, в экзон-2 локуса <i>CXCR4</i> , модифицированный по 5'-концу полноразмерным CTS	Эта работа
pKS-donor_MT-C34-CTS-3'	Плазмидный донор для нокина конструкции, кодирующей пептид MT-C34, в экзон-2 локуса <i>CXCR4</i> , модифицированный по 3'-концу полноразмерным CTS	Эта работа
pKS-donor_MT-C34-CTS-5'-3'	Плазмидный донор для нокина конструкции, кодирующей пептид MT-C34, в экзон-2 локуса <i>CXCR4</i> , модифицированный по 5'- и 3'-концам полноразмерным CTS	Эта работа
pKS-donor_MT-C34-tCTS(4MM)-5'	Плазмидный донор для нокина конструкции, кодирующей пептид MT-C34, в экзон-2 локуса <i>CXCR4</i> , модифицированный по 5'-концу tCTS 4 некомплементарными основаниями	Эта работа
pKS-donor_MT-C34-tCTS(4MM)-3'	Плазмидный донор для нокина конструкции, кодирующей пептид MT-C34, в экзон-2 локуса <i>CXCR4</i> , модифицированный по 3'-концу tCTS с 4 некомплементарными основаниями	Эта работа
pKS-donor_MT-C34-tCTS(4MM)-5'-3'	Плазмидный донор для нокина конструкции, кодирующей пептид MT-C34, в экзон-2 локуса <i>CXCR4</i> , модифицированный по 5'- и 3'-концам tCTS с 4 некомплементарными основаниями	Эта работа
pKS-donor_MT-C34-tCTS(6MM)-5'	Плазмидный донор для нокина конструкции, кодирующей пептид MT-C34, в экзон-2 локуса <i>CXCR4</i> , модифицированный по 5'-концу tCTS с 6 некомплементарными основаниями	Эта работа
pKS-donor_MT-C34-tCTS(6MM)-3'	Плазмидный донор для нокина конструкции, кодирующей пептид MT-C34, в экзон-2 локуса <i>CXCR4</i> , модифицированный по 3'-концу tCTS с 6 некомплементарными основаниями	Эта работа
pKS-donor_MT-C34-tCTS(6MM)-5'-3'	Плазмидный донор для нокина конструкции, кодирующей пептид MT-C34, в экзон-2 локуса <i>CXCR4</i> , модифицированный по 5'- и 3'-концам tCTS с 6 некомплементарными основаниями	Эта работа
pKS-donor_MT-C34-tCTS(8MM)-5'	Плазмидный донор для нокина конструкции, кодирующей пептид MT-C34, в экзон-2 локуса <i>CXCR4</i> , модифицированный по 5'-концу tCTS с 8 некомплементарными основаниями	Эта работа
pKS-donor_MT-C34-tCTS(8MM)-3'	Плазмидный донор для нокина конструкции, кодирующей пептид MT-C34, в экзон-2 локуса <i>CXCR4</i> , модифицированный по 3'-концу tCTS с 8 некомплементарными основаниями	Эта работа
pKS-donor_MT-C34-tCTS(8MM)-5'-3'	Плазмидный донор для нокина конструкции, кодирующей пептид MT-C34, в экзон-2 локуса <i>CXCR4</i> , модифицированный по 5'- и 3'-концам tCTS с 8 некомплементарными основаниями	Эта работа

США) по сайтам PstI/ClaI. Все эксперименты проводили с донорами в векторе pBluescript KS(+).

Нуклеотидные последовательности всех полученных плазмид верифицировали секвенированием по Сэнгеру в компании “Евроген” (Россия). Донорные плазмиды выделяли с помощью набора Plasmid Midi Kit (100) (#12145; “Qiagen”, Германия). Все остальные плазмиды выделяли с использованием набора Plasmid Midiprep 2.0 (#BC124; “Евроген”).

**Электropорация.** Для нокина *MTC34* в локус *CXCR4* электропорировали  $1 \times 10^6$  клеток СЕМ/R5, используя 3 мкг плазмиды pcDNA3.3-hCas9 (#41815; “Addgene”, США) и 1 мкг плазмиды pKS-gRNA-X4ex2, кодирующей гРНК против экзона-2 гена *CXCR4* (спейсер: 5'-САСТТСАГАТААСТАСАССГ-3', РАМ: АGG), которая описана ранее [23], вместе с 1 пмоль (~2.9 мкг) соответствующего плазмидного донора. Кроме гРНК против *CXCR4*, использовали гРНК против последовательности, кодирующей белок р24 ВИЧ-1 (спейсер: 5'-GTТААААGAGАССАТСААТG-3', РАМ: АGG), описанную ранее [23], а также контрольную гРНК с рандомизированной последовательностью, не имеющей комплементарных участков в геноме человека (спейсер: 5'-GСАСТАССАGAGСТААСТСА-3'). Для электропорации использовали прибор Neon electroporation system со 100-микролитровыми накопниками (“Invitrogen”, США) и следующими настройками: 1230 В, 40 мс, один импульс. На 5 сутки после электропорации клетки окрашивали соответствующими антителами и анализировали методом проточной цитофлуориметрии.

**Проточная цитофлуориметрия.** Для иммунофлуоресцентного окрашивания клетки СЕМ/R5 инкубировали в течение 30 мин при 4°C в фосфатно-солевом буфере (PBS) с мышинными моноклональными антителами С24 против пептида МТ-С34, описанными ранее [23], или с мышинными моноклональными антителами против *CXCR4* (клон 12G5; “Santa Cruz Biotechnology”, США). Затем клетки дважды отмывали в PBS и инкубировали с козьими антителами против IgG мыши, конъюгированными с Alexa488 или Alexa546 (“Thermo Fisher Scientific”, США) в течение 30 мин при 4°C. Далее клетки два раза промывали в PBS и анализировали на проточном цитофлуориметре CytoFLEX S (“Beckman Coulter”, США). В качестве отрицательного контроля использовали клетки, окрашенные только вторичными антителами.

**Расщепление плазмидной ДНК под действием рибонуклеопротеина (РНП) *in vitro*.** Для сборки РНП-комплексов использовали белок Cas9, неработанный и очищенный в лаборатории по протоколу, описанному ранее [23]. Cas9 разводили в буфере Orange (BO5; “Thermo Fisher Scientific”) до концентрации 3 мкМ. гРНК получали с помощью набора для *in vitro* транскрипции HiScribe T7 High Yield RNA synthesis kit (#E2040S; “NEB”, США), как описано в статье А. Масленниковой

(Maslennikova) с соавт. [23]. гРНК разводили в воде (“Ambion Thermo Scientific”, США) до концентрации 3 мкМ. На одну реакцию для сборки РНП брали 3 мкл 10× Orange, по 3 мкл (9 пмоль) Cas9 и гРНК и 18 мкл воды, инкубировали смесь 15 мин при комнатной температуре, после чего делали 27-микролитровые аликвоты и в каждую добавляли 3 мкл плазмидной ДНК с концентрацией 100 мкг/мл (0.037 мкМ). Реакционная смесь содержала 9 пмоль РНП и 0.111 пмоль плазмиды, то есть 81-кратный молярный избыток РНП. Реакцию проводили в течение 30 мин при 37°C, затем добавляли 1 мкл протеиназы К (20 мг/мл) и инкубировали 10 мин при 56°C. В образцы добавляли 6× буфер для нанесения образцов и 1/2 образца разделяли в 1%-ном ТАЕ-агарозном геле.

Для эксперимента по титрованию РНП собирали комплексы, как описано выше, получали общую смесь с концентрацией РНП 0.3 мкМ и готовили 5 последовательных двукратных разведений. Реакционная смесь содержала 3 мкл (0.9 пмоль) РНП и 5 мкл (0.185 пмоль) плазмидной ДНК, инкубацию проводили в описанных выше условиях.

**Статистический анализ** проводили с помощью программы GraphPad Prism 8.0 (США). Для сравнения средних использовали однофакторный ANOVA с последующим тестом Даннета для множественных сравнений с контрольной группой.

## РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ

### *Влияние CTS в донорной плазмиде на уровень нокина*

В качестве модели использовали нокин конструкции, кодирующей пептидный ингибитор слияния ВИЧ-1 МТ-С34, в экзон-2 гена *CXCR4* человека, как описано нами ранее [23]. Донорная конструкция несет плечи гомологии для локуса *CXCR4*, фланкирующие последовательность, кодирующую пептид МТ-С34 (рис. 1а). В результате корректного нокина в начало кодирующей последовательности гена *CXCR4* клетки теряют корецептор *CXCR4* (нокаут) и начинают экспрессировать пептид МТ-С34 на своей поверхности (нокин).

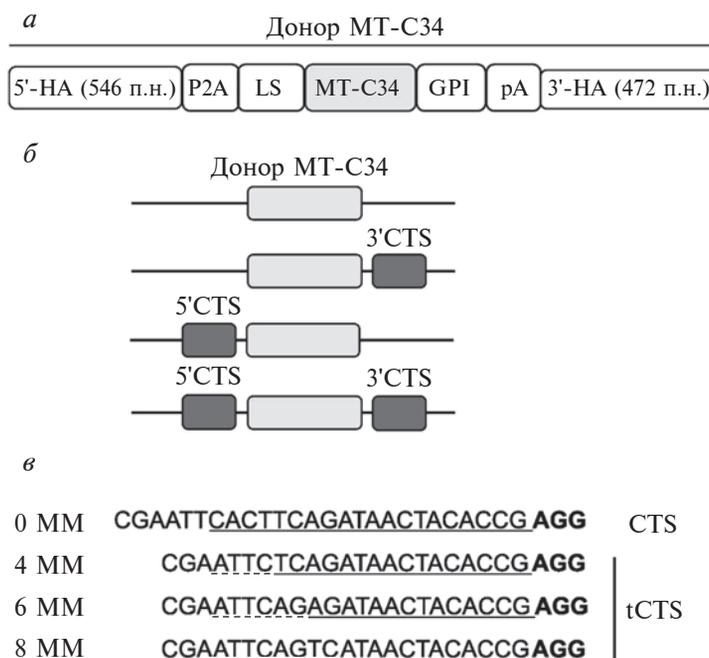
Для оценки влияния полноразмерных CTS на уровень нокина были получены плазмиды, несущие мишени для Cas9 на 5'- или 3'-конце донорной последовательности или с обеих сторон от нее (рис. 1б). Сайты CTS представляли собой прото-спейсер длиной 20 п.н. и РАМ, соответствующие выбранной целевой последовательности в локусе *CXCR4*. Для оценки влияния укороченных сайтов, tCTS, на уровень нокина были получены донорные плазмиды, модифицированные tCTS с 4, 6 или 8 некомплементарными основаниями в протоспейсере в дистальной области от РАМ (рис. 1в).

Полученные донорные плазмиды вместе с плазмидами, кодирующими Cas9 и гРНК, электропорировали в клетки СЕМ/R5 и на 5 сутки измеряли уровень нокина и нокаута с помощью проточной цитофлуориметрии. Обнаружено, что конструкции с полноразмерными сайтами CTS повышали уровень нокина в 1.5–2 раза по сравнению с немодифицированным донором (контроль) вне зависимости от количества или их положения относительно донорной последовательности, тогда как все сайты tCTS, содержащие некомплементарные основания, не влияли на уровень нокина (рис. 2а). Уровень нокаута для всех донорных конструкций не отличался от такового для немодифицированного донора (рис. 2б). В результате повышалось и соотношение уровней нокина к нокауту: примерно в 2 раза для доноров с сайтами CTS (рис. 2в).

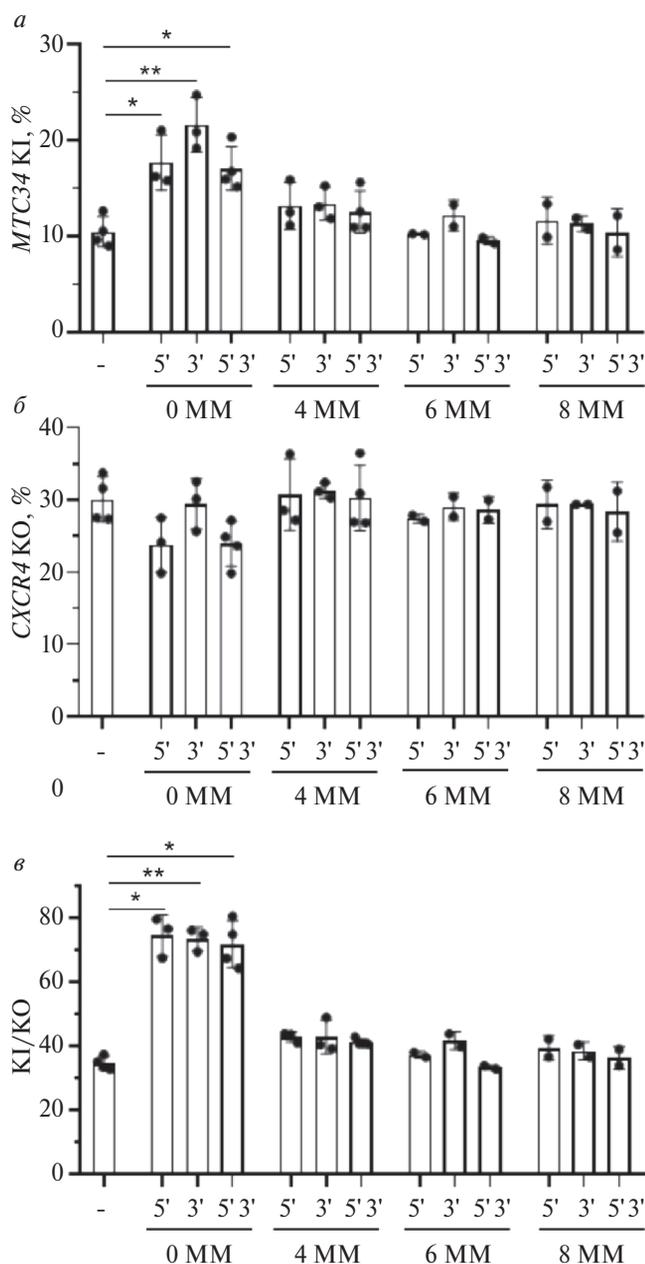
*Расщепление CTS-сайтов под действием Cas9 in vitro*

С целью оценить способность Cas9 в комплексе с гРНК расщеплять донорные плазмиды с CTS- или tCTS-сайтами *in vitro* мы использовали плазмиды, в которых эти сайты находились по обеим сторонам донорной последовательности. По нашим расчетам, в этом случае под действием Cas9

из плазмид будет вырезаться фрагмент размером около 1400 п.н. В результате было обнаружено, что донорная плазида с 4 некомплементарными основаниями (4ММ) расщеплялась Cas9 так же эффективно, как и плазида с полноразмерными CTS (рис. 3а). Плазида с 6ММ в протоспейсере также расщеплялась под действием Cas9, однако менее эффективно (рис. 3а, дорожка 3), и лишь в случае 8ММ расщепления не происходило (рис. 3а, дорожка 2). Важно отметить, что контрольная плазида с немодифицированной донорной последовательностью (рис. 3а, дорожка 1) не расщеплялась Cas9, как и все плазмиды в комплексе с контрольной гРНК (scrambled guide RNA, гРНК-Scr), не имеющей мишеней в геноме человека (рис. 3б). При расщеплении ДНК с помощью Cas9 *in vitro* обычно используют молярный избыток РНП, так как Cas9 – фермент одного цикла [24]. В эксперименте, представленном на рис. 3, использовали 81-кратный молярный избыток РНП. Предположив, что расщепление некомплементарных сайтов связано с избытком РНП, мы провели эксперимент с титрованием количества РНП. Как видно из рис. 3в и 3г, плазмидный субстрат с 4 некомплементарными основаниями (4ММ) расщеплялся так же эффективно, как и не имеющий



**Рис. 1.** Схема донорной конструкции для нокина *MTC34* в экзон-2 локуса *CXCR4* (а) и ее модификаций (б, в). *Обозначения:* 5'-НА – 5'-плечо гомологии; P2A – сигнал пропуска рибосомы; LS – лидерная последовательность; GPI – последовательность для модификации пептида GPI-якорем; pA – сигнал полиаденилирования; 3'-НА – 3'-плечо гомологии. *б* – Схематичное изображение плазмид, несущих CTS на 5'- или 3'- конце донорной ДНК (перед 5'-НА или после 3'-НА соответственно); *в* – последовательности CTS и tCTS. 0ММ – в протоспейсере нет некомплементарных оснований; 4ММ, 6ММ и 8ММ – на 5'-конце протоспейсера находятся соответственно 4, 6 или 8 некомплементарных оснований. Сайт РАМ показан жирным шрифтом, протоспейсер подчеркнут, некомплементарные основания в его составе подчеркнуты пунктирной линией. Здесь: CTS – таргетная последовательность Cas9, соответствующая протоспейсеру вместе с РАМ; tCTS – таргетная последовательность Cas9, соответствующая протоспейсеру с некомплементарными основаниями.



**Рис. 2.** Влияние CTS-модификаций донорной ДНК на уровень нокина *MTC34* (а), нокаута *CXCR4* (б) и соотношения нокин/нокаут (в) в клетках СЕМ/R5. Клетки электропорировали плазмидами pDNA3.3-hCas9, pKS-gRNA-X4ex2 и одним из вариантов донорной плазмиды pKS-don\_MТ-C34, на 5 сутки оценивали уровень нокаута (КО) *CXCR4* и нокина (КИ) *MTC34* по экспрессии соответствующих белка *CXCR4* и пептида МТ-C34 на поверхности клеток с помощью проточной цитофлуориметрии. КИ/КО – соотношение нокина к нокауту; контроль (–) – донорная плазида pKS-don\_MТ-C34 без модификаций; (0, 4, 6, 8) ММ – число некомплементарных оснований в протоспейсере CTS. Приведены средние значения  $\pm$  стандартное отклонение (SD) и отдельные результаты трех независимых экспериментов. \* $p < 0.01$ , \*\* $p < 0.0001$ .

некомплементарных оснований (0ММ), даже при молярном соотношении РНП: субстрат, равном 2.43 (рис. 3з).

## ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

В этой работе мы оценили влияние модификаций плазмидной донорной ДНК путем добавления таргетных сайтов для Cas9 с протоспейсерами без некомплементарных оснований (CTS, 0ММ, “double-cut donor”) или с протоспейсерами, имеющими 4, 6 или 8 некомплементарных оснований на 5'-конце (tCTS), на уровень нокина конструкции, кодирующей пептидный ингибитор слияния ВИЧ-1 МТ-C34 (*MTC34*), в локус *CXCR4*. В результате обнаружено повышение уровня нокина при фланкировании донорной последовательности сайтами CTS, но не tCTS.

Выявленное нами повышение уровня нокина *MTC34* при введении CTS согласуется с результатами J. Zhang и др. [10], впервые систематически сравнивших немодифицированные и “double-cut” доноры. Однако это не совпадает с результатами D. Nguyen и др. [17], которые регистрировали повышение уровня нокина в 2–4 раза при введении в донорную ДНК tCTS-сайтов с 4 ММ. Авторы предположили, что повышение эффективности нокина за счет CTS-сайтов обусловлено одновременным расщеплением последовательности-мишени Cas9 в геноме и в донорной ДНК [10]. Однако механизм действия “double-cut” доноров на уровень нокина остается неизвестным. D. Nguyen с соавт. [16] объяснили эффект сайтов tCTS с 4ММ тем, что они не подвергаются расщеплению, но вызывают связывание донора с нуклеазой и облегчают его транспорт в ядро за счет NLS в Cas9. Авторы не исследовали расщепление донорной ДНК, несущей такие сайты, ни *in vitro*, ни *in vivo*. Мы установили, что доноры, модифицированные tCTS с 4 ММ, расщеплялись под влиянием Cas9 *in vitro* так же эффективно, как и доноры с полноразмерными CTS, но не вызывали повышения уровня нокина *MTC34*. Можно предположить, что эффект от введения мишеней Cas9 в донорную ДНК связан или с процессом линеаризации плазмидного донора, или с его взаимодействием с Cas9, или с обоими факторами.

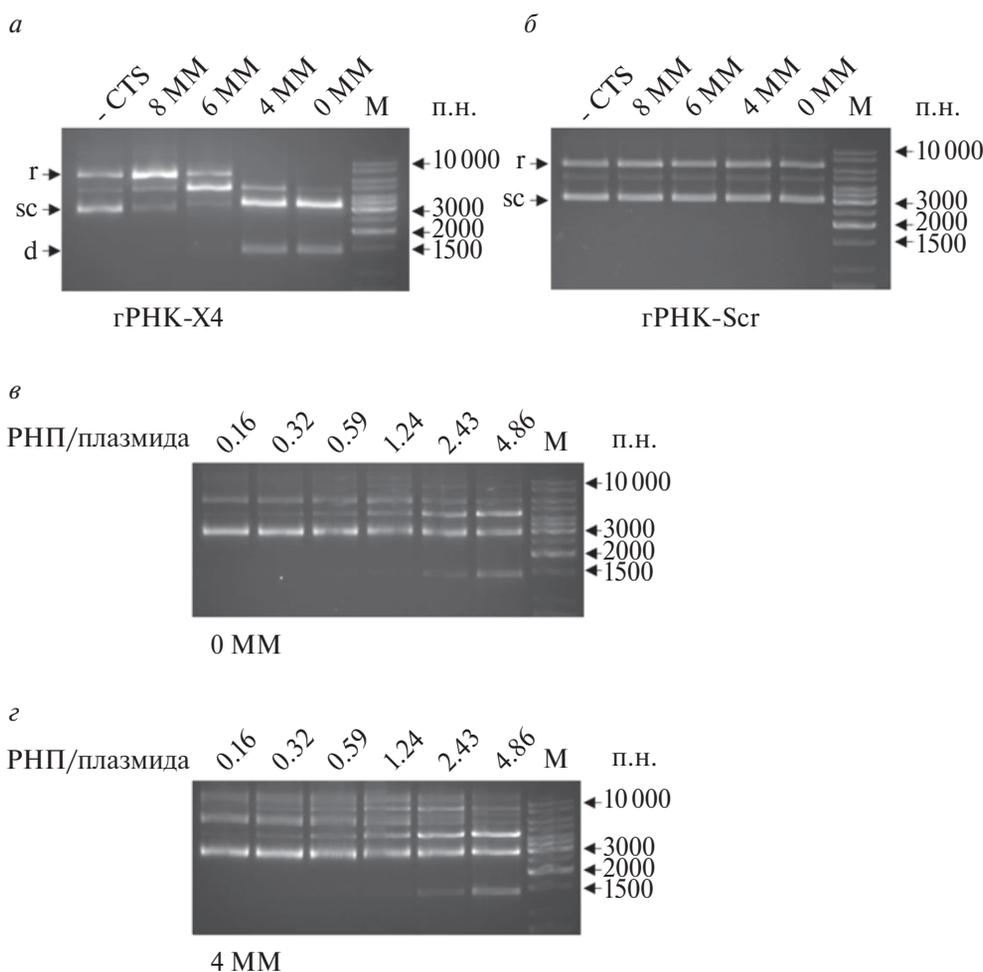
Нами показано, что повышение уровня нокина происходило при введении даже одного CTS-сайта по 5'- или 3'-концу донорной последовательности, тогда как J. Zhang и др. [10] исследовали только парные сайты. Можно допустить, что для повышения уровня нокина достаточно расщепления донорной ДНК в одном месте, с любой стороны от донорной последовательности, а наличие второго сайта расщепления не влияет на конечный результат. Недавно R. Jing и др. [25] показали, что уровень нокина повышался в два раза при введении

в плазмидный донор одного сайта CTS и в 10 раз при введении двух сайтов, в то время как третий сайт не приводил к дополнительному увеличению эффективности. Кроме того, при введении полно-размерных CTS повышение эффективности нокина регистрировали даже для линейных доноров. Все вышеперечисленное позволяет сделать вывод о том, что объяснить механизм действия CTS-сайтов исключительно линейризацией донорной ДНК невозможно, так как необходимо учитывать фактор взаимодействия донорной ДНК с Cas9.

Известно, что нуклеаза Cas9 — фермент одного цикла [24] и после расщепления субстрата *in vitro* в течение нескольких часов остается связанным с концами ДНК [26], поэтому при проведении *in vitro* расщепления ДНК по завершении

процесса реакцию смесь обрабатывают протеиназой К — для деградации Cas9 и высвобождения продуктов реакции из РНП-комплекса [27, 28]. Учитывая это, можно предположить, что *in vivo* Cas9 расщепляет таргетный сайт в донорной плазмиде, после чего остается связанной со свободным концом расщепленной ДНК. Если расщепление ДНК произошло в цитоплазме, то за счет NLS в белке Cas9 комплекс РНП и ДНК-матрицы может проникать в ядро. В ядре гРНК направляет Cas9 на целевой locus, что приводит к транспорту связанной ДНК-матрицы к месту репарации DSB. Кроме того, часть ДНК-матрицы может попасть в ядро и другим способом, где подвергнется расщеплению Cas9.

Нами показано, что доноры, модифицированные tCTS с 4 некомплементарными основаниями,



**Рис. 3.** Расщепление CTS-сайтов под действием Cas9 *in vitro*. Донорные плазмиды с CTS-сайтами, содержащими 0, 4, 6 или 8 некомплементарных оснований, а также донорную плазмиду без модификаций (-CTS) инкубировали *in vitro* с комплексом Cas9 и гРНК против *CXCR4* (гРНК-Х4) (а) или контрольной (гРНК-Scr) (б) и оценивали уровень расщепления плазмиды с помощью электрофореза в 1%-ном ТАЕ-агарозном геле. Условные обозначения: r (relaxed) — релаксированная форма плазмиды, sc (super-coiled) — суперскрученная форма, d — фрагмент ~1400 п.н. с донорной последовательностью, фланкированный CTS- или tCTS-сайтами. Донорные плазмиды с CTS-сайтами, содержащими 0 (в) или 4 (z) некомплементарных оснований, инкубировали *in vitro* с возрастающим количеством РНП и оценивали уровень расщепления плазмиды с помощью электрофореза в 1%-ном ТАЕ-агарозном геле. РНП/плазмиды — молярное соотношение РНП к плазмидной ДНК. М — маркеры длины ДНК: 10000, 8000, 6000, 5000, 4000, 3500, 3000, 2500, 2000, 1500, 1000, 750 п.н. (#SM1163; “Thermo Fisher Scientific”, США).

*in vitro* расщеплялись Cas9 так же, как и доноры с полноразмерными сайтами, однако это не влияло на уровень нокина. Если предположить, что для повышения уровня нокина необходимо именно расщепление донора, то введенные по его концам сайты tCTS с некомплементарными основаниями, скорее всего, не расщепляются Cas9 *in vivo* в клетках. Это предположение требует экспериментальной проверки. К сожалению, надежные способы оценки расщепления донорных плазмид в клетках под действием Cas9 пока не разработаны. Нам не удалось показать расщепление донорной ДНК *in vivo* с помощью количественной ПЦР (данные не приведены). Мы предполагаем, что при электропорации и транзиторной экспрессии в клетках образуется очень небольшое количество расщепленной донорной плазмиды, в результате уровень исходной плазмиды снижается настолько незначительно, что чувствительности количественной ПЦР для детекции этих изменений не хватает.

Стоит еще раз отметить, что ни в одной из опубликованных ранее работ, в которых использовали CTS или tCTS для модификации донорной ДНК, ее расщепление *in vivo* не исследовали. Кроме того, наличие или отсутствие эффекта от tCTS можно объяснить различиями в аффинности Cas9 к таким сайтам, которая варьирует в зависимости от нуклеотидной последовательности. В результате нуклеаза Cas9 может образовывать короткоживущий комплекс с tCTS донора, что не позволит ей эффективно переносить донорную ДНК в ядро. Эта гипотеза также требует экспериментальной проверки.

Отсутствие эффекта от tCTS в использованной нами модели можно связать с особенностью выбранного локуса — *CXCR4*. Так, В. Shy и др. [19] продемонстрировали, что влияние модификации донорной ДНК tCTS-сайтами на уровень нокина сильно варьировало для разных локусов: от отсутствия эффекта до повышения в 10 раз. Среди локусов с минимальным влиянием tCTS-сайтов на нокин был и *CXCR4*, хотя последовательность протоспейсера для этого локуса в работе не указана. Можно предположить, что в каждом индивидуальном случае, в зависимости от последовательности протоспейсера и некомплементарных оснований, *in vivo* расщепление донорной ДНК под действием Cas9 происходит либо эффективно, и в этом случае регистрируют повышение уровня нокина, либо неэффективно — тогда повышения нокина не детектируют.

Необходимо отметить, что в большинстве опубликованных работ, в которых изучали расщепление мишеней под действием Cas9 при формировании укороченного гетеродуплекса, использовали укороченные гРНК [18, 29]. В работе D. Nguyen и др. [17] для формирования связи между донором и Cas9 использовали укороченные протоспейсеры. Однако данных о расщеплении ДНК-мишени при

использовании укороченных протоспейсеров, а не укороченных гРНК, мы не нашли. В этом случае гетеродуплекс, формирующийся между ДНК-мишенью и гРНК, на самом деле нельзя считать укороченным в полной мере — укорочена лишь область комплементарности в протоспейсере, за которой следуют нуклеотиды, уникальные для каждого таргетного локуса генома и способные вносить различный вклад в стабильность гетеродуплекса. По этой причине мы предполагаем, что для понимания механизма действия tCTS-сайтов необходимо для каждой конкретной таргетной последовательности оценивать ее процессинг под действием Cas9 *in vitro* и *in vivo*.

Вероятно, при эффективном расщеплении *in vivo* таргетных последовательностей Cas9 сайты tCTS повышают уровень нокина, а в случае неэффективного расщепления *in vivo* эффект от tCTS отсутствует. В таком случае наблюдаемые J. Zhang и др. [10] и D. Nguyen и др. [17] эффекты от соответственно CTS- и tCTS-сайтов имеют одинаковую природу. Подтверждение этого предположения потребует анализа нокина в ряд дополнительных локусов с одновременной оценкой расщепления донорной плазмиды *in vivo* и *in vitro*, а также измерения кинетических параметров связывания Cas9 с таргетными последовательностями.

Таким образом, на клинически значимой модели нокина в Т-клетках нам удалось повысить уровень нокина *MTC34* в локус *CXCR4* путем модификации донорной плазмиды мишенями Cas9, фланкирующими донорные последовательности. Выяснение молекулярных механизмов этого эффекта позволит усовершенствовать технологию нокина за счет модификации донорной ДНК.

Работа выполнена при финансовой поддержке Российского научного фонда (грант № 22-25-00310).

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. Doudna J.A. (2020) The promise and challenge of therapeutic genome editing. *Nature*. **578**, 229–236. <https://doi.org/10.1038/s41586-020-1978-5>
2. Jiang F., Doudna J.A. (2017) CRISPR-Cas9 structures and mechanisms. *Annu. Rev. Biophys.* **46**, 505–529. <https://doi.org/10.1146/annurev-biophys-062215-010822>
3. Antoniani C., Meneghini V., Lattanzi A., Felix T., Romano O., Magrin E., Weber L., Pavani G., El Hoss S., Kurita R., Nakamura Y., Cradick T.J., Lundberg A.S., Porteus M., Amendola M., El Nemer W., Cavazzana M., Mavilio F., Miccio A. (2018) Induction of fetal hemoglobin

- synthesis by CRISPR/Cas9-mediated editing of the human  $\beta$ -globin locus. *Blood*. **131**, 1960–1973.  
<https://doi.org/10.1182/blood-2017-10-811505>
4. Pavlovic K., Tristán-Manzano M., Maldonado-Pérez N., Cortijo-Gutierrez M., Sánchez-Hernández S., Justicia-Lirio P., Carmona M.D., Herrera C., Martin F., Benabdelah K. (2020) Using gene editing approaches to fine-tune the immune system. *Front. Immunol.* **11**, 570672.  
<https://doi.org/10.3389/fimmu.2020.570672>
  5. Kotagama O.W., Jayasinghe C.D., Abeyasinghe T. (2019) Era of genomic medicine: a narrative review on CRISPR technology as a potential therapeutic tool for human diseases. *Biomed. Res. Int.* **2019**, 1369682.  
<https://doi.org/10.1155/2019/1369682>
  6. Sun W., Liu H., Yin W., Qiao J., Zhao X., Liu Y. (2022) Strategies for enhancing the homology-directed repair efficiency of CRISPR-Cas systems. *CRISPR J.* **5**, 7–18.  
<https://doi.org/10.1089/crispr.2021.0039>
  7. Shams F., Bayat H., Mohammadian O., Mahboudi S., Vahidnezhad H., Soosanabadi M., Rahimpour A. (2022) Advance trends in targeting homology-directed repair for accurate gene editing: an inclusive review of small molecules and modified CRISPR-Cas9 systems. *Bioimpacts.* **12**, 371–391.  
<https://doi.org/10.34172/bi.2022.23871>
  8. Smirnikhina S.A., Zaynitdinova M.I., Sergeeva V.A., Lavrov A.V. (2022) Improving homology-directed repair in genome editing experiments by influencing the cell cycle. *Int. J. Mol. Sci.* **23**, 5992.  
<https://doi.org/10.3390/ijms23115992>
  9. Richardson C.D., Ray G.J., DeWitt M.A., Curie G.L., Corn J.E. (2016) Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. *Nat. Biotechnol.* **34**, 339–344.  
<https://doi.org/10.1038/nbt.3481>
  10. Zhang J.P., Li X.L., Li G.H., Chen W., Arakaki C., Botimer G.D., Baylink D., Zhang L., Wen W., Fu Y.W., Xu J., Chun N., Yuan W., Cheng T., Zhang X.B. (2017) Efficient precise knockin with a double cut HDR donor after CRISPR/Cas9-mediated double-stranded DNA cleavage. *Genome Biol.* **18**, 35.  
<https://doi.org/10.1186/S13059-017-1164-8>
  11. Ghanta K.S., Chen Z., Mir A., Dokshin G.A., Krishnamurthy P.M., Yoon Y., Gallant J., Xu P., Zhang X.O., Ozturk A.R., Shin M., Idrizi F., Liu P., Gneid H., Edraki A., Lawson N.D., Rivera-Pérez J.A., Sontheimer E.J., Watts J.K., Mello C.C. (2021) 5'-Modifications improve potency and efficacy of DNA donors for precision genome editing. *Elife.* **10**, e72216.  
<https://doi.org/10.7554/eLife.72216>
  12. Haraguchi T., Koujin T., Shindo T., Bilir Ş., Osakada H., Nishimura K., Hirano Y., Asakawa H., Mori C., Kobayashi S., Okada Y., Chikashige Y., Fukagawa T., Shibata S., Hiraoka Y. (2022) Transfected plasmid DNA is incorporated into the nucleus via nuclear envelope reformation at telophase. *Commun. Biol.* **5**, 78.  
<https://doi.org/10.1038/s42003-022-03021-8>
  13. Carlson-Stevermer J., Abdeen A.A., Kohlenberg L., Goeland M., Molugu K., Lou M., Saha K. (2017) Assembly of CRISPR ribonucleoproteins with biotinylated oligonucleotides via an RNA aptamer for precise gene editing. *Nat. Commun.* **8**, 1711.  
<https://doi.org/10.1038/s41467-017-01875-9>
  14. Ma M., Zhuang F., Hu X., Wang B., Wen X.Z., Ji J.F., Xi J.J. (2017) Efficient generation of mice carrying homozygous double-floxed alleles using the Cas9-Avidin/Biotin-donor DNA system. *Cell Res.* **27**, 578–581.  
<https://doi.org/10.1038/cr.2017.29>
  15. Savic N., Ringnalda F.C., Lindsay H., Berk C., Bargsten K., Li Y., Neri D., Robinson M.D., Ciaudo C., Hall J., Jinek M., Schwank G. (2018) Covalent linkage of the DNA repair template to the CRISPR-Cas9 nuclease enhances homology-directed repair. *Elife.* **7**, e33761.  
<https://doi.org/10.7554/eLife.33761>
  16. Aird E.J., Lovendahl K.N., St. Martin A., Harris R.S., Gordon W.R. (2018) Increasing Cas9-mediated homology-directed repair efficiency through covalent tethering of DNA repair template. *Commun. Biol.* **1**, 54.  
<https://doi.org/10.1038/s42003-018-0054-2>
  17. Nguyen D.N., Roth T.L., Li P.J., Chen P.A., Apathy R., Mamedov M.R., Vo L.T., Tobin V.R., Goodman D., Shifrut E., Bluestone J.A., Puck J.M., Szoka F.C., Marson A. (2020) Polymer-stabilized Cas9 nanoparticles and modified repair templates increase genome editing efficiency. *Nat. Biotechnol.* **38**, 44–49.  
<https://doi.org/10.1038/s41587-019-0325-6>
  18. Zhang J.P., Li X.L., Neises A., Chen W., Hu L.P., Ji G.Z., Yu J.Y., Xu J., Yuan W.P., Cheng T., Zhang X.B. (2016) Different effects of sgRNA length on CRISPR-mediated gene knockout efficiency. *Sci. Rep.* **6**, 28566.  
<https://doi.org/10.1038/srep28566>
  19. Shy B.R., Vykunta V.S., Ha A., Talbot A., Roth T.L., Nguyen D.N., Pfeifer W.G., Chen Y.Y., Blaeschke F., Shifrut E., Vedova S., Mamedov M.R., Chung J.J., Li H., Yu R., Wu D., Wolf J., Martin T.G., Castro C.E., Ye L., Esensten J.H., Eyquem J., Marson A. (2023) High-yield genome engineering in primary cells using a hybrid ssDNA repair template and small-molecule cocktails. *Nat. Biotechnol.* **41**, 521–531.  
<https://doi.org/10.1038/s41587-022-01418-8>
  20. Kath J., Du W., Pruene A., Braun T., Thommandru B., Turk R., Sturgeon M.L., Kurgan G.L., Amini L., Stein M., Zittel T., Martini S., Ostendorf L., Wilhelm A., Akyüz L., Rehm A., Höpken U.E., Pruß A., Künkele A., Jacobi A.M., Volk H.D., Schmueck-Henneresse M., Stripecke R., Reinke P., Wagner D.L. (2022) Pharmacological interventions enhance virus-free generation of TRAC-replaced CAR T cells. *Mol. Ther. Methods Clin. Dev.* **25**, 311–330.  
<https://doi.org/10.1016/j.omtm.2022.03.018>
  21. Oh S.A., Senger K., Madireddi S., Akhmetzyanova I., Ishizuka I.E., Tarighat S., Lo J.H., Shaw D., Haley B., Rutz S. (2022) High-efficiency nonviral CRISPR/Cas9-mediated gene editing of human T cells using plasmid donor DNA. *J. Exp. Med.* **219**, e20211530.  
<https://doi.org/10.1084/jem.20211530>
  22. Lin-Shiao E., Pfeifer W.G., Shy B.R., Saffari Doost M., Chen E., Vykunta V.S., Hamilton J.R., Stahl E.C., Lopez D.M., Sandoval Espinoza C.R., Deyanov A.E., Lew R.J., Poirer M.G., Marson A., Castro C.E., Doudna J.A. (2022) CRISPR-Cas9-mediated nuclear transport and genomic integration of nanostructured genes in human primary cells. *Nucleic Acids Res.* **50**, 1256–1268.  
<https://doi.org/10.1093/nar/gkac049>
  23. Maslennikova A., Golubova N., Kalinichenko S., Komkov D., Shepelev M., Golubev D., Siniavin A., Vzorov A., Filatov A., Mazurov D. (2022) Engineering T-cell resistance

- to HIV-1 infection via knock-in of peptides from the heptad repeat 2 domain of gp41. *mBio*. **13**, e0358921. <https://doi.org/10.1128/mbio.03589-21>
24. Sternberg S.H., Redding S., Jinek M., Greene E.C., Doudna J.A. (2014) DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature*. **507**, 62–67. <https://doi.org/10.1038/nature13011>
  25. Jing R., Jiao P., Chen J., Meng X., Wu X., Duan Y., Shang K., Qian L., Huang Y., Liu J., Huang T., Jin J., Chen W., Zeng X., Yin W., Gao X., Zhou C., Sadelain M., Sun J. (2021) Cas9-cleavage sequences in size-reduced plasmids enhance nonviral genome targeting of CARs in primary human T cells. *Small Methods*. **5**, e2100071. <https://doi.org/10.1002/smt.202100071>
  26. Aldag P., Welzel F., Jakob L., Schmidbauer A., Rutkauskas M., Fettes F., Grohmann D., Seidel R. (2021). Probing the stability of the SpCas9–DNA complex after cleavage. *Nucleic Acids Res.* **49**, 12411–12421. <https://doi.org/10.1093/nar/gkab1072>
  27. Zou R., Liu Y., Ha T. (2021) *In vitro* cleavage and electrophoretic mobility shift assays for very fast CRISPR. *Bio Protoc.* **11**, e4138. <https://doi.org/10.21769/BioProtoc.4138>
  28. Liu Y., Zou R.S., He S., Nihongaki Y., Li X., Razavi S., Wu B., Ha T. (2020) Very fast CRISPR on demand. *Science*. **368**, 1265–1269. <https://doi.org/10.1126/science.aay8204>
  29. Fu Y., Sander J.D., Reyon D., Cascio V.M., Joung J.K. (2014) Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat. Biotechnol.* **32**, 279. <https://doi.org/10.1038/NBT.2808>

## Increasing the Level of Knock-In of the MT-C34-Encoding Construct into the *CXCR4* Locus by Modifying Donor DNA with Cas9 Target Sites

M. V. Shepelev<sup>1</sup>, D. S. Komkov<sup>1, 2</sup>, D. S. Golubev<sup>1</sup>,  
S. E. Borovikova<sup>3</sup>, D. V. Mazurov<sup>1, 4</sup>, N. A. Kruglova<sup>1, \*</sup>

<sup>1</sup>Center for Precision Genome Editing and Genetic Technologies for Biomedicine, Institute of Gene Biology, Russian Academy of Sciences, Moscow, 119334 Russia

<sup>2</sup>Department of Physiology and Cell Biology, Faculty of Health Sciences, Ben-Gurion University of the Negev, Beer-Sheva, 8410501 Israel

<sup>3</sup>Institute of Gene Biology, Russian Academy of Sciences, Moscow, 119334 Russia

<sup>4</sup>Division of Infectious Diseases and International Medicine, Department of Medicine, University of Minnesota, Minneapolis, 55455 USA

\*e-mail: natalya.a.kruglova@yandex.ru

For successful application of genome editing technology using CRISPR/Cas9 system in clinical practice, it is necessary to achieve high efficiency of knock-in, the insertion of a genetic construct into a given locus in the genome of a target cell. One approach to increasing knock-in efficiency involves modifying donor DNA with the same targets for Cas9 (Cas9 targeting sequence, CTS) that are used for induction of double-strand breaks in the cell genome (the “double-cut donor” method). Another approach is based on introducing truncated targets for Cas9 (truncated CTS, tCTS), including a PAM site and 16 nucleotides proximal to it, into the donor DNA. Presumably, tCTS sites do not induce cleavage of the donor plasmid, but can support its transport into the nucleus by Cas9. However, the exact mechanisms for the increase in knock-in levels with both types of donor DNA modifications are unknown. Here, we evaluated the effect of these modifications on the knock-in efficiency of the *MTC34* genetic construct encoding the HIV-1 fusion inhibitor, MT-C34 peptide, into the *CXCR4* locus of the CEM/R5 T cell line. When full-length CTS sites were introduced into the donor plasmid DNA, the knock-in level increased twofold, regardless of the number of CTSs or their position relative to the donor sequence. Modifications of donor plasmids with tCTS sites did not affect knock-in levels. It was found that *in vitro* both types of sites were efficiently cleaved by Cas9. In order to study the mechanism of action of these modifications in detail, it is necessary to evaluate their cleavage *in vitro* and *in vivo*.

**Keywords:** CRISPR/Cas9, genome editing, T lymphocytes, knock-in, donor DNA, Cas9 targeting site, DNA nuclear transport

## ЧУВСТВИТЕЛЬНОСТЬ ПЕРВИЧНЫХ КЛЕТОЧНЫХ ЛИНИЙ ГЛИОБЛАСТОМЫ ЧЕЛОВЕКА К ВАКЦИННОМУ ШТАММУ ВИРУСА ПАРОТИТА

© 2024 г. Е. Ю. Николаева<sup>а</sup>, Ю. Р. Желаева<sup>а</sup>, О. Ю. Сусова<sup>б</sup>, А. А. Митрофанов<sup>б</sup>, В. О. Варачев<sup>с</sup>, Т. В. Наседкина<sup>с</sup>, В. В. Зверев<sup>а</sup>, О. А. Свитич<sup>а</sup>, Ю. И. Аммура<sup>а, \*</sup>

<sup>а</sup>Научно-исследовательский институт вакцин и сывороток им. И.И. Мечникова  
Министерства науки и высшего образования Российской Федерации, Москва, 105064 Россия

<sup>б</sup>Национальный медицинский исследовательский центр онкологии им. Н.Н. Блохина  
Министерства здравоохранения Российской Федерации, Москва, 115478 Россия

<sup>с</sup>Институт молекулярной биологии им. В.А. Энгельгардта Российской академии наук, Москва, 119991 Россия

\*e-mail: yulia.ammour@yahoo.fr

Поступила в редакцию 15.11.2023 г.

После доработки 11.12.2023 г.

Принята к печати 11.12.2023 г.

Чувствительность клеток глиобластомы человека к вирус-опосредованному онколизу изучена на пяти линиях клеток, полученных от пациентов. Первичные клеточные линии глиобластомы (Gbl13n, Gbl16n, Gbl17n, Gbl25n и Gbl27n) инфицировали 10-кратными серийными разведениями вируса паротита (штамм Ленинград-3), репродукцию вируса и цитотоксичность отслеживали в течение 96–120 ч. Иммутизированные неопухолевые клетки NKE человека использовали в качестве контроля для определения специфичности вируса. Четыре из пяти клеточных линий глиобластомы были чувствительны к заражению вирусом паротита, тогда как в неопухолевой линии клеток репродукцию вируса не наблюдали. При этом уровень активности проапоптотической каспазы-3 повышался во всех инфицированных клетках через 48 ч после заражения. Кинетика накопления вирусной РНК в исследованных линиях клеток глиобластомы была сопоставимой со скоростью гибели клеток. Полученные данные свидетельствуют о том, что клеточные линии глиобластомы перmissive для вируса паротита. Продукция интерферона типа I в ответ на заражение вирусом паротита клеточных линий глиобластомы различалась. Кроме того, показано, что заражение вирусом паротита способно вызывать иммуногенную гибель клеток глиобластомы.

**Ключевые слова:** глиобластома, онколитические вирусы, вирус паротита

DOI: 10.31857/S0026898424040068, EDN: INBJDQ

### ВВЕДЕНИЕ

Глиобластома 4-й степени (Glioblastoma, Grade 4) – наиболее распространенная и наиболее летальная первичная злокачественная опухоль головного мозга у взрослых. Медиана общей выживаемости при глиобластоме составляет 15 месяцев, а 5-летняя выживаемость обычно не превышает 5–7% после постановки первоначального диагноза [1]. Стандартом лечения уже более 10 лет является проведение полной хирургической резекции с последующей химиотерапией

темозоломидом и лучевой терапией, что позволяет увеличить среднюю выживаемость до 14.6 месяцев [2, 3]. Глиобластома характеризуется быстрым диффузным, инфильтративным ростом и высоким уровнем клеточной гетерогенности, что затрудняет лечение и ухудшает прогноз выживаемости [4]. Перспективным подходом к лечению глиобластомы является применение онколитических вирусов (ОВ), способных специфически инфицировать и лизировать опухолевые клетки, минимально повреждая при этом нормальные клетки. Кроме того, ОВ могут преодолевать иммуносупрессивное

Сокращения: ОВ – онколитические вирусы; МОИ – множественность инфекции; PBS – фосфатно-солевой раствор; FBS – эмбриональная сыворотка крупного рогатого скота; УФ – ультрафиолетовое излучение; МТТ – 3-(4,5-диметилтиазол-2-ил)-2,5-дифенилтетразолийбромид; ПЦР-РВ – ПЦР в режиме реального времени; ТЦД – тканевая цитопатическая доза; ЦПД – цитопатическое

действие; БОЕ – бляшкообразующие единицы; IFN – интерферон; ED50 – эффективная доза; РВМС (peripheral blood mononuclear cell) – мононуклеарные клетки периферической крови, NK (natural killer) – естественные киллерные клетки, E:T (Effector to target ratio) – соотношение числа эффекторных клеток (NK) к числу клеток глиобластомы, ns (not significant) – незначимое различие.

действие микроокружения опухоли, стимулируя врожденный иммунитет, а возможность применения в комбинации с другими методами лечения существенно повышает их противоопухолевую эффективность [5]. На данный момент ОВ используются в трех клинических исследованиях терапии глиобластомы, восемь клинических испытаний уже завершены (<https://clinicaltrials.gov/search?cond=Glioblastoma&term=Glioblastoma%20Multiforme;%20&distance=50&aggFilters=status:act&viewType=Table&intr=oncolytic>).

Ранее нами был показан онколитический потенциал вакцинного штамма Ленинград-16 вируса кори против первичных клеток глиобластомы человека и охарактеризованы генетические закономерности, определяющие чувствительность первичных клеток глиобластомы человека к онколитической терапии вирусом кори [6]. Так, проведение РНК-Seq выявило около 1200 дифференциально экспрессированных генов (FDR < 0.05) как минимум с двукратным изменением уровня экспрессии между инфицированными и неинфицированными клетками, с наиболее значительным повышением уровня экспрессии генов интерфероновой ответа, апоптозного и цитокинового сигнальных путей. В данной работе рассмотрен онколитический потенциал вируса паротита (Mumps virus, MV) в отношении первичных клеток глиобластомы и оценено влияние интерфероновой ответа на репликацию MV в культурах клеток глиобластомы. Впервые онколитические свойства

MV были показаны в клинических исследованиях, проведенных в Японии в 1970–1980-х гг., в которых оценивали действие клинического изолята MV дикого типа в отношении различных видов рака [7–9], включая опухоли центральной нервной системы [10]. Однако терапевтический потенциал живых вакцинных штаммов MV в отношении клеток солидных опухолей человека (меланомы [11], фибросаркомы и рака шейки матки [12], рака яичников [13], трижды негативного рака молочной железы [14]) показан относительно недавно.

Таким образом, цель работы заключалась в оценке онколитических свойств отечественного вакцинного штамма MV Ленинград-3 (L-3), которым заражали первичные клетки глиобластомы человека. Исследование проведено на пяти клеточных линиях и неопухолевой линии клеток человека, используемой в качестве контроля. Показано, что MV реплицируется в большинстве клеточных линий глиобластомы, вызывая их гибель. Таким образом, MV, наряду с другими парамиксовирусами, можно рассматривать в качестве потенциального онколитического агента в отношении клеток глиобластомы.

## ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

**Вирусы и клетки.** В работе использован вакцинный штамм MV L-3 с титром 6.4–6.5 lgТЦД<sub>50</sub>/мл или 6.9 lgБОЕ/мл (коллекция НИИ вакцин и сывороток (НИИВС) им. И. И. Мечникова). MV

**Таблица 1.** Использованные культуры клеток глиобластомы человека

Культура клеток	Пол и возраст пациента	Диагноз (по ВОЗ)	Мутированный ген	Мутация*	Пассажный уровень
Gbl 13n	М, 45	Глиобластома G4, рецидив	<i>PTEN</i> <i>MRE11</i> <i>TP53</i>	c.T821C c.579_582del c.C460T	10
Gbl 16n	М, 47	Глиобластома G4	<i>TP53</i> <i>PTEN</i> <i>TAF1L</i> <i>LCK</i>	c.A419G c.A588T c.C2746T c.G1332A	9
Gbl 17n	М, 68	Глиобластома G4	<i>PTEN</i> <i>ERBB3</i> <i>PIK3CG</i> <i>TOP1</i>	c.1524_1525ins c.C2150T c.A766T c.533_538del	8
Gbl 25n	Ж, 70	Глиобластома G4	<i>PTEN</i> <i>TP53</i>	c.635–1G>A c.C263T	9
Gbl 27n	М, 56	Глиобластома G4	<i>PTEN</i> <i>TP53</i>	c.C195A c.C263T	7

\*Секвенирование проводили на приборе NextSeq2000 методом парноконцевых прочтений на сформированной панели зондов, включающей последовательности 812 генов, мутации в которых ассоциированы с развитием опухоли.

Примечание. Панели для таргетного секвенирования формировали на основе панелей Foundation One (“Roche”, Швейцария) и TruSight Cancer Sequencing Panel (“Illumina”, США), базы данных cBioPortal (<https://www.cbioportal.org>) и данных о спектре соматических мутаций в опухолях мозга [15].

культивировали и титровали по ЦПД и реакции бляшкообразования на клеточной культуре Vero (CCL-81, ATCC, США).

Линии клеток Gbl13n, Gbl16n, Gbl17n, Gbl25n и Gbl27n на ранних пассажах (7–10 пассажи) получены в Российском онкологическом научном центре (РОНЦ) им. Н.Н. Блохина (Москва) из операционного материала или биопсии пациентов с подтвержденным клиническим диагнозом глиобластома. Молекулярно-генетический профиль клеточных линий изучен с помощью NGS-секвенирования [15]. Обнаруженные мутации приведены в табл. 1.

Иммортализованные эпителиальные клетки NKE человека (нормальные эпителиальные клетки почки) получены из коллекции РОНЦ им. Н.Н. Блохина. Перевиваемая культура клеток Vero получена из коллекции НИИВС им. И.И. Мечникова.

Мононуклеарные клетки периферической крови (PBMC) здоровых доноров выделяли с использованием Ficoll (F5415, “Sigma”, США) согласно инструкции.

Все клеточные линии культивировали во флаконах T25 (430372, “Corning”, Нидерланды) в питательной среде DMEM/F12 (31330095, “Thermo Fisher Scientific” (“TFS”), США), содержащей 10% эмбриональной сыворотки крупного рогатого скота (FBS; A3160802, “TFS”), 1 мМ глутамин (35050, “TFS”) и 100 мкг/мл гентамицина (15710064, “TFS”). Клетки инкубировали при температуре 37°C в атмосфере 5% CO<sub>2</sub>. При достижении монослы клеток проводили пересев культуры. Для этого сливали супернатанты, обрабатывали 10%-ным фосфатно-солевым раствором (PBS) для удаления остатков сыворотки и мертвых клеток. Клетки снимали раствором TrypLE Express (12605, “TFS”). Клетки в клеточной суспензии для определения их жизнеспособности окрашивали 0.2%-ным раствором трипанового синего (15250061, “TFS”). Подсчет клеток проводили в камере Горяева.

**Заражение клеток.** Клетки, достигшие 80–90% монослы во флаконах T25, промывали 2 мл 10%-ного раствора PBS. Исходя из числа клеток во флаконе и исходного титра MV, делали серию разведений вируса в соответствующей каждой культуре клеток питательной среде (без FBS) и вносили в монослой клеток. Каждую клеточную линию инкубировали с 2 мл вирусного материала с различной множественностью заражения (MOI) при 37°C в атмосфере 5% CO<sub>2</sub> в течение 90 мин. В качестве контроля использовали те же культуры, инкубированные с предварительно УФ-обработанным MV. После адсорбции вируса монослой клеток промывали раствором Хэнкса и культивировали в 7 мл (T25) питательной среды DMEM/F12 с 2% FBS в течение 120 ч при 37°C в атмосфере 5% CO<sub>2</sub>.

Для определения кинетики внутриклеточного накопления вируса клетки за день до заражения высаживали в 6-луночный планшет (353046, “Corning”) из расчета  $2 \times 10^6$  клеток на лунку и по вышеприведенной схеме заражали MV с MOI 0.1 или 1.0, а также УФ-инактивированным вирусом (MOI 1.0). Клеточный монослой лизировали через 0, 24, 48, 72 и 96 ч после заражения, используя 300 мкл лизирующего буфера RLT RNAeasy kit (74104, “Qiagen”, Германия) на лунку. Эксперимент проводили в трех повторах. Лизаты клеток центрифугировали при 1000 об./мин в течение 5 мин. Супернатант хранили при –70°C до использования.

**Оценка жизнеспособности клеток. МТТ-тест.** Выживаемость клеток оценивали в МТТ-тесте согласно [11]. В лунки 96-луночных планшетов вносили по 150 мкл ростовой среды, содержащей  $1.5 \times 10^4$  клеток. Клетки заражали вирусом, содержащим супернатантом, как описано выше, с 10-кратными разведениями вируса и инкубировали при 37°C в атмосфере 5% CO<sub>2</sub> в течение 3–120 ч. Через каждые 24 ч в лунки добавляли 25 мкл раствора 3-(4,5-диметилтиазол-2-ил)-2,5-дифенилтетразолийбромид (МТТ, M-2128, “Sigma”), разведенного до концентрации 5 мг/мл в аналогичной ростовой среде, и инкубировали в течение 3 ч при 37°C в атмосфере 5% CO<sub>2</sub>. Среда удаляла и добавляли по 60 мкл DMSO в лунку для растворения образовавшихся кристаллов формазана; для гомогенного распределения красителя использовали шейкер (10 мин, 300 об./мин). Оптическую плотность определяли на микропланшетном детекторе Varioskan Flash (“TFS”) при длине волны 570 нм (при эталонной длине волны 650 нм). Все эксперименты проводили в трех независимых повторах, значения оптической плотности зараженных культур рассчитывали в процентах от значений для аналогичной незараженной культуры непосредственно перед заражением (0 ч).

**Оценка активности каспазы 3.** Параллельно с МТТ-тестом колориметрическим методом оценивали активность каспазы 3, измеряя окрашенный субстрат согласно инструкции производителя (E-CK-A311, “Elabscience”, Китай) через 48 ч после инокуляции вируса в 6-луночный планшет. Схема заражения идентична схеме, используемой для определения кинетики накопления вируса. Колориметрический сигнал от каждого образца измеряли с помощью микропланшетного ридера Varioskan Flash.

**Выделение тотальной РНК.** Тотальную РНК выделяли из лизатов  $(1–6) \times 10^6$  клеток с помощью коммерческого набора RNeasy Mini Kit (74104, “Qiagen”) согласно инструкции производителя. Очищенную РНК элюировали с мембраны колонок RNeasy Mini Spin (2 раза по 50 мкл воды, не содержащей РНКаз), концентрацию РНК определяли на приборе NanoDrop 8000 (“TFS”).

**Обратная транскрипция и количественная ПЦР с детекцией в режиме реального времени (ОТ-ПЦР-РВ).** Репродукцию MV оценивали по уровню вирусной РНК, который определяли методом ОТ-ПЦР-РВ [16]. В каждой постановке реакции параллельно определяли пороговые циклы для разведений стандартного образца MV с известным титром вируса, по результатам которых строили калибровочную кривую. На калибровочной кривой откладывали значения пороговых циклов образцов и рассчитывали значения в  $\lg(\text{ТЦД}_{50}/\text{мл})$ , которые условно принимали за титры [16]. Полученные значения выборочно проверяли методом предельных разведений по цитопатическому действию на чувствительной культуре клеток (Vero), расхождения между значениями, полученными двумя методами, не превышали  $0.5 \lg(\text{ТЦД}_{50}/\text{мл})$ . Эксперименты проводили в трех независимых повторах.

**Иммуноферментный анализ.** Уровень IFN- $\beta$  измеряли в супернатантах клеточных культур с помощью набора ELISA Kit for Interferon Beta (IFN $\beta$ ) (SEA222Hu, "Cloud-Clone Corp.", Китай) согласно инструкции производителя.

**Проточная цитометрия.** Клетки РВМС оценивали по представленности лейкоцитарных маркеров (BD Multitest IMK Kit, 340503, "BD Biosciences", США), а также маркеров CD3 (E-AB-F1001C FITC-меченое антитело к CD3 [ОКТ3], "Elabscience"), CD56 (E-AB-F1239D PE-меченое антитело к CD56 [5.1H11], "Elabscience"), CD69 (E-AB-F1138E APC-меченое антитело к CD69 [FN50], "Elabscience") в присутствии изотипических контролей (E-AB-F09802C FITC-меченое антитело для IgG2a, каппа [C1.18.4]; E-AB-F09792D PE-меченое антитело для IgG1, каппа [MOPC-21]; E-AB-F09792E APC-меченое антитело для IgG1, каппа [MOPC-21], "Elabscience") согласно инструкциям. Анализ проводили на приборе BD FACSCanto II ("BD Biosciences").

## РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

*Клетки глиобластомы чувствительны к онколизу, индуцированному вирусом паротита штамм L-3*

Чувствительность клеток глиобластомы человека к MV-опосредованному онколизису оценивали на пяти клеточных линиях: Gbl13n, Gbl16n, Gbl17n, Gbl25n и Gbl27n, полученных от пациентов (табл. 1). При оценке специфичности действия MV на клетки глиобластомы человека в качестве контроля использовали линию иммортализованных клеток эпителия почки – NKE, темп пролиферации которых сопоставим со временем удвоения большинства линий клеток глиобластомы (за исключением Gbl17n), использованных в данном исследовании.

Клеточные линии глиобластомы заражали 10-кратными серийными разведениями материала,

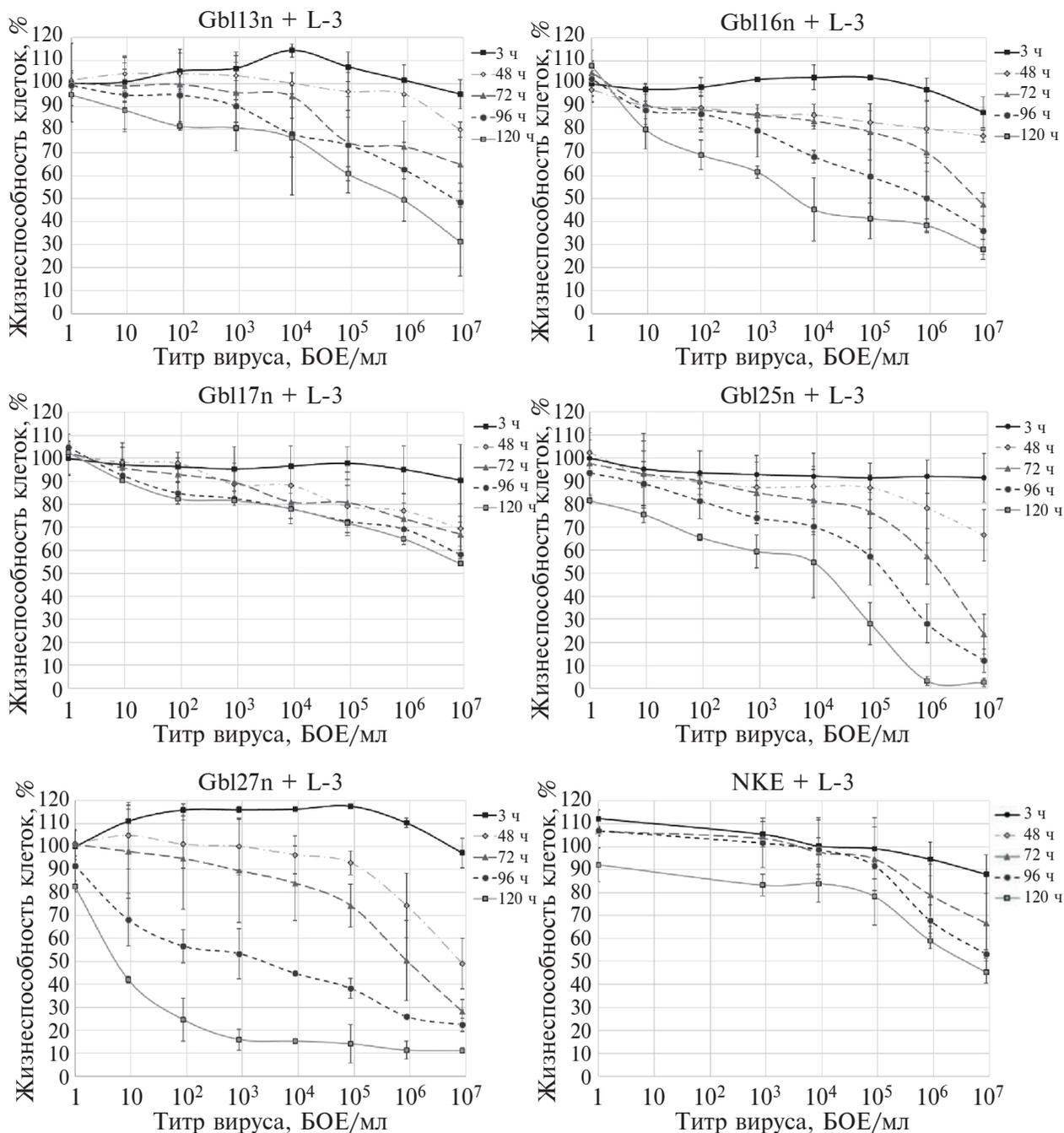
содержащего штамм L-3 MV. Титр вируса варьировал от 0.5 до 6.5  $\lg\text{ТЦД}_{50}/\text{мл}$  или от 0.9 до 6.9  $\lg\text{БОЕ}/\text{мл}$ . Контрольные клетки инкубировали с УФ-инактивированным вирусом (МОИ 1.0) в течение такого же промежутка времени.

Уже через 48 ч после заражения наблюдали цитопатическое действие (ЦПД) MV L-3 на клетки, которое выражалось в грануляции цитоплазмы, округлении и постепенном отслаивании клеток всех исследованных культур глиобластомы, кроме Gbl17n и контрольной культуры клеток NKE. ЦПД коррелировало с разведением MV, становясь более очевидным с течением времени.

Выживаемость клеток оценивали, измеряя митохондриальную активность с помощью МТТ-теста через 3, 36, 48, 72, 96 и 120 ч после заражения, и выражали в процентах на указанной временной точке относительно незараженной культуры клеток непосредственно перед заражением (рис. 1).

Первичные клетки глиобластомы проявляли разную чувствительность к MV-опосредованному онколизу. На основании данных о выживаемости клеток после заражения вирусом все культуры глиобластомы условно разделили на резистентные, умеренно-чувствительные и восприимчивые к заражению. Так, культура Gbl17n принята за условно резистентную, так как при заражении даже неразведенным MV (с наибольшим титром) выживаемость клеток через 120 ч после заражения составила 54%. Культура Gbl13n принята за умеренно-чувствительную, поскольку только при воздействии неразведенного вируса (с наибольшим титром) выживаемость клеток была менее 50%: 48.4 и 31.3% через 96 и 120 ч после заражения соответственно. При заражении 10-кратным разведением вируса выживаемость клеток глиобластомы спустя 120 ч после заражения составила 49.3%. Три остальные культуры, Gbl16n, Gbl25n и Gbl27n, были восприимчивыми к MV-опосредованному онколизу, при этом наиболее чувствительной оказалась Gbl27n. Используя полученные кривые, рассчитали индекс ED50 разведения MV для каждой клеточной линии с высокой статистической значимостью ( $R^2 > 0.98$ ). ED50 разведения вводимого MV составила: **2.54**, **4.07** и **0.68**  $\lg\text{БОЕ}/\text{мл}$  для клеточных культур Gbl16n, Gbl25n и Gbl27n соответственно; **5.91**  $\lg\text{БОЕ}/\text{мл}$  – для культуры клеток Gbl13n; и **7.20**  $\lg\text{БОЕ}/\text{мл}$  – для культуры клеток Gbl17n через 120 ч после заражения.

Выживаемость клеток контрольной неопухольевой культуры NKE составила 53.2 и 45.3% через 96 и 120 ч после заражения неразведенным вирусом с исходным титром соответственно. Рассчитанная ED50 ( $R^2 = 0.98$ ) составила **6.67**  $\lg\text{БОЕ}/\text{мл}$  через 120 ч после заражения, что соответствует множественности инфекции (МОИ) более 10 инф.ед. на клетку. Такая МОИ *in vivo* недостижима,



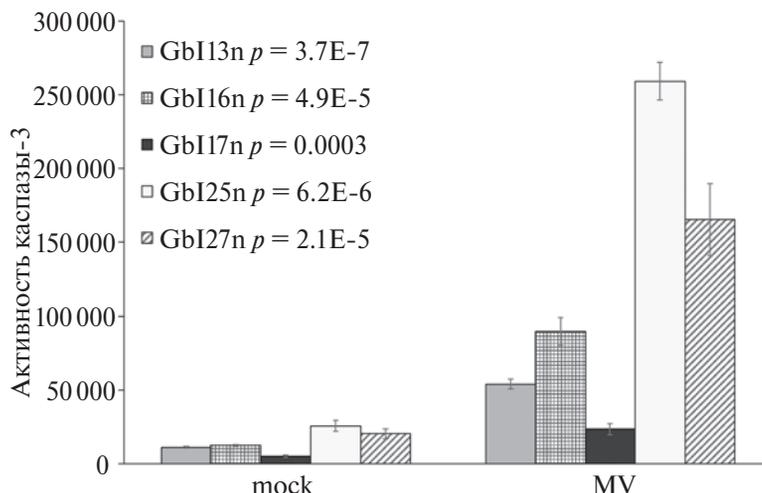
**Рис. 1.** Жизнеспособность пяти первичных культур клеток глиобластомы и иммортализованной культуры клеток NKE, инкубированных с вирусом паротита (MV) штамм Ленинград-3 (L-3), через 3–120 ч после заражения. По оси абсцисс отложен титр инокулированного вируса (БОЕ/мл), по оси ординат – жизнеспособность клеток (%), соответствующая значениям оптической плотности зараженной культуры клеток, выраженным в процентах от значений в незараженной культуре непосредственно перед заражением (0 ч), принятым за 100%. Значение “0” по оси абсцисс соответствует УФ-инактивированному MV. Полосы погрешностей обозначают стандартное отклонение.

поэтому клетки NKE также рассматриваются как резистентные.

Выявлено значительное увеличение активности проапоптотической каспазы-3 во всех MV-инфицированных ( $p < 0.0005$ ) клеточных линиях глиобластомы, включая Gbl17n, через 48 ч после заражения (рис. 2). Эти результаты показывают, что все

клетки глиобластомы, включая Gbl17n, отвечают на заражение MV.

Таким образом, можно предположить, что, используя гибель клеток в качестве защитного механизма, глиобластома может запускать апоптотический сигнальный путь в ответ на заражение MV, на что может указывать активация



**Рис. 2.** Активность каспазы-3 в неинфицированных (mock) и MV-инфицированных клетках глиобластомы. Клетки глиобластомы заражали MV с MOI 1.0, через 48 ч измеряли уровень активности каспазы-3. Одна единица измерения активности каспазы-3 – количество фермента, необходимое для расщепления 1.0 нМ субстрата Ac-DEVD-pNA в течение 1 ч при 37°C. Полосы погрешностей обозначают стандартное отклонение.

каспазы-3, возможно, за счет вирусного матричного белка [17].

*Штамм L-3 вируса паротита реплицируется в клетках глиобластомы человека*

Чувствительность клеток глиобластомы к MV – это способность к заражению небольшим количеством MV, тогда как перmissивность – это способность поддерживать размножение MV, т.е. обеспечивать транскрипцию и репликацию генетического материала, а также полноценную сборку вирионов. Известно, что чувствительность не всегда коррелирует с перmissивностью [18]. Связан ли MV-опосредованный онколиз с активной репликацией вируса внутри клеток глиобластомы или гибель клеток вызвана присутствием вирусной РНК? Чтобы понять это, с помощью ОТ-ПЦР-РВ определили кинетику репликации вируса в клетках глиобластомы [16] по содержанию вирусной РНК через 24, 48, 72 и 96 ч после заражения MV с MOI 1.0 и 0.1. В качестве контроля клетки инкубировали с УФ-инактивированным вирусом (MOI 1.0) в течение такого же времени (рис. 3). Ранее было показано, что выживаемость клеток, инкубированных с УФ-инактивированным вирусом, не отличалась значимо от выживаемости незараженных клеток (рис. 1).

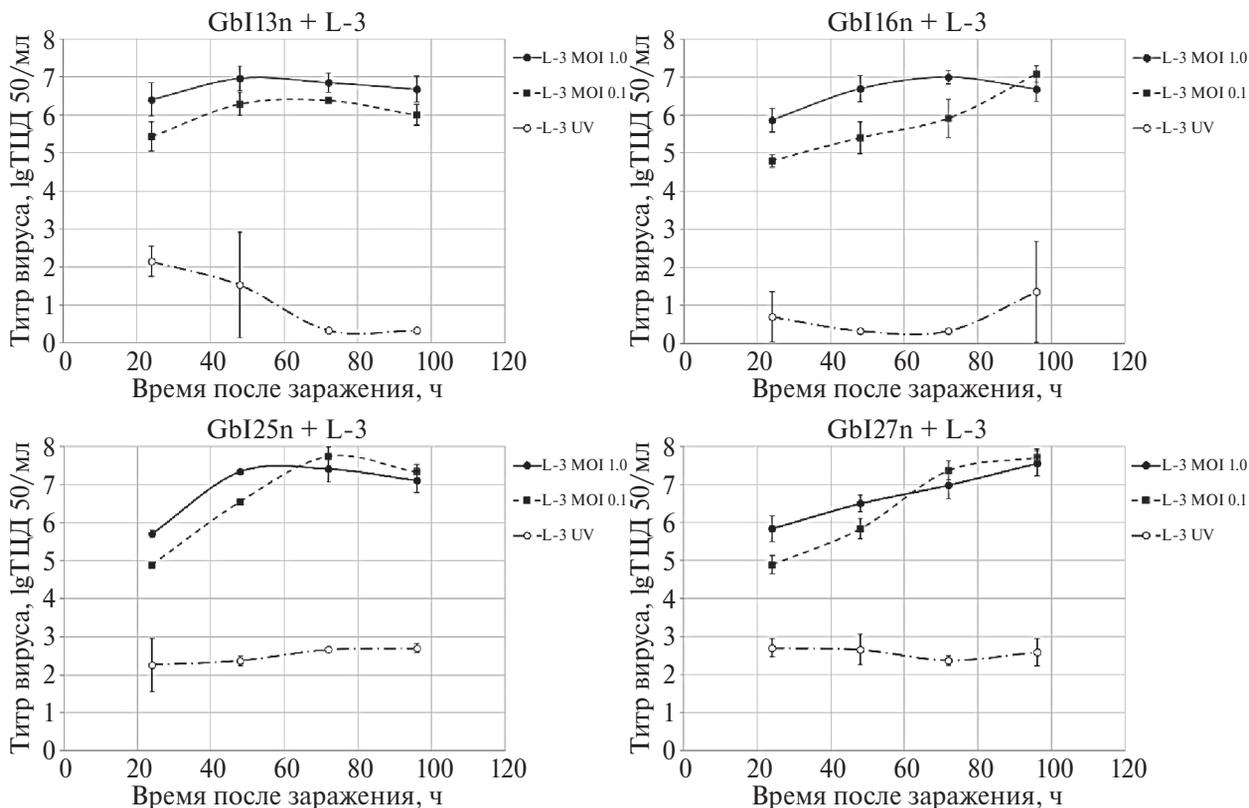
Как видно из рис. 3, вирусная РНК накапливается в культурах зараженных клеток GbI16n, GbI25n и GbI27n, что свидетельствует о чувствительности этих клеток к MV L-3, при этом в клетках GbI13n количество вирусной РНК оставалось практически постоянным в течение 96 ч культивирования. Содержание вирусной РНК в клетках GbI16n, GbI25n и GbI27n повышалось на 1–3 lgТЦД<sub>50</sub>/мл в течение 24–72 ч после заражения (рис. 3). В контрольных

клетках NKE и резистентных клетках глиобластомы GbI17n репликацию вируса не наблюдали.

Отмечено, что кинетика репликации вируса во всех протестированных клеточных линиях в целом коррелирует с показателями жизнеспособности. Таким образом, MV-опосредованный онколиз чувствительных культур клеток глиобластомы (GbI16n, GbI25n и GbI27n) вызван репликацией вирусной РНК. Однако присутствие репликативно-компетентной вирусной РНК также способно приводить к умеренному онколизу в культуре клеток GbI13n. При этом репликативно-некомпетентная вирусная РНК (УФ-инактивированный MV) не вызывает онколиз клеток глиобластомы.

*Инфицированные вирусом паротита L-3 клетки глиобластомы продуцируют интерферон*

Чувствительность опухолевых клеток к специфическим вирусам зависит как от вируса, прикрепленного к рецепторам клеточной мембраны, так и от генов клетки-хозяина, связанных с противовирусным ответом. Ранее сообщалось о дефектах сигнального пути IFN типа I во многих опухолевых клетках [19]. Для опухолевых клеток характерны специфические мутации в генах, кодирующих продукты IFN-системы. В свою очередь, повреждение этой системы в опухолевых клетках приводит, с одной стороны, к их усиленной пролиферации по сравнению с нормальными клетками, а с другой – к снижению противовирусной защиты, что делает опухолевую ткань относительно перmissивным субстратом для вирусного роста, в то время как соседние нормальные клетки остаются защищенными от заражения вирусом [11]. Чтобы объяснить причины различий в чувствительности клеточных линий глиобластомы к заражению MV, мы проанализировали экспрессию



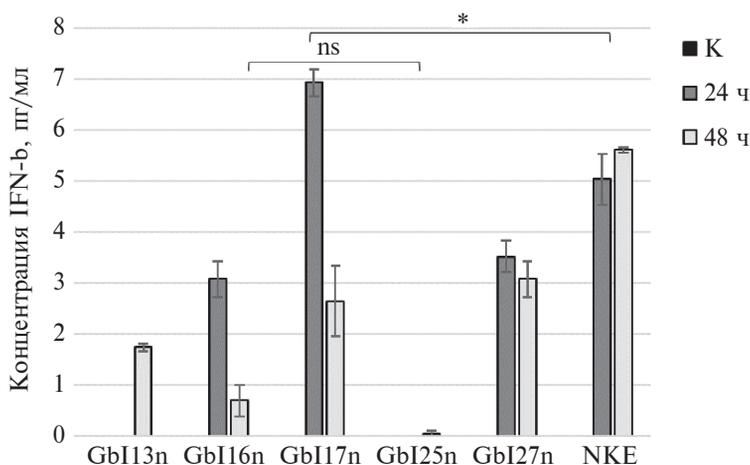
**Рис. 3.** Титры вируса паротита (MV) штамм Ленинград-3 (L-3), рассчитанные по накоплению вирусной РНК в клетках глиобластомы Gbl13n, Gbl16n, Gbl25n и Gbl27n. MV собирали с клеток глиобластомы, зараженных с MOI 0.1 (—■—) либо 1.0 (—●—) или зараженных УФ-инактивированным MV в качестве контроля. В каждой постановке реакции параллельно определяли пороговые циклы для разведений стандартного образца MV с известным титром вируса, по результатам которых строили калибровочную кривую. На нее наносили значения пороговых циклов исследуемых образцов и рассчитывали значения, выраженные в lgТЦД<sub>50</sub>/мл, которые условно принимали за титры [16].

главного ключевого компонента противовирусного ответа клетки: IFN-β (рис. 4).

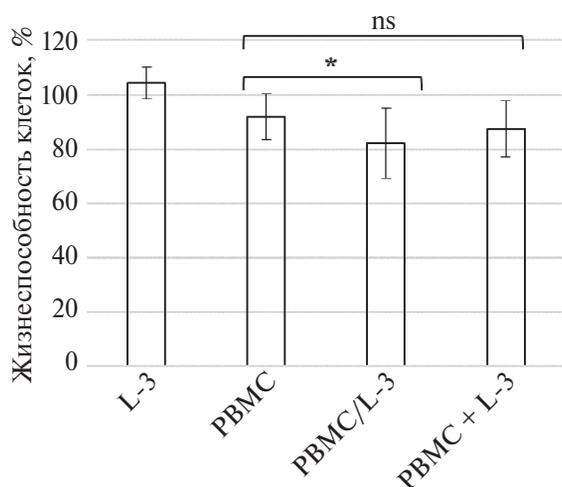
Чтобы добиться относительно большого количества инфицированных клеток глиобластомы и избежать асинхронной инфекции из-за

распространения вирусного потомства, клетки инфицировали MOI 1.0, а продукцию IFN анализировали через 24 и 48 ч после заражения.

Из рис. 4 видно, что только культура клеток глиобластомы Gbl25n не продуцировала IFN-β ни



**Рис. 4.** Продукция IFN-β MV-зараженными клетками глиобластомы. Клетки глиобластомы заражали MV с MOI 1.0, через 24 и 48 ч измеряли концентрацию IFN-β с помощью ИФА. К — незараженные культуры клеток. Полосы погрешностей обозначают стандартное отклонение. \**p* = 0.0244, ns: *p* = 0.0961.



**Рис. 5.** Жизнеспособность клеток глиобластомы Gbl13n через 9–12 ч после инокуляции вирусом паротита (MV) штамм Ленинград-3 (L-3); инкубации с мононуклеарными клетками периферической крови (PBMC); PBMC, предварительно стимулированными MV (PBMC/L-3), а также нестимулированными PBMC с одновременной инокуляцией MV (PBMC+L-3). По оси ординат отложена жизнеспособность клеток (%), соответствующая значениям оптической плотности, выраженным в процентах от значений для незараженной культуры, принятых за 100%. Полосы погрешностей обозначают стандартное отклонение. \* $p = 0.0192$ , ns:  $p = 0.1857$ .

через 24, ни через 48 ч после заражения MV, что позволяет считать клетки этой культуры дефектными по противовирусной защите. Однако культуры клеток Gbl13n, Gbl16n, Gbl17n и Gbl27n, а также неопухолевая культура клеток NKE продуцировали IFN- $\beta$  в ответ на заражение MV. Так, продукция IFN- $\beta$  клетками Gbl17n через 24 ч после заражения была значимо выше, чем клетками контрольной культуры NKE. Культуры клеток Gbl13n и Gbl25n не продуцировали IFN- $\beta$  через 24 ч после

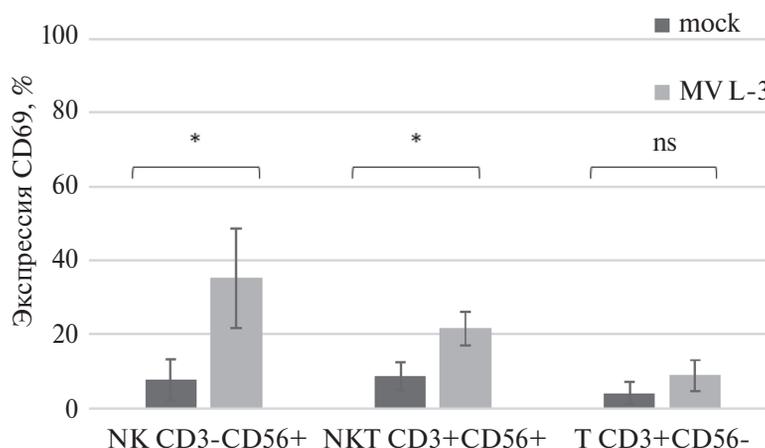
заражения, при этом через 48 ч уровень IFN- $\beta$  значимо повышался. Спустя 48 ч после заражения уровень IFN- $\beta$  в клетках глиобластомы Gbl16n снижался, не отличаясь статистически значимо от уровня в культуре клеток Gbl25n, дефектной по продукции IFN. Таким образом, чувствительность культуры клеток глиобластомы Gbl25n, а также резистентность культуры Gbl17n можно объяснить зависимостью онколиза от продукции IFN типа I. Однако эта зависимость была неочевидной для других культур клеток глиобластомы.

Таким образом, первичные клетки глиобластомы различались по способности индуцировать сигнальный путь IFN- $\beta$ , что указывает на различия в ответе исследованных линий на инфекцию MV.

*Вирус паротита L-3 стимулирует противоопухолевый иммунитет за счет активации NK-клеток ex vivo*

Далее мы оценили способность MV стимулировать противоопухолевый иммунитет в эффекторных PBMC здоровых доноров. PBMC выделяли из свежей периферической крови, отделяли от макрофагов, инкубировали с MV L-3 (MOI 1.0) в течение ночи и анализировали их способность убивать опухолевые клетки. Клеточную линию Gbl13n инкубировали с вирусом (MOI 0.1), с клетками PBMC не обработанными вирусом и в присутствии MV, а также с PBMC, стимулированными MV. Гибель клеток оценивали через 9–12 ч после заражения.

Обработка клеток Gbl13n PBMC, предварительно инкубированными с MV (PBMC/L-3), в соотношении E:T 10:1–20:1 более существенно усиливала лизис клеток глиобластомы, чем обработка нестимулированными PBMC (рис. 5). Однако не выявлено значимых отличий в жизнеспособности клеток Gbl13n, инкубируемых с нестимулированными PBMC и нестимулированными PBMC в комбинации с MV (PBMC+L-3).



**Рис. 6.** Активация NK-клеток. Результаты проточной цитометрии клеток PBMC до и после индукции вирусом паротита (MV). Экспрессия CD69 (%) на поверхности CD3-CD56+ клеток. \* $p = 0.014$ ; 0.027 и 0.122 соответственно.

Поскольку лизис клеток глиобластомы после инкубации с РВМС может быть опосредован НК-клетками [20], антигенспецифическую активацию НК-клеток после обработки вирусом оценили с помощью проточной цитометрии в трехпараметрическом анализе экспрессии маркера ранней фенотипической активации НК, CD69, на поверхности CD56-положительных CD3-негативных клеток. Известно, что активировать НК-клетки способен поверхностный белок HN парамиксовирусов, который связывается с поверхностными рецепторами НКр44 и НКр46 НК-клеток [21]. При этом взаимодействие белка HN с НКр44 или НКр46 на одной и той же НК-клетке может оказывать синергический эффект на активацию этих клеток в отношении опухолевых клеток. Таким образом, HN-опосредованная активация НК-клеток усиливает лизис клеток глиобластомы. В свою очередь, экспрессия маркера CD69 указывает на цитолитическую активность НК-клеток в отношении опухолевых мишеней [22]. Экспрессия CD69 повышалась на поверхности обработанных вирусом НК-клеток у всех доноров (рис. 6). Так, содержание CD69+ клеток, инкубированных с вирусом, составило 27–59%, из них в среднем 25–44% это клетки CD3-CD56+CD69+ (НК), 18–24% – CD3+CD56+CD69+ (NKT) и 9% – CD3+CD56-CD69+ (Т). Повышение экспрессии CD69 на поверхности НК-клеток указывает на их антигенспецифическую активацию и возможную иммуногенную гибель клеток глиобластомы за счет, в частности, цитотоксической активности НК-клеток.

Таким образом, нами впервые показана онколитическая активность отечественного вакцинного штамма MV L-13 в отношении клеток глиобластомы. Четыре из пяти исследованных клеточных линий глиобластомы оказались чувствительными к заражению MV L-3 в отличие от неопухолевой линии, где не наблюдали репродукцию вируса. Наиболее чувствительными к онколизу были клеточные линии Gbl25n и Gbl27n, в то время как культура клеток Gbl17n была резистентна к MV-опосредованному онколизу. Интересно отметить, что Gbl17n – это единственная из использованных нами культур клеток глиобластомы, не несущая мутаций в гене TP53. Мутации в гене TP53, кодирующем белок-супрессор опухолевого роста, характерны для клеток глиобластомы. Ранее Gil-Ranedo J. и соавт. показали, что TP53-мутированные стволовые клетки глиобластомы служат селективными мишенями для парвовируса [23]. Кинетика накопления вирусной РНК в исследованных линиях клеток глиобластомы коррелировала со скоростью гибели клеток. Клеточные линии глиобластомы различались по уровню продукции IFN типа I в ответ на заражение MV L-3. Кроме того, показано, что заражение MV способно вызывать иммуногенную гибель клеток глиобластомы.

Таким образом, вакцинный штамм MV L-3 эффективно реплицируется в клетках глиобластомы человека с мутантным TP53, что приводит к гибели опухолевых клеток. Полученные данные могут стать основой для дальнейшего исследования MV как потенциального онколитического агента.

Работа выполнена с использованием вакцинного штамма Ленинград-3 из коллекции Центра коллективного пользования “НИИВС им. И. И. Мечникова” (соглашение № 075-15-2021-676 от 28.07.2021).

Работа выполнена при финансовой поддержке Российского научного фонда (грант № 22-15-00304).

Все процедуры, выполненные в исследовании с участием людей, соответствуют этическим стандартам институционального и национального комитета по исследовательской этике и Хельсинкской декларации 1964 г. и ее последующим изменениям или сопоставимым нормам этики. Протокол исследования одобрен локальным этическим комитетом ФГБУ “Национальный медицинский исследовательский центр онкологии им. Н. Н. Блохина” Минздрава Российской Федерации (№ 7 от 27 июля 2023 г.). От каждого пациента, принимавшего участие в исследовании, получено информированное добровольное согласие.

Авторы заявляют об отсутствии конфликта интересов.

## СПИСОК ЛИТЕРАТУРЫ

1. Николаева Е.Ю., Щетинина Ю.Р., Шохин И.Е., Зверев В.В., Свитич О.А., Сусова О.Ю., Митрофанов А.А., Аммури Ю.И. (2022) Вирус кори как векторная платформа для иммунотерапии опухолей головного мозга (обзор). *Разработка и регистрация лекарственных средств*. 11, 51–58.
2. Stupp R., Mason W.P., Van den Bent M.J., Weller M., Fisher B., Taphoorn M.J., Belanger K., Brandes A.A., Marosi C., Bogdahn U., Curschmann J., Janzer R.C., Ludwin S.K., Gorlia T., Allgeier A., Lacombe D., Cairncross J.G., Eisenhauer E., Mirimanoff R.O.; European organization for research and treatment of cancer brain tumor and radiotherapy groups; National Cancer Institute of Canada clinical trials group. (2005) Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. *N. Engl. J. Med.* 352, 987–996.
3. Stupp R., Hegi M.E., Mason W.P., Van den Bent M.J., Taphoorn M.J., Janzer R.C., Ludwin S.K., Allgeier A., Fisher B., Belanger K., Hau P., Brandes A.A., Gijtenbeek J., Marosi C., Vecht C.J., Mokhtari K., Wesseling P., Villa S., Eisenhauer E., Gorlia T., Weller M., Lacombe D., Cairncross J.G., Mirimanoff R.O.; European organization for research and treatment of cancer brain tumour and radiation oncology groups; National Cancer Institute of Canada clinical trials group. (2009) Effects of radiotherapy with concomitant and adjuvant temozolomide versus radiotherapy alone on survival in glioblastoma in a randomised phase III study:

- 5-year analysis of the EORTC-NCIC trial. *Lancet Oncol.* **10**, 459–466.
4. Taylor O.G., Brzozowski J.S., Skelding K.A. (2019) Glioblastoma multiforme: an overview of emerging therapeutic targets. *Front. Oncol.* **9**, 963.
  5. Suryawanshi Y.R., Schulze A.J. (2021) Oncolytic viruses for malignant glioma: on the verge of success? *Viruses.* **13**, 1294.
  6. Ammour Y., Susova O., Krasnov G., Nikolaeva E., Varachev V., Schetinina Y., Gavrilova M., Mitrofanov A., Poletaeva A., Bekyashev A., Faizuloev E., Zverev V.V., Svitich O.A., Nasedkina T.V. (2022) Transcriptome analysis of human glioblastoma cells susceptible to infection with the Leningrad-16 vaccine strain of measles virus. *Viruses.* **14**, 2433.
  7. Asada T. (1974) Treatment of human cancer with mumps virus. *Cancer.* **34**, 1907–1928.
  8. Okuno Y., Asada T., Yamanishi K., Otsuka T., Takahashi M., Tanioka T., Aoyama H., Fukui O., Matsumoto K., Uemura F., Wada A. (1978) Studies on the use of mumps virus for treatment of human cancer. *Biken J.* **21**, 37–49.
  9. Shimizu Y., Hasumi K., Okudaira Y., Yamanishi K., Takahashi M. (1988) Immunotherapy of advanced gynecologic cancer patients utilizing mumps virus. *Cancer Detect Prev.* **12**, 487–495.
  10. Oka N. (1988) Experimental studies of antineoplastic therapy using mumps virus for malignant brain tumor. *J. Kansai Med. Univ.* **40**, 19–43.
  11. Аммур Ю.И., Рябая О.О., Милованова А.В., Сидоров А.В., Шохин И.Е., Зверев В.В., Наседкина Т.В. (2018) Исследование онколитических свойств вакцинного штамма вируса паротита на клеточных линиях меланомы человека. *Молекуляр. биология.* **52**, 659–666.
  12. Alirezaie B., Mohammadi A., Ghalyanchi Langeroudi A., Fallahi R., Khosravi A.R. (2020) Intrinsic oncolytic activity of Hoshino mumps virus vaccine strain against human fibrosarcoma and cervical cancer cell lines. *Int. J. Cancer Manag.* **13**, e103111.
  13. Myers R., Greiner S., Harvey M., Soeffker D., Frenzke M., Abraham K., Shaw A., Rozenblatt S., Federspiel M.J., Russell S.J., Peng K.W. (2005) Oncolytic activities of approved mumps and measles vaccines for therapy of ovarian cancer. *Cancer Gene Ther.* **12**, 593–599.
  14. Behrens M.D., Stiles R.J., Pike G.M., Sikkink L.A., Zhuang Y., Yu J., Wang L., Boughey J.C., Goetz M.P., Federspiel M.J. (2022) Oncolytic Urabe mumps virus: a promising virotherapy for triple-negative breast cancer. *Mol. Ther. Oncolytics.* **27**, 239–255.
  15. Nasedkina T., Varachev V., Susova O., Krasnov G., Poletaeva A., Mitrofanov A.A., Naskhletashvili D., Bekyashev A. (2021) 350P Molecular profiling of tumor tissue and tumor-derived cell lines in patients with glioblastoma. *Ann. Oncol.* **32**, S519.
  16. Ammour Y., Faizuloev E., Borisova T., Nikonova A., Dmitriev G., Lobodanov S., Zverev V. (2013) Quantification of measles, mumps and rubella viruses using real-time quantitative TaqMan-based RT-PCR assay. *J. Virol. Methods.* **187**, 57–64.
  17. Morovati S., Mohammadi A., Masoudi R., Heidari A.A., Asad Sangabi M. (2023) The power of mumps virus: matrix protein activates apoptotic pathways in human colorectal cell lines. *PLoS One.* **18**, e0295819.
  18. Laksono B.M., Grosserichter-Wagener C., de Vries R.D., Langeveld S.A.G., Brem M.D., van Dongen J.J.M., Katsikis P.D., Koopmans M.P.G., van Zelm M.C., de Swart R.L. (2018) *In vitro* measles virus infection of human lymphocyte subsets demonstrates high susceptibility and permissiveness of both naive and memory B cells. *J. Virol.* **92**, e00131–18.
  19. Lichty B.D., Breitbach C.J., Stojdl D.F., Bell J.C. (2014) Going viral with cancer immunotherapy. *Nat. Rev. Cancer.* **14**, 559–567.
  20. Marden C.M., North J., Anderson R., Bakhsh I.A., Addison E., Pittman H., Mackinnon S., Lowdell M.W. (2005) CD69 is required for activated NK cell-mediated killing of resistant targets. *Blood.* **106**, 3322.
  21. Jarahian M., Watzl C., Fournier P., Arnold A., Djandji D., Zahedi S., Cerwenka A., Paschen A., Schirrmacher V., Momburg F. (2009) Activation of natural killer cells by newcastle disease virus hemagglutinin-neuraminidase. *J. Virol.* **83**, 8108–8121.
  22. Donnelly O.G., Errington-Mais F., Steele L., Hadac E., Jennings V., Scott K., Peach H., Phillips R.M., Bond J., Pandha H., Harrington K., Vile R., Russell S., Selby P., Melcher A.A. (2013) Measles virus causes immunogenic cell death in human melanoma. *Gene Ther.* **20**, 7–15.
  23. Gil-Ranedo J., Gallego-García C., Almendral J.M. (2021) Viral targeting of glioblastoma stem cells with patient-specific genetic and post-translational p53 deregulations. *Cell Rep.* **36**, 109673.

## Sensitivity of Primary Human Glioblastoma Cell Lines to Mumps Virus Vaccine Strain

E. Yu. Nikolaeva<sup>1</sup>, Y. R. Zhelayeva<sup>1</sup>, O. Yu. Susova<sup>2</sup>, A. A. Mitrofanov<sup>2</sup>, V. O. Varachev<sup>3</sup>, T. V. Nasedkina<sup>3</sup>, V. V. Zverev<sup>1</sup>, O. A. Svitich<sup>1</sup>, Y. I. Ammour<sup>1, \*</sup>

<sup>1</sup>*Mechnikov Research Institute for Vaccines and Sera, Moscow, 105064 Russia*

<sup>2</sup>*Blokhin Russian Cancer Research Center, Moscow, 115478 Russia*

<sup>3</sup>*Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Moscow, 119991 Russia*

\**e-mail: yulia.ammour@yahoo.fr*

The sensitivity of human glioblastoma cells to virus-mediated oncolysis was investigated on five patient-derived cell lines. Primary glioblastoma cells (Gbl13n, Gbl16n, Gbl17n, Gbl25n, and Gbl27n) were infected with 10-fold serial dilutions of the Leningrad-3 strain of mumps virus, virus reproduction and cytotoxicity were monitored for 96–120 hours. Immortalized human non-tumor NKE cells were used as controls to determine virus specificity. Four out of the five glioblastoma cell lines examined were susceptible to mumps virus infection, whereas no virus reproduction was observed in the non-tumor cell line. Moreover, the level of proapoptotic caspase-3 activity was increased in all infected cells 48 hours after infection. The kinetics of viral RNA accumulation in the studied glioblastoma cell lines was comparable with the rate of cell death. The data suggest that glioblastoma cell lines are permissive for mumps virus. Glioblastoma cell lines differed in type I IFN production in response to mumps virus infection. In addition, it was shown that MV infection was able to induce immunogenic death of glioblastoma cells.

**Keywords:** glioblastoma, oncolytic viruses, mumps virus

УДК 577.21

## БЕЛОК CG9609 ДРОЗОФИЛЫ, СОДЕРЖАЩИЙ ДОМЕНЫ ЦИНКОВЫХ ПАЛЬЦЕВ, ВЗАИМОДЕЙСТВУЕТ С ДЕУБИКВИТИНИРУЮЩИМ (DUB) МОДУЛЕМ КОМПЛЕКСА SAGA И УЧАСТВУЕТ В РЕГУЛЯЦИИ ТРАНСКРИПЦИИ

© 2024 г. Ю. В. Николенко<sup>a</sup>, М. М. Куршакова<sup>a</sup>, Д. В. Копытова<sup>a</sup>, Ю. А. Вдовина<sup>a</sup>, Н. Е. Воробьева<sup>b</sup>, А. Н. Краснов<sup>b, \*</sup>

<sup>a</sup>Институт молекулярной биологии им. В. А. Энгельгардта Российской академии наук, Москва, 119991 Россия

<sup>b</sup>Институт биологии гена Российской академии наук, Москва, 119334 Россия

\*e-mail: krasnov@genebiology.ru

Поступила в редакцию 17.11.2023 г.

После доработки 12.12.2023 г.

Принята к публикации 12.12.2023 г.

Ранее мы показали, что белки Su(Hw) и CG9890, содержащие домены цинковых пальцев, взаимодействуют с комплексом SAGA дрозофилы, участвуют в формировании активной структуры хроматина и регуляции транскрипции. В представленном исследовании обнаружено взаимодействие DUB-модуля комплекса SAGA с белком CG9609 – еще одним белком с доменами цинковых пальцев. Проведен ChIP-Seq-анализ и идентифицированы сайты связывания белка CG9609 в геноме дрозофилы, локализованные, согласно анализу сайтов связывания, преимущественно на промоторах генов. Показано, что белок CG9609 участвует в регуляции экспрессии тех генов, на промоторах которых он локализован.

**Ключевые слова:** белок CG9609, цинковые пальцы, активация транскрипции, DUB-модуль, комплекс SAGA

DOI: 10.31857/S0026898424040072, EDN: INAPWW

### ВВЕДЕНИЕ

Регуляция экспрессии генов эукариот – сложный процесс, который происходит на нескольких последовательных этапах транскрипции, процессинга мРНК, экспорта мРНК из ядра, трансляции и фолдинга белков [1]. Локальная структура хроматина, положение гена относительно функциональных ядерных компартментов и дальние взаимодействия регуляторных элементов являются дополнительным уровнем регуляции генетических процессов у эукариот в контексте сложной организации их генома в трехмерном пространстве ядра [2–5]. Белок ENY2 – мультифункциональный фактор, участвующий в различных стадиях экспрессии генов [6–11]. Разнообразие клеточных функций ENY2 определяется активностями тех белковых комплексов, в состав которых он входит. Так показано, что ENY2 является субъединицей деубиквитирующего (DUB) модуля комплекса SAGA – важного коактиватора транскрипции у дрозофилы [12, 13]. В состав DUB-модуля входят также белки Sgf11 и Nonstop [14, 15]. Этот комплекс вместе с комплексом SWI/SNF ремоделирования хроматина присутствует на промоторах генов и участвует в создании активной структуры

хроматина, что необходимо для регуляции транскрипции [16, 17].

Ранее нами было показано, что белок Su(Hw), содержащий домены цинковых пальцев, взаимодействует с ENY2-содержащими комплексами дрозофилы и необходим для их привлечения на сайты связывания Su(Hw), что необходимо для создания активной структуры хроматина и позиционирования участков начала репликации (ориджинов) [6, 7, 10]. Высказано предположение о существовании целого ряда “цинковых” белков, которые взаимодействуют с ENY2-содержащими комплексами дрозофилы, что необходимо для привлечения этих комплексов на различные регуляторные элементы генома. Так, обнаружено взаимодействие ENY2 с еще одним цинковым белком CG9890 [18]. Этот белок локализован преимущественно на промоторах активных генов и колокализуется с комплексами SAGA, ORC и dSWI/SNF на своих сайтах связывания. Показано, что белок CG9890 участвует в регуляции транскрипции, включая регуляцию экспрессии генов эдизонового каскада [19, 20].

Данная работа посвящена идентификации и изучению новых белков, содержащих домены цинковых пальцев, которые взаимодействуют

с DUB-модулем комплекса SAGA дрозофилы. Такие белки, подобно белку Su(Hw), могут быть важны для привлечения SAGA и других комплексов на различные регуляторные элементы генома, включая промоторы, для формирования активной структуры хроматина и регуляции транскрипции.

## ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

**Культивирование клеток линии Schneider 2 (S2). РНК-интерференция.** Клетки культивировали при 25°C в среде Schneider's Insect Medium ("Sigma", США), содержащей 10% фетальной сыворотки крупного рогатого скота ("HyClone", США). Транфекцию клеток проводили при помощи Effectene Transfection Reagent ("Qiagen", Германия) по протоколу производителя. дцРНК синтезировали с использованием TranscriptAid T7 High Yield Transcription Kit ("ThermoScientific", США). Нокдаун генов проводили с помощью РНК-интерференции согласно протоколу [21]. Для нокдауна гена *CG9609* синтезировали дцРНК длиной 600 п.н., соответствующую 3'-концевому участку кодирующей последовательности. В качестве контроля использовали дцРНК, соответствующую фрагменту плазмиды pBluescript II SK(-) ("Stratagene", США) длиной 500 п.н. дцРНК синтезировали с использованием следующих праймеров:

*CG9609* – CGACTCACTATAGGGAGAAGCTCGT TGCACGGCAAGAAT и CGACTCACTATAGGGAG ATGCTCCTTTTCTCTTCAT;

*pBluescriptIIISK(-)* – GAATTAATACGACTCACTA TAGGGAGAGTTACATGATCCCCCATG и GAATT AATACGACTCACTATAGGGAGATTTTCGCCCCG AAGAACG.

В каждый эксперимент по РНК-интерференции дцРНК брали из расчета 30 мкг/млн клеток. По прошествии 5 дней инкубации выделяли РНК и измеряли уровни мРНК исследуемых генов методом обратной транскрипции с последующей ПЦР в реальном времени. Были использованы следующие праймеры:

*bml* – TCGTTTAGATGCGAGCAATTTCG и AGCAATTTTCAGCTTGTATGCG;  
*cv-2* – AGCAGCAGCAGTAGAAGCACCT и САСТТСТСГТТАААТГСТСГГС;  
*pnt* – GGGCTACTTCAATGATATGGCG и GTTGTTGATGCGGTCGTGTG;  
*exn* – AGAGTGAAACAAGCAAGTGGATTG и CGATCACGTTCTCGGCTATCC;  
*Kek1* – ACCCTGCTGCCTGGAATGAT и GCATGTAGCCATCCTCAGTCATC;  
*sgl* – CGTGATAACTCTGGTGGACAAG и GTCTCGATGTCCGTTGAGAAG;  
*wg* – АСТГСТСГСГСГАААСТТСТС и TGTAGTCGCAGGTGCAGGAC;  
*Psc* – CGCTGACTCAAACTCAAGTGTG и TCGATTCTGGCGTCATCATAGTA;

*CG1907* – CCGCTAGACTTGGTCAAGACC и TGTCTCAGCAGAGCAGCTCC;  
*Ras64B* – GAGGGATTCTGCTCGTCTTCG и GTCGCACTTGTACCCACCATC.

**Клонирование конструкций для экспрессии белков, слитых с 3xFLAG-эпитопом.** Конструкции создавали в векторе pAc 5.1/V5-His, разработанном для экспрессии белков под контролем актинового промотора в клетках дрозофилы. В данный вектор мы клонировали последовательность 3xFLAG-эпитопа с сайтом рестрикции HindIII непосредственно после нее. Далее с использованием новосозданного вектора клонировали различные последовательности для получения полноразмерных белков (без первого остатка метионина). Гены клонировали после 3xFLAG-эпитопа по сайту HindIII. Второй сайт рестрикции выбирали из полилинкера исходного вектора. Таким образом 3xFLAG-эпитоп оказывался на N-конце слитых белков. кДНК требуемых участков генов синтезировали с помощью ОТ-ПЦР, сайты клонирования интегрировали в праймеры. Правильность полученных конструкций проверяли секвенированием.

**Иммунопреципитация. Получение антител.** Для выделения ядер клетки центрифугировали в течение 5 мин (500 g, +4°C). Осадок клеток промыли 1 мл буфера LB cyto 3 (3 mM MgCl<sub>2</sub>, 20 mM Hepes-NaOH, pH 8.0) с добавлением бутирата натрия (ингибитор деацетилазы) до конечной концентрации 20 mM. Затем проводили повторное центрифугирование в тех же условиях. Осадок клеток тщательно ресуспендировали в 200 мкл буфера LB cyto 3 с добавлением бутирата натрия до конечной концентрации 10 mM и ингибитора протеаз (Protease Inhibitor Cocktail (PIC), "Roche", Швейцария), после чего инкубировали на льду в течение 15 мин. После центрифугирования избавлялись от супернатанта и использовали далее только ядерную фракцию.

Осадок ядер ресуспендировали в 500 мкл буфера MN III (20 mM Hepes-KOH, 3 mM MgCl<sub>2</sub>, 0.1% NP40, 0.1 M KCl, pH 8.0) с добавлением бутирата натрия до конечной концентрации 20 mM и ингибитора протеаз (PIC, "Roche"). ДНК фрагментировали обработкой ядер ультразвуком (2 раза по 10 с, перерыв – 1 мин, средняя мощность прибора) на льду, инкубировали с 2 ед. ДНКазы I на льду в течение 30 мин и центрифугировали (16000 g, 20 мин, +4°C).

Коиммунопреципитацию проводили с использованием поликлональных антител к белкам Sgf11 и ENY2 и иммуноглобулинов сыворотки крови неиммунизированного кролика в качестве отрицательного контроля. Антитела иммобилизовали на Mab-сефарозе.

Белковый экстракт (200 мкл) из клеток линии S2 инкубировали с 15 мкл 50%-ной Mab-сефарозы с иммобилизованными антителами в течение 3 ч на

ротаторе при +4°C. Сефарозу отмывали буфером MN III (3 раза по 10 мин, +4°C), иммунные комплексы элюировали с сефарозы в 28 мкл 2 × буфера для нанесения по Лэммли с добавлением 3 мкл 1 М дитиотреитола (DTT). Результаты эксперимента анализировали с помощью Вестерн-блот-анализа.

Поликлональные антитела α-CG9609 получены из сыворотки крови крысы, иммунизированной С-концевым фрагментом (102 аминокислотных остатка) белка CG9609, экспрессированным в *E. coli*. Этот фрагмент соответствует области, расположенной после доменов цинковых пальцев, и до конца белка. При получении антител к цинковым белкам мы не используем домены цинковых пальцев, так как это может дать перекрестную реакцию с другими цинковыми белками.

**Идентификация сайтов связывания белка CG9609 в геноме дрозофилы.** Исходные данные геномного секвенирования (секвенатор Illumina NovaSeq 6000) выравнивали на геном дрозофилы *dmel\_r6.40* с помощью программы Hisat2 в режиме end-to-end с ключом no-spliced-alignment. Также устанавливали ключ “-a”, который позволяет искать множественные выравнивания, с целью их исключения из анализа. В дальнейшую работу брали только уникально картированные риды, обрабатывая флаг “NH: i” в результатах вывода программы Hisat2. Также брали только те риды, уровень мисматчей в которых не превышал 5%, обрабатывая флаг “XM: i” в результатах вывода программы Hisat2. Фильтрацию осуществляли с помощью собственной программы. Сайты связывания белка CG9609 определяли с помощью программы MACS2 с использованием параметров по умолчанию ( $q$ -value < 0.05).

**Анализ сайтов связывания белка CG9609 в промоторных областях генов.** Промоторные области определяли из аннотации транскриптов генома дрозофилы *dmel\_r6.40*. Промоторной областью считали интервал ±500 п.н. от старта транскрипции. Считали, что сайт связывания находится в промоторной области гена, если пик сайта связывания попадает в заданный интервал. Анализ проводили с помощью программы bedtools. Кроме

того, сайты связывания белка CG9609 в промоторных областях генов дополнительно контролировали путем совмещения структуры транскриптов из геномного браузера и профиля ChIP-Seq.

Результаты ChIP-Seq размещены в базе данных Gene Expression Omnibus (идентификатор эксперимента GSE250183). Кроме этих результатов, в базе содержится также протокол иммунопреципитации хроматина, создания библиотек и обработки данных.

Профили ChIP-Seq получали с помощью программы bamCoverage 3.5.4. Значения в профиле представляют собой нормированное (CPM – Counts Per Million mapped reads) значение ДНК-фрагментов в каждой геномной позиции.

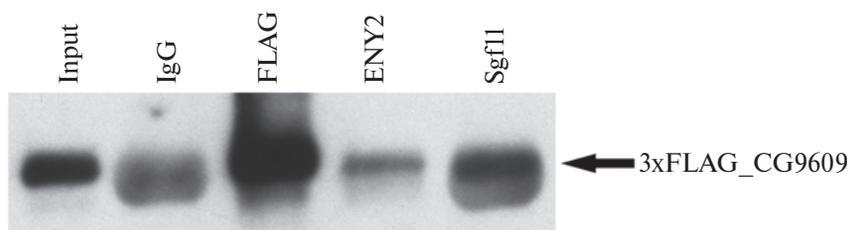
## РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

### *Идентификация цинкового белка CG9609 – нового взаимодействующего партнера DUB-модуля*

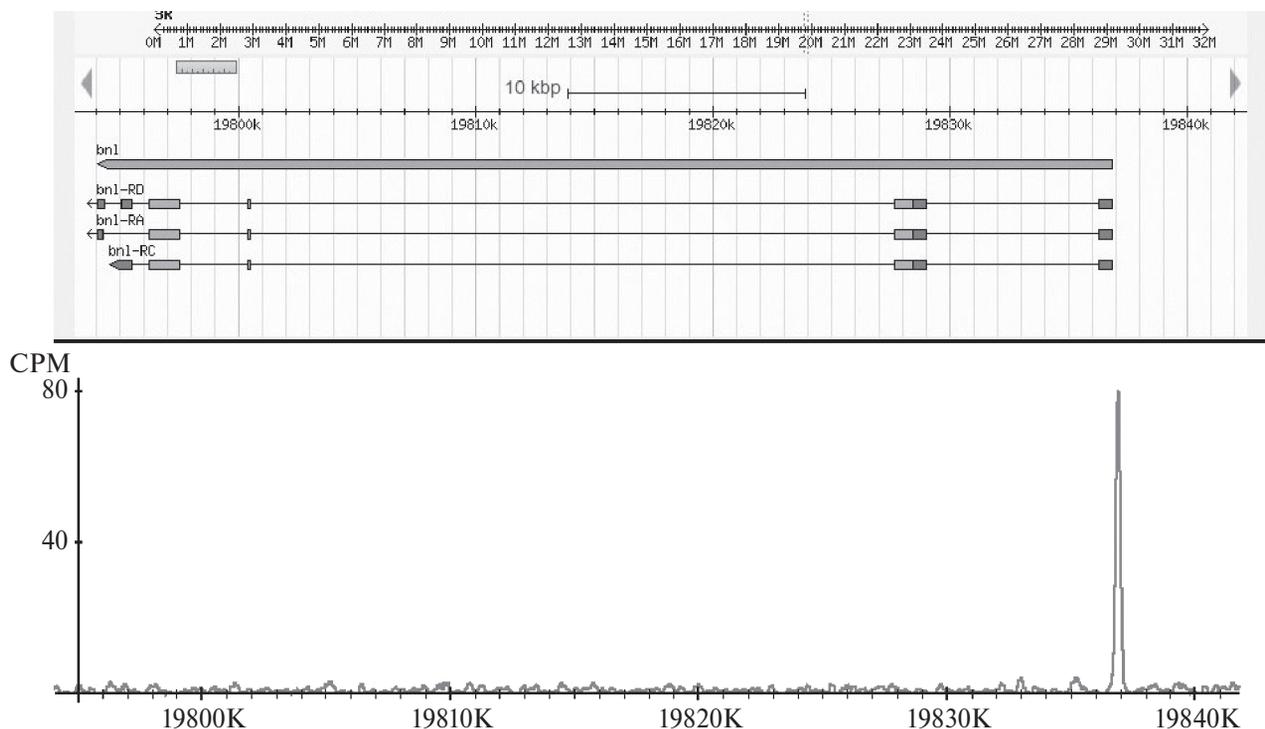
Ранее нами был очищен DUB-модуль комплекса SAGA дрозофилы. С помощью MALDITOF MS идентифицировано несколько сотен белков, которые соосаждались с DUB-модулем за антитела к белку Sgf11 [22].

Проведение масс-спектрометрического анализа показало, что в состав фракции, содержащей DUB-модуль комплекса SAGA, входят также 55 белков с доменами цинковых пальцев. Из числа белков этой фракции были исключены белки, представленные также в отрицательном контроле (белковый комплекс, очищенный с использованием неспецифических антител), а затем отобраны белки, содержащие кластер близкорасположенных цинковых пальцев с каноническим линкером (TGE[K/R]P) между ними, так как именно такая структура с максимальной вероятностью предсказывает ДНК-связывающие свойства белка со специфичным сайтом связывания [23, 24]. На этом этапе были отобраны 26 белков с доменами цинковых пальцев.

В ходе экспериментов по очистке комплексов с последующей масс-спектрометрией мы анализировали несколько фракций, полученных после гель-фильтрации. Показано разделение фракций,



**Рис. 1.** Коиммунопреципитация белка 3xFLAG\_CG9609 с белками ENY2, Sgf11. Указаны антитела, которые использовали в иммунопреципитации. Input – исходный экстракт, IgG – иммунопреципитация неспецифическими антителами. Вестерн-блот окрашивали антителами к 3xFLAG-эпитопу. Стрелкой отмечен белок 3xFLAG\_CG9609. Видно, что антитела к ENY2 и Sgf11 преципитируют 3xFLAG\_CG9609. В образце IgG широкая полоса соответствует тяжелой цепи антител, белок 3xFLAG\_CG9609 в нем отсутствует.



**Рис. 2.** ChIP-Seq профиль белка CG9609 в области гена *bnl*. На верхней панели представлена структура гена из геномного браузера. На нижней панели представлен сам ChIP-Seq профиль в единицах CPM. Видно, что белок CG9609 локализован на промоторе гена *bnl*.

содержащих DUB-модуль комплекса SAGA (включает белки ENY2, Sgf11), и остальной части SAGA (содержит белки GCN5, Ada2b) без DUB-модуля. Были отобраны цинковые белки, которые присутствуют во фракции DUB-модуля. Таким образом, в дальнейший анализ взяли 12 белков: CG9609, CG10543, CG4903, CG34100, CG9932, CG12701, CG3998, CG4936, CG11696, CG5683, CG8474, CG11352. В этой фракции обнаружен также белок Su(Hw), взаимодействие которого с DUB-модулем показано нами ранее.

Взаимодействие этих белков с DUB-модулем комплекса SAGA проверяли с помощью генетических конструкций, созданных для экспрессии каждого из 12 белков, маркированных 3xFLAG-эпитопом. Каждую из этих генетических конструкций трансфицировали по отдельности в клетки S2 дрозофилы для экспрессии слитых белков, а затем проводили иммунопреципитацию из лизата этих клеток с использованием антител к белкам ENY2 и Sgf11 и неспецифических антител в качестве отрицательного контроля. Результаты иммунопреципитации анализировали с помощью Вестерн-блоттинга с помощью антител к 3xFLAG эпитопу (“Sigma”). Как видно из рис. 1, антитела к белку ENY2 и Sgf11 преципитируют белок 3xFLAG\_CG9609 в отличие от неспецифических антител (IgG). В образце IgG широкая полоса соответствует тяжелой цепи антител, белок 3xFLAG\_CG9609 в нем отсутствует. Таким образом, подтверждено взаимодействие

белка CG9609 с DUB-модулем комплекса SAGA. Обнаружено также слабое взаимодействие с белками CG10543 и CG5683 (Aef1) (данные не представлены). Взаимодействие остальных 9 белков с DUB-модулем подтвердить не удалось.

С целью дальнейшего изучения CG9609 получены поликлональные антитела к этому белку, которые аффинно очистили на колонке, содержащей рекомбинантный белок CG9609.

#### *Определение сайтов связывания белка CG9609 в геноме дрозофилы*

Сайты локализации CG9609 в геноме идентифицировали с помощью ChIP-Seq-анализа. Для этого провели иммунопреципитацию хроматина с антителами к CG9609 (в двух повторах) и получили библиотеки для высокопроизводительного секвенирования на геномном секвенаторе Illumina NovaSeq 6000. Вероятные места локализации белка идентифицировали с помощью программы Macs2 (всего выявлено около 1500 сайтов). Анализ сайтов локализации CG9609 в геноме показал, что 75% из них совпадают с промоторами генов. У сотни наиболее сильных сайтов связывания (в списке, отсортированном по значению ChIP/Input) этот показатель достигает 95%. Таким образом, белок CG9609 локализуется преимущественно на промоторах генов. Характерный ChIP-Seq-профиль белка CG9609 на одном из его сайтов связывания представлен на рис. 2.



Рис. 3. Консенсусная последовательность потенциального сайта связывания белка CG9609.

Далее была определена потенциальная консенсусная последовательность сайта связывания белка CG9609. С этой целью с помощью программы bedtools получены последовательности длиной 500 п.н. вокруг каждого пика ChIP-Seq, которые проанализировали в программе MEME-ChIP. На рис. 3 представлена консенсусная последовательность, имеющая наибольшее обогащение во фрагментах ChIP-Seq. Эта консенсусная последовательность имеет высокую статистическую значимость, центральное обогащение в ChIP-Seq-фрагментах и не является сайтом связывания какого-либо известного белка. Эту последовательность можно рассматривать как потенциальный сайт связывания белка CG9609.

*Белок CG9609 принимает участие в регуляции экспрессии генов*

Учитывая, что CG9609 найден преимущественно на промоторах генов, мы решили исследовать, к каким изменениям в экспрессии

CG9609-ассоциированных генов приведет снижение внутриклеточного уровня белка CG9609 с помощью РНК-интерференции. В результате оптимизации условий РНК-интерференции удалось добиться эффективного снижения экспрессии изучаемого белка в клетках: более чем в 9 раз по количеству мРНК и фактически полную деплецию по количеству белка (ниже предела детекции с помощью Вестерн-блоттинга).

Используя метод ОТ-кПЦР, мы проанализировали количество РНК девяти CG9609-ассоциированных генов в образцах, выделенных из контрольных клеток и клеток после РНК-интерференции. Результаты этого эксперимента представлены на рис. 4. После нокдауна гена *CG9609* количество мРНК пяти из девяти генов достоверно возросло. Таким образом, белок CG9890 действительно участвует в регуляции экспрессии тех генов, на промоторах которых он локализуется, выступая в роли репрессора.

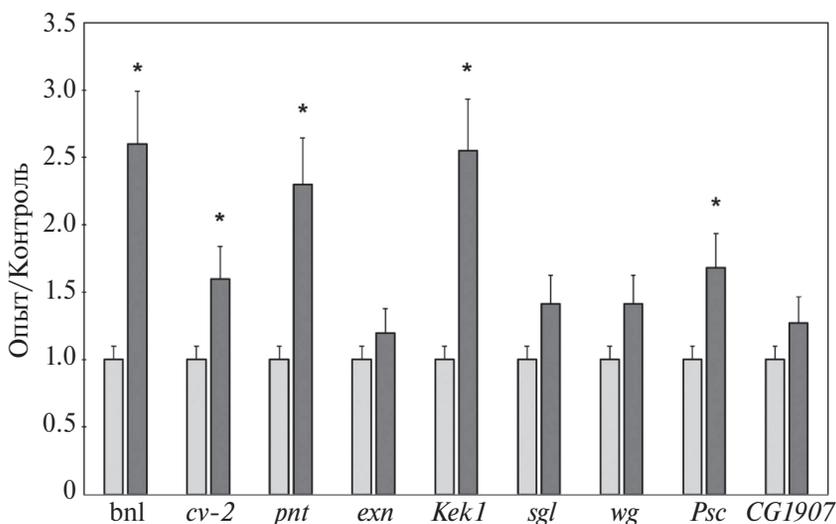


Рис. 4. Транскрипция некоторых генов, содержащих CG9609 на промоторах, при РНК-интерференции *CG9609*. Названия генов указаны внизу. Светлые столбцы соответствуют транскрипции гена в норме (принято за единицу). Темные столбцы соответствуют транскрипции генов при РНК-интерференции *CG9609*. По оси ординат указано соотношение транскрипции в опыте и в контроле. Уровень транскрипции гена *CG9609* при РНК-интерференции был в 9 раз меньше, чем в норме. В качестве нормировочного гена использовали ген *ras64B*. Эксперименты проводили в трех повторах. Планки погрешностей отмечают стандартную ошибку среднего. Звездочкой указаны статистически значимые изменения ( $p < 0.05$ ,  $t$ -критерий Стьюдента).

Ранее мы обнаружили, что инсуляторный белок Su(Hw), содержащий домен цинковых пальцев, взаимодействует с белком ENY2 (субъединица DUB-модуля) и привлекает ENY2-содержащие комплексы на Su(Hw)-зависимые инсуляторы дрозофилы, участвуя одновременно в регуляции транскрипции и в позиционировании ориджинов репликации [6, 7, 10]. Обнаружено также взаимодействие ENY2 с еще одним цинковым белком CG9890 [18]. Этот белок локализован преимущественно на промоторах активных генов и кокаллизуется с комплексами SAGA, ORC и dSWI/SNF на своих сайтах связывания. Показано, что белок CG9890 участвует в регуляции транскрипции генов, включая гены эдизонового каскада [19, 20]. В данной работе нами обнаружено взаимодействие DUB-модуля с еще одним белком CG9609, который содержит домен цинковых пальцев, как и Su(Hw). Мы предполагаем, что по аналогии с Su(Hw) белок CG9609 является ДНК-связывающим белком, который привлекает ENY2-содержащие комплексы на свои сайты связывания, организуя таким образом регуляторные элементы генома, необходимые для функционирования клетки.

Нами показано, что белок CG9609 принимает участие в регуляции транскрипции, зависимой от РНК-полимеразы II. После нокдауна гена *CG9609* количество мРНК пяти из девяти генов возросло статистически значимо (рис. 4). По всей видимости, на промоторах генов присутствует множество цинковых белков, которые стабилизируют транскрипционные комплексы, и нокдаун одного из них не всегда приводит к заметному влиянию на транскрипцию. Белок GTF3A человека – гомолог CG9609, является транскрипционным фактором РНК-полимеразы III, который принимает участие в транскрипции 5S рРНК человека. По-видимому, CG9609 может также принимать участие и в регуляции транскрипции, зависимой от РНК-полимеразы III.

Работа выполнена с использованием оборудования ЦКП ИБГ РАН и при поддержке Российского научного фонда (грант № 20-14-00269).

Данная работа выполнена без привлечения людей и животных в качестве объектов исследования.

Авторы заявляют об отсутствии конфликта интересов.

## СПИСОК ЛИТЕРАТУРЫ

- Orphanides G., Reinberg D. (2002) A unified theory of gene expression. *Cell*. **108**, 439–451.
- Maksimenko O., Georgiev P. (2014) Mechanisms and proteins involved in long-distance interactions. *Front. Genet.* **5**, 28.
- van Bommel J.G., Pagie L., Braunschweig U., Brugman W., Meuleman W., Kerkhoven R.M., van Steensel B. (2010) The insulator protein SU(HW) fine-tunes nuclear lamina interactions of the *Drosophila* genome. *PLoS One*. **5**, e15013.
- Rando O.J., Chang H.Y. (2009) Genome-wide views of chromatin structure. *Annu. Rev. Biochem.* **78**, 245–271.
- Tchurikov N.A., Krasnov A.N., Ponomarenko N.A., Golova Y.B., Chernov B.K. (1998) Forum domain in *Drosophila melanogaster* cut locus possesses looped domains inside. *Nucl. Acids Res.* **26**, 3221–3227.
- Vorobyeva N.E., Mazina M.U., Golovnin A.K., Kopytova D.V., Gurskiy D.Y., Nabirochkina E.N., Georgieva S.G., Georgiev P.G., Krasnov A.N. (2013) Insulator protein Su(Hw) recruits SAGA and Brahma complexes and constitutes part of Origin Recognition Complex-binding sites in the *Drosophila* genome. *Nucl. Acids Res.* **41**, 5717–5730.
- Мазина М.Ю., Воробьева Н.Е., Краснов А.Н. (2013) Способность Su(Hw) создавать платформу для формирования ориджинов репликации не зависит от типа окружающего хроматина. *Цитология*. **55**(4), 218–224.
- Vorobyeva N.E., Nikolenko J.V., Krasnov A.N., Kuzmina J.L., Panov V.V., Nabirochkina E.N., Georgieva S.G., Shidlovskii Y.V. (2011) SAYP interacts with DHR3 nuclear receptor and participates in ecdysone-dependent transcription regulation. *Cell Cycle*. **10**, 1821–1827.
- Kopytova D.V., Krasnov A.N., Orlova A.V., Gurskiy D.Y., Nabirochkina E.N., Georgieva S.G., Shidlovskii Y.V. (2010) ENY2: couple, triple...more? *Cell Cycle*. **9**, 479–481.
- Kurshakova M., Maksimenko O., Golovnin A., Pulina M., Georgieva S., Georgiev P., Krasnov A. (2007) Evolutionarily conserved E(y)2/Sus1 protein is essential for the barrier activity of Su(Hw)-dependent insulators in *Drosophila*. *Mol. Cell*. **27**, 332–338.
- Krasnov A.N., Kurshakova M.M., Ramensky V.E., Mardanov P.V., Nabirochkina E.N., Georgieva S.G. (2005) A retrocopy of a gene can functionally displace the source gene in evolution. *Nucl. Acids Res.* **33**, 6654–6661.
- Zhao Y., Lang G., Ito S., Bonnet J., Metzger E., Sawat-subashi S., Suzuki E., Le Guezennec X., Stunnenberg H.G., Krasnov A., Georgieva S.G., Schule R., Takeyama K., Kato S., Tora L., Devys D. (2008) A TFTC/STAGA module mediates histone H2A and H2B deubiquitination, coactivates nuclear receptors, and counteracts heterochromatin silencing. *Mol. Cell*. **29**, 92–101.
- Helmlinger D., Tora L. (2017) Sharing the SAGA. *Trends Biochem. Sci.* **42**, 850–861.
- Samara N.L., Datta A.B., Berndsen C.E., Zhang X., Yao T., Cohen R.E., Wolberger C. (2010) Structural insights into the assembly and function of the SAGA deubiquitinating module. *Science*. **328**, 1025–1029.
- Kohler A., Zimmermann E., Schneider M., Hurt E., Zheng N. (2010) Structural basis for assembly and activation of the heterotetrameric SAGA histone H2B deubiquitinase module. *Cell*. **141**, 606–617.
- Мазина М.Ю., Николенко Ю.В., Краснов А.Н., Воробьева Н.Е. (2016) Транскрипция гена *HSP70* дрозофилы на этапах инициации и элонгации происходит с участием белковых комплексов SWI/SNF. *Генетика*. **52**, 164–169.
- Krasnov A.N., Mazina M.Y., Nikolenko J.V., Vorobyeva N.E. (2016) On the way of revealing coactivator complexes cross-talk during transcriptional activation. *Cell. Biosci.* **6**, 15.
- Фурсова Н.А., Николенко Ю.В., Сошникова Н.В., Мазина М.Ю., Воробьева Н.Е., Краснов А.Н. (2018) Белок

- CG9890 с доменами цинковых пальцев — новый компонент ENY2-содержащих комплексов дрозофилы. *Acta Naturae*. **10**, 110–114.
19. Фурсова Н.А., Мазина М.Ю., Николенко Ю.В., Воробьева Н.Е., Краснов А.Н. (2020) Белок CG9890 дрозофилы, содержащий домены цинковых пальцев, локализуется с комплексами модификации и ремоделирования хроматина на промоторах генов и участвует в регуляции транскрипции. *Acta Naturae*. **12**, 114–119.
  20. Николенко Ю.В., Фурсова Н.А., Мазина М.Ю., Воробьева Н.Е., Краснов А.Н. (2022) Белок CG9890 дрозофилы участвует в регуляции эклизонзависимой транскрипции. *Молекуляр. биология*. **56**(4), 557–563.
  21. Clemens J.C., Worby C.A., Simonson-Leff N., Muda M., Maehama T., Hemmings B.A., Dixon J.E. (2000) Use of double-stranded RNA interference in *Drosophila* cell lines to dissect signal transduction pathways. *Proc. Natl. Acad. Sci. USA*. **97**, 6499–6503.
  22. Николенко Ю.В., Вдовина Ю.А., Фелелова Е.И., Глухова А.А., Набирочкина Е.Н., Копытова Д.В. (2021) Деубиквитирующий (DUB) модуль SAGA участвует в Pol III-зависимой транскрипции. *Молекуляр. биология*. **55**(3), 500–509.
  23. Enuameh M.S., Asriyan Y., Richards A., Christensen R.G., Hall V.L., Kazemian M., Zhu C., Pham H., Cheng Q., Blatti C., Brasfield J.A., Basciotta M.D., Ou J., McNulty J.C., Zhu L.J., Celniker S.E., Sinha S., Stormo G.D., Brodsky M.H., Wolfe S.A. (2013) Global analysis of *Drosophila* Cys(2)-His(2) zinc finger proteins reveals a multitude of novel recognition motifs and binding determinants. *Genome Res*. **23**, 928–940.
  24. Laity J.H., Dyson H.J., Wright P.E. (2000) DNA-induced alpha-helix capping in conserved linker sequences is a determinant of binding affinity in Cys(2)-His(2) zinc fingers. *J. Mol. Biol.* **295**, 719–727.

## The *Drosophila* Zinc Finger Protein CG9609 Interacts with the Deubiquitinating (DUB) Module of the SAGA Complex and Participates in the Regulation of Transcription

J. V. Nikolenko<sup>1</sup>, M. M. Kurshakova<sup>1</sup>, D. V. Kopytova<sup>1</sup>,  
Y. A. Vdovina<sup>1</sup>, N. E. Vorobyova<sup>2</sup>, A. N. Krasnov<sup>2, \*</sup>

<sup>1</sup>*Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Moscow, 119991 Russia*

<sup>2</sup>*Institute of Gene Biology, Russian Academy of Sciences, Moscow, 119334 Russia*

\*e-mail: krasnov@genebiology.ru

In previous studies, we found that the zinc finger proteins Su(Hw) and CG9890 interact with the *Drosophila* SAGA complex and participate in the formation of active chromatin structure and transcriptional regulation. In this research, we discovered the interaction of the DUB module of the SAGA complex with another zinc finger protein, CG9609. ChIP-Seq analysis was performed and CG9609 binding sites in the *Drosophila* genome were identified. Analysis of binding sites showed that they are localized predominantly at gene promoters. The CG9609 protein has been shown to be involved in the regulation of gene expression.

**Keywords:** CG9609, zinc fingers, transcription activation, DUB-module, SAGA complex

УДК 577.21

## БЕЛКИ AEF1 И CG10543 ДРОЗОФИЛЫ, СОДЕРЖАЩИЕ ДОМЕНЫ ЦИНКОВЫХ ПАЛЬЦЕВ, КОЛОКАЛИЗУЮТСЯ С КОМПЛЕКСАМИ SAGA, SWI/SNF И ORC НА ПРОМОТОРАХ ГЕНОВ И УЧАСТВУЮТ В РЕГУЛЯЦИИ ТРАНСКРИПЦИИ

© 2024 г. Ю. В. Николенко<sup>а</sup>, М. М. Куршакова<sup>а</sup>, Д. В. Копытова<sup>а</sup>, Ю. А. Вдовина<sup>а</sup>,  
Н. Е. Воробьева<sup>б</sup>, А. Н. Краснов<sup>б, \*</sup>

<sup>а</sup>Институт молекулярной биологии им. В. А. Энгельгардта Российской академии наук, Москва, 119991 Россия

<sup>б</sup>Институт биологии гена Российской академии наук, Москва, 119334 Россия

\*e-mail: krasnov@genebiology.ru

Поступила в редакцию 03.12.2023 г.

После доработки 11.12.2023 г.

Принята к публикации 12.12.2023 г.

Ряд белков с доменами типа цинковых пальцев, включая Aef1 и CG10543, взаимодействуют, как показано нами ранее, с DUB-модулем комплекса SAGA дрозофилы. В представленной работе проведено полногеномное исследование локализации белков Aef1 и CG10543 и показано, что они находятся преимущественно на промоторах активных генов. Сайты связывания этих белков колокализируются с комплексами модификации и ремоделирования хроматина SAGA и dSWI/SNF, а также с репликационным комплексом ORC. Показано, что белки Aef1 и CG10543 участвуют в регуляции экспрессии части генов, на промоторах которых они находятся. Таким образом, белки Aef1 и CG10543 являются новыми участниками транскрипционной сети клетки, которые колокализируются с основными транскрипционными и репликационными комплексами дрозофилы.

**Ключевые слова:** Aef1, CG10543, домены цинковых пальцев, регуляция транскрипции, DUB-модуль, комплекс SAGA, SWI/SNF, ORC

DOI: 10.31857/S0026898424040087, EDN: IMOGAY

### ВВЕДЕНИЕ

Высококонсервативный коактиваторный комплекс SAGA, основной функцией которого является модификация (ацетилирование и деубиквитинирование) гистонов хроматина, содержит более 20 белковых субъединиц [1]. Субъединицы комплекса SAGA взаимодействуют с различными активаторами транскрипции, привлекая таким образом комплекс на промоторы определенных генов [2, 3]. В действии комплексов, ремоделирующих и модифицирующих хроматин, существует определенный синергизм. Показано, что комплекс SAGA ацетилюет нуклеосомы на промоторах генов в процессе активации транскрипции. Это приводит к привлечению комплекса ремоделирования хроматина dSWI/SNF и стимулированию его ремоделирующей активности [4, 5].

Комплексы SAGA и SWI/SNF дрозофилы локализованы на различных регуляторных элементах генома, включая промоторы, где зачастую колокализируются с репликационным комплексом ORC [6]. Инициация репликации происходит на множестве сайтов в геноме, которые называются участками начала (ориджинами) репликации. На ориджины

репликации привлекается комплекс ORC, который состоит из шести субъединиц (ORC1–6). Этот комплекс участвует в создании платформы для связывания белков пререпликативного комплекса и в инициации репликации. Чтобы обеспечить правильный временной паттерн репликации генома в S-фазе, в клетках эукариот создается множество ориджинов репликации, определенная часть которых запускается в зависимости от стадии развития организма и условий роста. Комплекс ORC связывает определенные участки генома [7], однако субъединицы данного комплекса не проявляют явной специфичности к последовательностям ДНК. Поэтому встает вопрос, какие факторы определяют позиционирование ORC-комплексов в геноме.

В ходе предыдущих исследований нами обнаружено, что инсуляторный белок Su(Hw), содержащий домен цинковых пальцев, взаимодействует с белком ENY2 (субъединица DUB-модуля комплекса SAGA) и привлекает комплексы SAGA, SWI/SNF и ORC на Su(Hw)-зависимые инсуляторы дрозофилы, участвуя одновременно в регуляции транскрипции и в позиционировании ориджинов репликации [8–11]. Белок ENY2 является субъединицей деубиквитинирующего (DUB)

модуля комплекса SAGA, связанной с различными стадиями экспрессии генов [8–15]. Высказано предположение, что существуют и другие “цинковые белки”, которые взаимодействуют с комплексом SAGA дрозофилы и функционируют сходным образом на других регуляторных элементах генома, включая промоторы. В ходе дальнейших экспериментов был идентифицирован еще один подобный цинковый белок – CG9890 [16–18]. С целью поиска новых цинковых белков, взаимодействующих с комплексом SAGA дрозофилы, был очищен DUB-модуль комплекса SAGA [19]. С использованием масс-спектрометрического анализа показано, что в состав фракций, содержащих DUB-модуль комплекса SAGA, входят несколько десятков белков с доменами цинковых пальцев, из которых отобрали белки с кластером близкорасположенных цинковых пальцев и каноническим линкером (TGE[K/R]P) между ними, так как именно такая структура с максимальной вероятностью предсказывает ДНК-связывающие свойства белка со специфичным сайтом связывания [20, 21]. В результате дальнейшего отбора осталось 12 белков, взаимодействие которых было проверено с помощью реакции коиммунопреципитации. Подтверждено взаимодействие белка CG9609, а также обнаружено его слабое взаимодействие с белками Aef1 (Adult enhancer factor 1) и CG10543. Чтобы дополнительно подтвердить взаимодействие белков Aef1 и CG10543 с комплексом SAGA дрозофилы было решено провести ChIP-Seq-анализ этих белков и определить колокализацию данных факторов на полногеномном уровне.

## ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

**Культивирование клеток линии Schneider 2 (S2). РНК-интерференция.** Клетки культивировали при 25°C в среде Schneider's insect medium (“Sigma”, США), содержащей 10% фетальной сыворотки крупного рогатого скота (“HyClone”, США). Клетки трансфицировали с использованием Effectene Transfection Reagent (“Qiagen”, Германия) по протоколу производителя. Нокдаун генов проводили с помощью РНК-интерференции согласно опубликованному протоколу [22]. дцРНК для нокдауна генов *Aef1* и *CG10543* синтезировали с использованием следующих праймеров:

*Aef1* – GAATTAATACGACTCACTATAGGGAGAA TGATGCATATCAAAAGCCT и GAATTAATACGACTCACTATAGGGAGATCCGGGATGCTCGCTA TGT;  
*CG10543* – GAATTAATACGACTCACTATAGGGA GATTGATGAGCTTCCAGCATAT и GAATTAATA CGACTCACTATAGGGAGAATCGTTCTTCACTG GGTAGTA.

В качестве контроля использовали дцРНК, соответствующую фрагменту плазмиды pBluescript II

SK(–) (“Stratagene”, США) длиной 500 п.н. *pBluescriptIIISK(–)*, GAATTAATACGACTCACTATA GGGAGAGTTACATGATCCCCCATG и GAATTA ATACGACTCACTATAGGGAGATTTCCGCCCCGA AGAACG.

В каждый эксперимент по РНК-интерференции дцРНК брали из расчета 30 мкг на 1 млн клеток. По прошествии 5 дней инкубации выделяли РНК и измеряли уровни мРНК исследуемых генов методом обратной транскрипции с последующей ПЦР в реальном времени. Использовали следующие праймеры:

*sgl* – CGTGATAACTCTGGTGGACAAG и GTCTCGATGTCCGTTGAGAAG;  
*D19B* – TTGCCGAAAAGGCTTTACCAC и CACTCCGTGCAGGGGAAGTC;  
*Kek1* – ACCCTGCTGCCTGGAATGAT и GCATGTAGCCATCCTCAGTCATC;  
*wg* – ACTGCTCGACGAGAACTTCTC и TGTAGTCGCAGGTGCAGGAC;  
*Psc* – CGCTGACTCAAACTCAAGTGTG и TCGATTCTGGCGTCATCATAGTA;  
*CG1907* – CCGCTAGACTTGGTCAAGACC и TGTCTCAGCAGAGCAGCTCC;  
*CG34136* – TGCCTTGCATTCTTTCGCTTG и GTCGTGGCGCTGTAGGGACT;  
*Ras64B* – GAGGGATTCCCTGCTCGTCTTCG и GTCGCACTTGTTACCCACCATC.

**Получение поликлональных антител к белкам Aef1 и CG10543.** Поликлональные антитела α-Aef1 и α-CG10543 получали из сыворотки крови крыс, иммунизированных фрагментами соответствующих белков. Фрагменты белков нарабатывали в клетках *Escherichia coli*: белок Aef1 с использованием N-концевого фрагмента длиной 162 аминокислотных остатка, а белок CG10543 – C-концевого фрагмента из 136 аминокислотных остатков. Эти участки были выбраны, чтобы исключить домены цинковых пальцев из фрагментов, которые будут использоваться для иммунизации животных. Домены цинковых пальцев при получении антител к цинковым белкам не использовали, так как это может дать перекрестную реакцию с другими цинковыми белками. Антитела очищали на колонках, которые содержали соответствующие рекомбинантные белки.

**Идентификация сайтов связывания белков Aef1 и CG10543 в геноме дрозофилы.** Сайты связывания белков Aef1 и CG10543 идентифицировали, как описано ранее. Идентифицировано примерно по 3500 сайтов связывания каждого из белков.

**Анализ распределения сайтов связывания белков Aef1 и CG10543 в геноме.** Геном разбивали на точки начала транскрипции (TSS), точки терминации транскрипции (TES), транскрибируемые области (участки генов за исключением TSS и TES) и межгенные области (все остальное). Областью TSS считали интервал ±500 п.н. от старта транскрипции.

Областью TES считали интервал  $\pm 500$  п.н. от точки терминации транскрипции. Данные области определяли из аннотации транскриптов генома дрозофилы dmel\_r6.40. Считали, что сайт связывания находится в данной области, если пик сайта связывания попадает в заданный интервал.

**Анализ привлечения комплексов SAGA, SWI/SNF, ORC на сайты связывания белков Aef1 и CG10543.** В анализе использовали ChIP-Seq профили белков GCN5 (комплекс SAGA), OSA (комплекс SWI/SNF), ORC2 (комплекс ORC), полученные нами в предыдущих работах. Был посчитан усредненный профиль каждого из этих белков на четырех группах сайтов: на сайтах связывания белков Aef1 и CG10543, на случайных промоторах и на случайных местах генома. Так как для каждого из белков Aef1 и CG10543 было идентифицировано около 3500 сайтов, мы отобрали такое же число случайных промоторов и случайных мест генома с помощью генератора псевдослучайных чисел. Для получения усредненного ChIP-Seq профиля белков GCN5, OSA и ORC2 на какой-либо группе сайтов сначала были получены локальные профили каждого из 3500 сайтов в группе в пределах  $\pm 5000$  п.н. от сайта. Затем эти 3500 локальных профилей усредняли по каждой нуклеотидной позиции для получения среднего профиля. Таким образом получено 12 усредненных профилей для трех белков – GCN5, OSA и ORC2 – на четырех группах сайтов.

Результаты ChIP-Seq-анализа размещены в базе данных Gene Expression Omnibus (идентификатор эксперимента GSE250183). Кроме самих данных, в этой базе приведены также протоколы иммунопреципитации хроматина, создания библиотек и обработки данных.

ChIP-Seq профили получали с помощью программы bamCoverage 3.5.4. Значения в профиле представляют собой нормированное

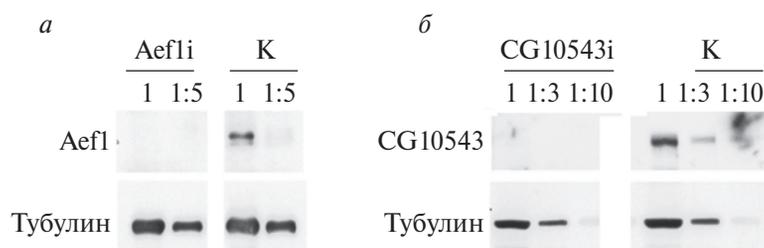
(CPM – Counts Per Million mapped reads) значение ДНК-фрагментов в каждой геномной позиции.

## РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

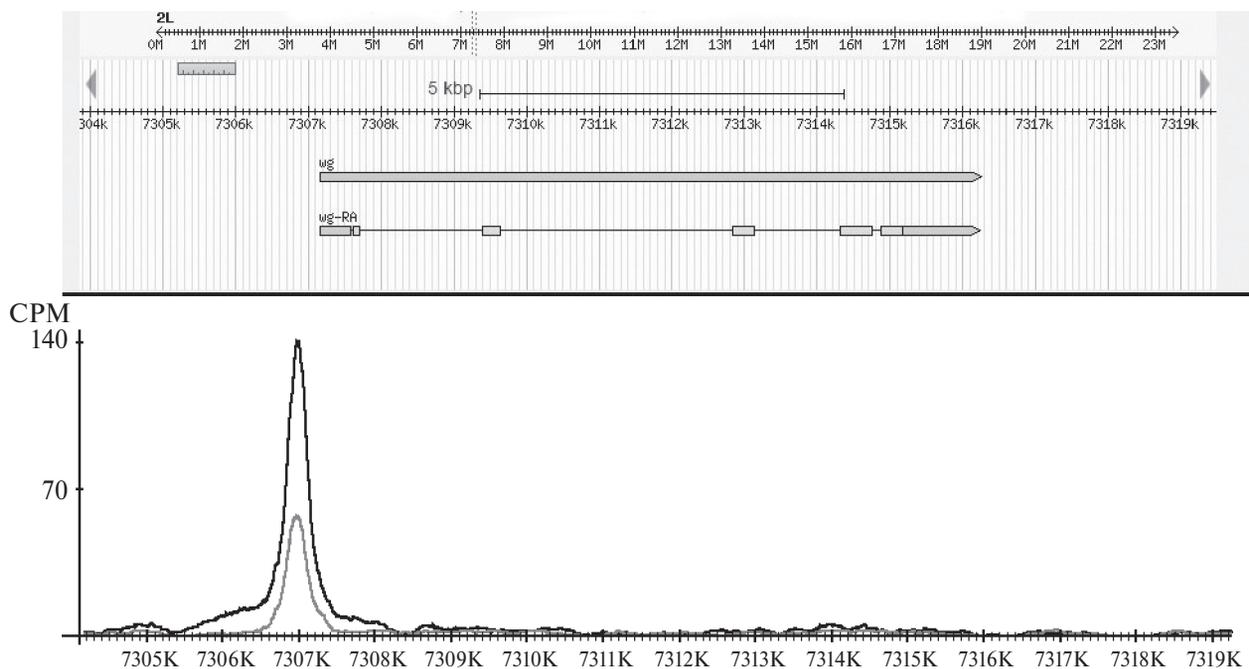
### *Белки Aef1 и CG10543 локализованы преимущественно на промоторах генов*

С целью идентификации сайтов локализации белков Aef1 и CG10543 в геноме проведен ChIP-Seq-анализ. Поликлональные антитела к белкам Aef1 и CG10543 получены и аффинно очищены на колонках, содержащих соответствующие рекомбинантные белки. Специфичность антител проверена с помощью Вестерн-блоттинга. На полиакриламидный гель наносили белковые экстракты, выделенные как из нормальных клеток, так и подвергнутых РНК-интерференции (рис. 1). Вестерн-блот инкубировали с проверяемыми антителами и антителами к  $\alpha$ -тубулину для контроля белковой нагрузки. На рис. 1 видно, что антитела узнают полосы, интенсивность которых значительно уменьшилась после РНК-интерференции соответствующих белков. При этом полосы  $\alpha$ -тубулина не изменились.

Затем была проведена иммунопреципитация хроматина из культуры клеток дрозофилы S2 с использованием антител к CG10543 и Aef1 в двух повторях и получены библиотеки для высокопроизводительного секвенирования, которое проводили на геномном секвенаторе Illumina Novaseq 6000. Полученные последовательности выравнивали на геном с помощью программы Hisat2. В дальнейший анализ брали только уникально картированные риды. Потенциальные места связывания идентифицировали с помощью программы Macs2. Характерный ChIP-seq профиль белков Aef1 и CG10543 представлен на рис. 2. В качестве примера приведен ген *wg*, на промоторе которого находятся оба



**Рис. 1.** Проверка специфичности антител к белку Aef1 (a) и CG10543 (б) с помощью Вестерн-блот-анализа. На панели a нанесен белковый экстракт из контрольных клеток (К) в двух разведениях (1 и 1:5), а также белковый экстракт из клеток после РНК-интерференции *Aef1* (Aef1i), также взятый в двух разведениях. Вестерн-блот (панель a) инкубировали с проверяемыми антителами к белку Aef1, а также с антителами к  $\alpha$ -тубулину для контроля нанесения. На панели б каждый из образцов нанесен в трех разведениях (1, 1:3, 1:10). К – экстракт из контрольных клеток, CG10543i – экстракт из клеток после РНК-интерференции *CG10543*. Вестерн-блот (панель б) инкубировали с антителами к белку CG10543, а также с антителами к  $\alpha$ -тубулину для контроля нанесения.

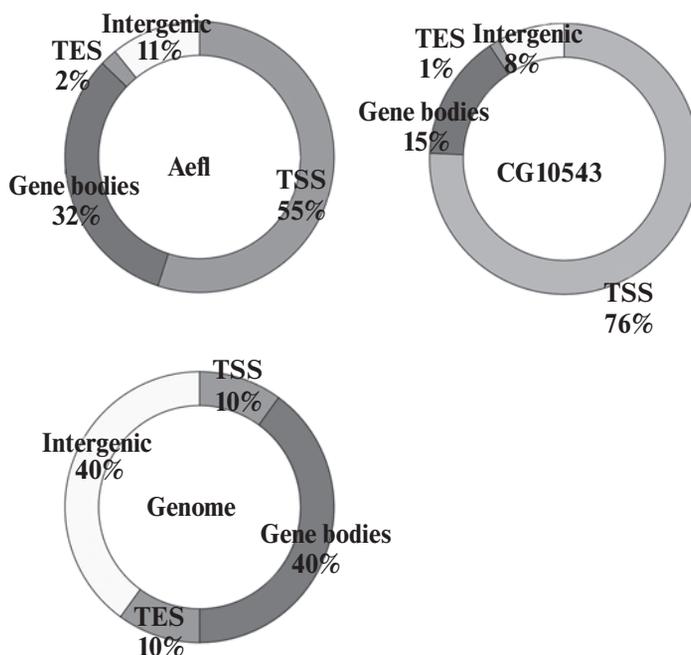


**Рис. 2.** ChIP-Seq профили белков CG10543 (черный) и Aef1 (серый) в области гена *wg*. На верхней панели представлена структура гена из геномного браузера. На нижней панели представлены ChIP-Seq профили в единицах CPM.

белка. Идентифицировано примерно по 3500 сайтов связывания каждого белка (FDR <5%).

Мы аннотировали полученные сайты по признаку их локализации в одном из следующих элементов генома дрозофилы: промоторы генов, 3'-концы генов, тела генов и межгенные области.

Согласно полученным данным (рис. 3), наибольшее количество ChIP-Seq-пиков белков Aef1 и CG10543 представлено в промоторных областях генов дрозофилы, 55 и 76% соответственно. Мы предполагаем, что, локализуясь преимущественно на промоторах генов, данные белки могут участвовать



**Рис. 3.** Распределение сайтов связывания белков Aef1 (верхнее левое изображение) и CG10543 (верхнее правое изображение) относительно аннотированных элементов генома дрозофилы. Для сравнения показана относительная представленность всех аннотированных элементов в геноме (Genome) (нижнее изображение). TSS-промоторная область, TES-конец гена, Gene bodies – область гена между TSS и TES, Intergenic – межгенные области.



Рис. 4. Консенсусная последовательность потенциального сайта связывания белков CG10543 (а) и Aef1 (б).

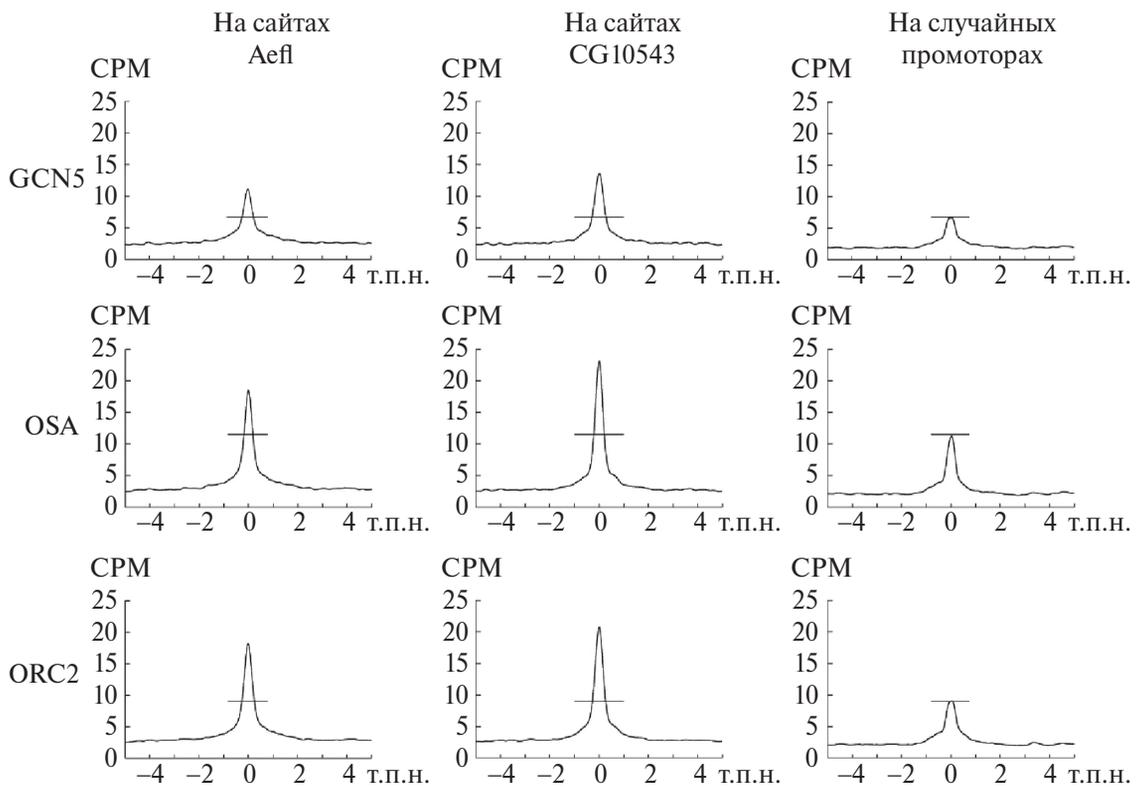
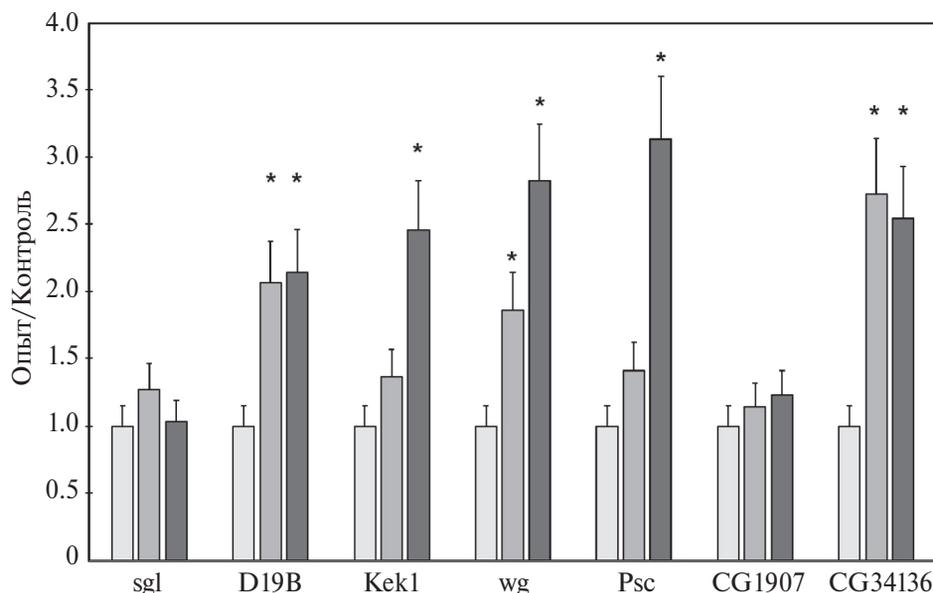


Рис. 5. Усредненные ChIP-Seq-профили белков ORC2 (комплекс ORC), OSA (комплекс SWI/SNF) и GCN5 (комплекс SAGA) на сайтах связывания белков Aef1, CG10543 и на случайных промоторах. Слева указаны субъединицы данных комплексов: ORC2 (ORC), OSA (SWI/SNF), GCN5 (SAGA). Представлен средний ChIP-Seq-профиль белков (указаны слева) на трех группах сайтов (указаны сверху). По оси ординат профили представлены в единицах CPM, по оси абсцисс указано расстояние в т.п.н. относительно центра пика. В каждой из групп анализировали примерно по 3500 сайтов. Горизонтальными отрезками отмечен уровень сигнала на случайных промоторах. Видно, что субъединицы комплексов SAGA, SWI/SNF и ORC привлекаются на сайты Aef1 и CG10543 в большем количестве, чем на случайные промоторы.



**Рис. 6.** Транскрипция некоторых генов, содержащих оба белка, CG10543 и Aef1, на своих промоторах, при нокадауне CG10543 и Aef1 в индивидуальных экспериментах. Названия генов указаны внизу. Светло-серые столбцы соответствуют транскрипции генов при РНК-интерференции Aef1, темно-серые столбцы – при РНК-интерференции CG10543. Белые столбцы соответствуют транскрипции гена в норме (контроль), принято за единицу. По оси ординат указано соотношение транскрипции в опыте и в контроле. В качестве нормировочного гена использовали ген *ras64B*. Эксперименты проводили в трех повторах. Планки погрешностей отмечают стандартную ошибку среднего. Звездочкой указаны статистически значимые изменения с  $p < 0.05$  (использован *t*-критерий Стьюдента).

в функционировании регуляторных генетических элементов этого типа.

Далее были определены потенциальные консенсусные последовательности связывания белков Aef1 и CG10543. Для этого получили нуклеотидные последовательности длиной 500 п.н. вокруг каждого ChIP-Seq-пика с помощью программы bedtools, а затем проанализировали их в программе MEME-ChIP. Были определены консенсусные последовательности, которыми обогащены фрагменты ChIP-Seq, представленные на рис. 4. Эти консенсусы имеют высокую статистическую значимость, центральное обогащение во фрагментах ChIP-Seq и не являются сайтами связывания какого-либо известного белка. Таким образом, их можно рассматривать как потенциальные сайты связывания белков Aef1 и CG10543.

*Белки CG10543 и Aef1 колокализуются с комплексами SAGA, SWI/SNF и ORC на своих сайтах связывания*

Геномная колокализация белков Aef1 и CG10543 с комплексами SAGA, SWI/SNF и ORC изучена нами с использованием программного обеспечения собственной разработки для построения усредненного профиля исследуемого фактора на заданных сайтах генома (см. раздел Экспериментальная часть) [8]. ChIP-Seq-профили белков ORC2, GCN5, OSA получены в нашей лаборатории ранее. Рассчитаны усредненные профили данных

белков на четырех группах сайтов: на сайтах связывания белка Aef1, на сайтах связывания белка CG10543, на случайных промоторах и на случайных участках генома (рис. 5). Изображение усредненных профилей на случайных сайтах генома не приведено. Эти профили представляют собой прямые линии на уровне фона (примерно 2).

Поскольку белки Aef1 и CG10543 локализованы преимущественно на промоторах генов, на их сайтах связывания можно будет найти обогащение любых промоторных факторов, включая комплексы SAGA, dSWI/SNF и ORC, относительно среднегеномного уровня. Однако, как можно видеть из рис. 5, белки GCN5 (комплекс SAGA), OSA (комплекс dSWI/SNF) и ORC2 (комплекс ORC) обогащены на сайтах связывания белков Aef1 и CG10543 не только по сравнению со среднегеномным уровнем (примерно 2), но и по сравнению со случайными промоторами. Полученный результат свидетельствует, что причиной данной колокализации является не просто случайное совпадение на промоторах генов, а то, что именно сайт связывания цинковых белков Aef1 и CG10543 способствует локализации комплексов SAGA, dSWI/SNF и ORC.

*Белки CG10543 и Aef1 принимают участие в регуляции экспрессии генов*

Учитывая, что исследуемые белки найдены преимущественно на промоторах генов, мы решили с помощью метода РНК-интерференции определить, к каким изменениям в экспрессии генов

приведет снижение внутриклеточного уровня этих белков. В эксперимент взяли гены, на промоторах которых присутствуют оба белка. Используя метод ОТ-кПЦР, мы проанализировали количество РНК семи генов в образцах, выделенных из контрольных клеток и клеток после РНК-интерференции. Результаты этого эксперимента представлены на рис. 6. После нокдауна *CG10543* и *Aef1* количество мРНК части генов возросло. Таким образом, эти белки действительно могут участвовать в регуляции экспрессии тех генов, на промоторах которых они локализуются, выступая в роли репрессоров.

В ходе предыдущих исследований мы обнаружили, что инсуляторный белок Su(Hw) дрозофилы, содержащий домен цинковых пальцев, взаимодействует с белком ENY2 (субъединица DUB-модуля комплекса SAGA) и привлекает комплексы SAGA, SWI/SNF и ORC на Su(Hw)-зависимые инсуляторы, участвуя одновременно в регуляции транскрипции и в позиционировании ориджинов репликации [8–10]. Предполагалось, что существуют и другие цинковые белки, которые взаимодействуют с комплексом SAGA дрозофилы и функционируют сходным образом на других регуляторных элементах генома, включая промоторы. В дальнейшем были идентифицированы еще два цинковых белка CG9890 [16–18] и CG9609. В данной работе показана полногеномная колокализация еще двух белков – *Aef1* и *CG10543* – с комплексами SAGA, SWI/SNF и ORC. Все четыре белка, CG9890, CG9609, *Aef1* и *CG10543*, содержат домены цинковых пальцев, локализованы на промоторах активных генов, колокализуются с основными транскрипционными и репликационными комплексами и участвуют в регуляции транскрипции. Мы предполагаем, что подобно Su(Hw) эти четыре белка могут участвовать в привлечении комплексов SAGA, SWI/SNF и ORC на свои сайты связывания, организуя, таким образом, регуляторные элементы генома, необходимые для функционирования клетки. Нокдаун генов *Aef1* и *CG10543* приводит к изменению транскрипции некоторых из проверенных генов (рис. 6). По всей видимости, на промоторах генов присутствует множество цинковых белков, которые стабилизируют транскрипционные комплексы, и нокдаун одного из них не всегда приводит к заметному влиянию на транскрипцию.

Работа выполнена с использованием оборудования ЦКП ИБГ РАН и при поддержке Российского научного фонда (грант № 20-14-00269).

Данная работа выполнена без привлечения людей и животных в качестве объектов исследования.

Авторы заявляют об отсутствии конфликта интересов.

## СПИСОК ЛИТЕРАТУРЫ

1. Koutelou E., Hirsch C.L., Dent S.Y. (2010) Multiple faces of the SAGA complex. *Curr. Opin. Cell. Biol.* **22**, 374–382.

2. Baker S.P., Grant P.A. (2007) The SAGA continues: expanding the cellular role of a transcriptional co-activator complex. *Oncogene*. **26**, 5329–5340.
3. Brown C.E., Howe L., Sousa K., Alley S.C., Carrozza M.J., Tan S., Workman J.L. (2001) Recruitment of HAT complexes by direct activator interactions with the ATM-related Tra1 subunit. *Science*. **292**, 2333–2337.
4. Chatterjee N., Sinha D., Lemma-Dechassa M., Tan S., Shogren-Knaak M.A., Bartholomew B. (2011) Histone H3 tail acetylation modulates ATP-dependent remodeling through multiple mechanisms. *Nucl. Acids Res.* **39**, 8378–8391.
5. Li B., Carey M., Workman J.L. (2007) The role of chromatin during transcription. *Cell*. **128**, 707–719.
6. MacAlpine D.M., Rodriguez H.K., Bell S.P. (2004) Coordination of replication and transcription along a *Drosophila* chromosome. *Genes Dev.* **18**, 3094–3105.
7. Eaton M.L., Prinz J.A., MacAlpine H.K., Tretyakov G., Kharchenko P.V., MacAlpine D.M. (2011) Chromatin signatures of the *Drosophila* replication program. *Genome Res.* **21**, 164–174.
8. Vorobyeva N.E., Mazina M.U., Golovnin A.K., Kopytova D.V., Gurskiy D.Y., Nabirochkina E.N., Georgieva S.G., Georgiev P.G., Krasnov A.N. (2013) Insulator protein Su(Hw) recruits SAGA and Brahma complexes and constitutes part of Origin Recognition Complex-binding sites in the *Drosophila* genome. *Nucl. Acids Res.* **41**, 5717–5730.
9. Мазина М.Ю., Воробьева Н.Е., Краснов А.Н. (2013) Способность Su(Hw) создавать платформу для формирования ориджинов репликации не зависит от типа окружающего хроматина. *Цитология*. **55**(4), 218–224.
10. Kurshakova M., Maksimenko O., Golovnin A., Pulina M., Georgieva S., Georgiev P., Krasnov A. (2007) Evolutionarily conserved E(y)2/Sus1 protein is essential for the barrier activity of Su(Hw)-dependent insulators in *Drosophila*. *Mol. Cell*. **27**, 332–338.
11. Vorobyeva N.E., Erokhin M., Chetverina D., Krasnov A.N., Mazina M.Y. (2021) Su(Hw) primes 66D and 7F *Drosophila* chorion genes loci for amplification through chromatin decondensation. *Sci. Rep.* **11**, 16963.
12. Vorobyeva N.E., Nikolenko J.V., Krasnov A.N., Kuzmina J.L., Panov V.V., Nabirochkina E.N., Georgieva S.G., Shidlovskii Y.V. (2011) SAYP interacts with DHR3 nuclear receptor and participates in ecdysone-dependent transcription regulation. *Cell Cycle*. **10**, 1821–1827.
13. Kopytova D.V., Krasnov A.N., Orlova A.V., Gurskiy D.Y., Nabirochkina E.N., Georgieva S.G., Shidlovskii Y.V. (2010) ENY2: couple, triple...more? *Cell Cycle*. **9**, 479–481.
14. Krasnov A.N., Kurshakova M.M., Ramensky V.E., Mardanov P.V., Nabirochkina E.N., Georgieva S.G. (2005) A retrocopy of a gene can functionally displace the source gene in evolution. *Nucl. Acids Res.* **33**, 6654–6661.
15. Gurskiy D., Orlova A., Vorobyeva N., Nabirochkina E., Krasnov A., Shidlovskii Y., Georgieva S., Kopytova D. (2012) The DUBm subunit Sgf11 is required for mRNA export and interacts with Cbp80 in *Drosophila*. *Nucl. Acids Res.* **40**, 10689–10700.
16. Николенко Ю.В., Фурсова Н.А., Мазина М.Ю., Воробьева Н.Е., Краснов А.Н. (2022) Белок CG9890 дрозофилы участвует в регуляции экдизонзависимой транскрипции. *Молекуляр. биология*. **56**(4), 557–563.

17. Фурсова Н.А., Мазина М.Ю., Николенко Ю.В., Воробьева Н.Е., Краснов А.Н. (2020) Белок CG9890 дрозофилы, содержащий домены цинковых пальцев, локализуется с комплексами модификации и ремоделирования хроматина на промоторах генов и участвует в регуляции транскрипции. *Acta Naturae*. **12**, 114–119.
18. Фурсова Н.А., Николенко Ю.В., Сошникова Н.В., Мазина М.Ю., Воробьева Н.Е., Краснов А.Н. (2018) Белок CG9890 с доменами цинковых пальцев – новый компонент ENY2-содержащих комплексов дрозофилы. *Acta Naturae*. **10**, 110–114.
19. Николенко Ю.В., Вдовина Ю.А., Фефелова Е.И., Глухова А.А., Набирочкина Е.Н., Копытова Д.В. (2021) Деубиквитирующий (DUB) модуль SAGA участвует в Pol III-зависимой транскрипции. *Молекуляр. биология*. **55**(3), 500–509.
20. Enuameh M.S., Asriyan Y., Richards A., Christensen R.G., Hall V.L., Kazemian M., Zhu C., Pham H., Cheng Q., Blatti C., Brasefield J.A., Basciotta M.D., Ou J., McNulty J.C., Zhu L.J., Celniker S.E., Sinha S., Stormo G.D., Brodsky M.H., Wolfe S.A. (2013) Global analysis of *Drosophila* Cys(2)-His(2) zinc finger proteins reveals a multitude of novel recognition motifs and binding determinants. *Genome Res*. **23**, 928–940.
21. Laity J.H., Dyson H.J., Wright P.E. (2000) DNA-induced alpha-helix capping in conserved linker sequences is a determinant of binding affinity in Cys(2)-His(2) zinc fingers. *J. Mol. Biol.* **295**, 719–727.
22. Clemens J.C., Worby C.A., Simonson-Leff N., Muda M., Maehama T., Hemmings B.A., Dixon J.E. (2000) Use of double-stranded RNA interference in *Drosophila* cell lines to dissect signal transduction pathways. *Proc. Natl. Acad. Sci. USA*. **97**, 6499–6503.

## The *Drosophila* Zinc Finger Proteins Aef1 and Cg10543 Are Co-Localized with SAGA, SWI/SNF and ORC Complexes on Gene Promoters and Involved in Transcription Regulation

J. V. Nikolenko<sup>1</sup>, M. M. Kurshakova<sup>1</sup>, D. V. Kopytova<sup>1</sup>,  
Y. A. Vdovina<sup>1</sup>, N. E. Vorobyova<sup>2</sup>, A. N. Krasnov<sup>2, \*</sup>

<sup>1</sup>Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Moscow, 119991 Russia

<sup>2</sup>Institute of Gene Biology, Russian Academy of Sciences, Moscow, 119334 Russia

\*e-mail: krasnov@genebiology.ru

In previous studies, we purified the DUB-module of the *Drosophila* SAGA complex and showed that a number of zinc proteins interact with it, including Aef1 and CG10543. In this work, we conducted a genome-wide study of the Aef1 and CG10543 proteins and showed that they are localized predominantly on the promoters of active genes. The binding sites of these proteins colocalize with the SAGA and dSWI/SNF chromatin modification and remodeling complexes, as well as with the ORC replication complex. It has been shown that the Aef1 and CG10543 proteins are involved in the regulation of the expression of some genes on the promoters of which they are located. Thus, the Aef1 and CG10543 proteins are new participants in the cell transcriptional network and colocalize with the main transcription and replication complexes of *Drosophila*.

**Keywords:** Aef1, CG10543, zinc fingers domains, transcription regulation, DUB-module, complex SAGA, SWI/SNF, ORC

УДК 577.21

## РЕГУЛЯЦИЯ ТРАНСЛЯЦИИ ФАКТОРА eRF1 ЧЕЛОВЕКА

© 2024 г. А. В. Шувалов<sup>а, б</sup>, А. А. Клишин<sup>а, с</sup>, Н. С. Бизяев<sup>а</sup>,  
Е. Ю. Шувалова<sup>а, б</sup>, Е. З. Алкалаева<sup>а, б, с, \*</sup><sup>а</sup>Институт молекулярной биологии им. В. А. Энгельгардта Российской академии наук, Москва, 119991 Россия<sup>б</sup>Центр высокоточного редактирования и генетических технологий для биомедицины, Москва, 119991 Россия<sup>с</sup>Московский физико-технический институт (национальный исследовательский университет),  
Долгопрудный, Московская обл., 141700 Россия

\*e-mail: alkalaeva@eimb.ru

Поступила в редакцию 10.11.2023 г.

После доработки 11.12.2023 г.

Принята к публикации 13.12.2023 г.

Фактор терминации трансляции эукариот eRF1 – важный клеточный белок, играющий ключевую роль в процессах терминации трансляции, нонсенс-опосредованного распада мРНК (NMD) и сквозного прочтения стоп-кодонов. Количество eRF1 в клетке влияет на все эти процессы. Механизм регуляции трансляции eRF1 через авторегуляторный NMD-зависимый контур экспрессии описан у растений и грибов, тогда как механизмы регуляции трансляции eRF1 человека до сих пор не изучены. С помощью репортерных конструкций нами изучено влияние элементов мРНК eRF1 на его трансляцию в бесклеточных системах трансляции и культуре клеток HEK293. Наши данные указывают на отсутствие NMD-зависимого авторегуляторного контура экспрессии eRF1 человека. Обнаружено, что на трансляцию основной рамки считывания eRF1 сильнее всего влияют 5'-нетранслируемая область мРНК этого фактора и старт-кодон короткой открытой рамки считывания. Согласно базе данных стартов транскрипции, для мРНК eRF1 характерна высокая гетерогенность начала транскрипции и варьирующая из-за этого длина 5'-нетранслируемой области. Кроме того, старт-кодон основной рамки считывания в мРНК eRF1 находится в пределах известного регулятора трансляции короткой 5'-нетранслируемой области (TISU), который также стимулирует транскрипцию мРНК генов с высокой гетерогенностью старта транскрипции. Мы предполагаем, что регуляция синтеза eRF1 человека осуществляется как на уровне транскрипции, так и на уровне трансляции. На уровне транскрипции регулируется длина 5'-нетранслируемой области eRF1 и количество коротких открытых рамок считывания в ней, регулирующих, в свою очередь, продукцию eRF1 на уровне трансляции.

**Ключевые слова:** eRF1, трансляция, мРНК, 3'-нетранслируемая область мРНК, 5'-нетранслируемая область мРНК, короткая открытая рамка считывания

DOI: 10.31857/S0026898424040091, EDN: IMNJKL

## ВВЕДЕНИЕ

Фактор терминации трансляции eRF1 – основной фактор, обеспечивающий терминацию трансляции в клетках эукариот. eRF1 состоит из трех функциональных доменов: N-домен вовлечен в распознавание стоп-кодона [1–6], M-домен обеспечивает гидролиз пептидил-тРНК в Р-сайте рибосомы [7], а С-домен связывается с дополнительным фактором терминации трансляции

eRF3 [8–13]. Кроме того, eRF1 вовлечен в регуляцию сквозного прочтения [14–16], в нонсенс-опосредованный распад (NMD) мРНК [17–20], преждевременную терминацию на некоторых редких смысловых кодонах [21, 22].

Терминация – это важный этап трансляции, на котором регулируется скорость и эффективность высвобождения новосинтезированных пептидов. Показано, что терминация трансляции на kORC в 5'-нетранслируемой области (5'-НТО) мРНК ряда генов обеспечивает регуляцию трансляции основного продукта [23]. Изучение терминации трансляции имеет и прикладное значение. До 10% орфанных генетических заболеваний человека связаны с возникновением преждевременных стоп-кодонов (ПСК), а терапия таких заболеваний направлена на увеличение сквозного прочтения ПСК [24].

Сокращения: NMD – нонсенс-опосредованный распад; 5'- и 3'-НТО – нетранслируемые области мРНК; ОРС – открытая рамка считывания; kОРС – короткая открытая рамка считывания; EJC – комплекс срачивания экзонов; Fluc (Fl) – люцифераза светлячка; Nluc (Nl) – нанолуцифераза; HbV – β-глобин; *ETF1* – символ гена, кодирующего eRF1; GARDH – глицеральдегидфосфат-дегидрогеназа; ПСК – преждевременный стоп-кодон; ССТ – сайты старта транскрипции; О.Е.Л. – относительные единицы люминесценции.

Сквозное прочтение стоп-кодонов и преждевременная терминация являются противоположными процессами, зависящими от концентрации eRF1 и конкурирующих tРНК [21, 25]. При сквозном прочтении происходит удлинение белкового продукта с С-конца за счет распознавания стоп-кодона конкурирующей близкородственной tРНК, что в некоторых случаях может быть частью нормальных физиологических процессов клетки [26–28]. Менее изученный процесс преждевременной терминации на смысловых кодонах описан у прокариот, грибов и дрожифил, обсуждается его более широкая представленность у эукариот [21, 22, 29]. Предполагается, что физиологический смысл преждевременной терминации на смысловых кодонах заключается в высвобождении “застывших” на мРНК рибосом, и этот механизм является альтернативой No-go-деградации мРНК (NGD).

NMD – процесс деградации мРНК, обусловленный наличием ПСК, может стимулировать комплекс сращивания экзонов (EJC – Exon Junction Complex) на определенном расстоянии после стоп-кодона. Основная функция NMD заключается в контроле качества и удалении ошибочно возникающих мРНК с ПСК в результате нарушений сплайсинга. Кроме того, имеется множество свидетельств участия NMD в регуляции нормальной экспрессии генов. мРНК могут подвергаться NMD и в отсутствие ошибочного соединения экзонов и появления ПСК. Такие события могут обуславливаться либо EJC-независимым NMD, либо неканоническим расположением EJC. Вероятность NMD при этом тем выше, чем длиннее 3'-НТО [19].

Таким образом, исследование продукции eRF1 имеет важное значение, так как уровень этого белка в клетке будет определять эффективность важных клеточных процессов, описанных выше. Экспрессия гена *ETF1*, кодирующего eRF1 человека, изучена на уровне транскрипции [30], однако отсутствуют работы о регуляции трансляции eRF1 у млекопитающих. Стоит отметить, что описан авторегуляторный контур экспрессии eRF1 у растений и грибов. При высоком содержании eRF1 запускается механизм распада собственной мРНК, обусловленный наличием EJC после первого стоп-кодона, компетентного для стимулирования NMD. Из-за этого снижается количество eRF1 в клетке, происходит сквозное прочтение стоп-кодона собственной мРНК, что ингибирует ее NMD [31, 32].

В представленной работе с помощью конструкций мРНК, включающих кодирующие последовательности репортеров люциферазы светлячка (Fluc – Firefly luciferase) или нанолуциферазы (Nluc – Nano luciferase), нами изучено влияние элементов мРНК eRF1 (5'-НТО, ОРС, 3'-НТО) на трансляцию. Показано, что основной вклад в регуляцию трансляции вносит 5'-НТО мРНК eRF1

и находящиеся в ней дополнительные старт-кодона. Наиболее вероятным механизмом контроля трансляции eRF1 является регуляция длины ее 5'-НТО за счет частоты использования сайтов инициации транскрипции.

## ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

**Клонирование 3'-НТО мРНК eRF1.** Тотальную РНК HEK293 выделяли с помощью ExtractRNA (“Evrogen”, ВС032, Россия) по протоколу производителя. Эту РНК использовали для получения кДНК с помощью AMV Reverse Transcription System (“Promega”, А3500, США) с олиго(dT)<sub>15</sub> или случайными праймерами (табл. S1). С помощью праймеров CDS F и олиго(dT)<sub>15</sub> (табл. S1, дополнительные материалы размещены в электронном виде по DOI: 10.31857/S0026898424040091 статьи, а также размещены на сайте [http://www.molecbio.ru/downloads/2024/4/supp\\_Shupalov\\_rus.pdf](http://www.molecbio.ru/downloads/2024/4/supp_Shupalov_rus.pdf)) был получен ПЦР-продукт, который клонировали в ТА-вектор pGEM®-T Easy Vector (“Promega”, А1360). Клонированный продукт содержал часть ОРС eRF1 и последовательность 3'-НТО мРНК eRF1, включающую 861 нуклеотид после стоп-кодона UAG ОРС eRF1 и далее А(15).

**Создание экспериментальных конструкций и синтез мРНК. Конструкции NI с ПСК и контроли с кодирующими кодонами.** Плазмида pNI-gl описана ранее [33]. В конструкцию pNI-gl по NcoI-сайту рестрикции вставили частичную последовательность ОРС β-глобина с ПСК UGA и контекстом CUAGUA (слабый), усиливающим сквозное прочтение [34, 35]. Старт-кодон в последовательности гена β-глобина заменили на АТА, используя набор QuikChange Site-Directed Mutagenesis Kit (“Agilent”, США), и получили плазмиду UGA\_NI. С использованием этого же набора получены и все остальные плазмиды: UGU\_NI, UAA\_NI/AAA\_NI, UAG\_NI/UAU\_NI.

**Конструкции NI для изучения сквозного прочтения стоп-кодонов мРНК eRF1.** На плазмиде pNI-gl с праймерами pNI F/R получен ПЦР-продукт NI. С помощью праймеров 5'1 F/R, 5'2 F/R, 5'2 F и 5'2sense R (табл. S1, см. Дополнительные материалы на сайте [http://www.molecbio.ru/downloads/2024/4/supp\\_Shupalov\\_rus.pdf](http://www.molecbio.ru/downloads/2024/4/supp_Shupalov_rus.pdf)) получены соответствующие ПЦР-продукты. С помощью системы Gibson Assembly (“NEB”, E2611L, США) из ПЦР-продуктов NI, 5'1 и 5'2 получена плазмида 5'\_NI (1), а из ПЦР-продуктов NI, 5'1 и 5'2sense плазмида 5'sense\_NI (1), в которой стоп-кодон TAG кОРС мРНК eRF1 заменен на TCG, кодирующий серин (табл. S2). На плазмиде pNI-gl с праймерами pNI F и pNI-gl R (табл. S1) получен ПЦР-продукт NI-gl. С помощью системы Gibson Assembly (“NEB”, E2611L) из ПЦР-продуктов NI-gl и 5'2 получена плазмида 5'\_NI (3), а из ПЦР-продуктов NI-gl и 5'2sense – плазмида 5'sense\_NI (3) (табл. S2). На полученной нами ранее плазмиде petSUMO-eRF1

с помощью пар праймеров pNI\_eRF1 F/R, pNI\_eRF1 F и pNI\_eRF1-Ser R (табл. S1) синтезировали ПЦР-продукты OPC и OPCsense. С помощью системы Gibson Assembly (“NEB”, E2611L) из ПЦР-продуктов NI-gl и OPC получена плаزمид OPC\_NI, а из ПЦР-продуктов NI-gl и OPCsense – плазмид OPCsense\_NI, в которой стоп-кодон TAG OPC eRF1 заменен на TCG, кодирующий серин (табл. S2). Конструкция с OPC eRF1 содержала дополнительно природную 3'-контекстную последовательность из 9 нуклеотидов после стоп-кодона UAG.

**Конструкции Fluc для изучения влияния на трансляцию 5'- и 3'-НТО мРНК eRF1.** На плазмиде pGEM-Fluc (“Promega”, кат. № E1541) с использованием праймеров F1 F/R получен ПЦР-продукт F1. С использованием праймеров 3'GAPDH F/R получен ПЦР-продукт 3'GAPDH. С помощью системы Gibson Assembly (“NEB”, E2611L) из ПЦР-продуктов F1 и 3'GAPDH получена плазмид F1\_3'GAPDH (табл. S2). С использованием праймеров 3' F и 3' R на плазмиде pGEM-3'eRF1 получен ПЦР-продукт 3'eRF1. С помощью системы Gibson Assembly (“NEB”, E2611L) из ПЦР-продуктов F1 и 3'eRF1 получена плазмид F1\_3'eRF1 (табл. S2). С использованием праймеров 5'\_F1 F/R и плазмид 5'\_NI (1) и 5'sense\_NI (1) получены ПЦР-продукты 5'\_F1 и 5'sense\_F1. ПЦР-продукт 5'\_F1 был вставлен в плазмиды F1\_3'GAPDH и F1\_3'eRF1 по сайтам рестрикции NheI и XbaI для получения конструкций 5'eRF1\_F1\_3'GAPDH и 5'eRF1\_F1\_3'eRF1, а ПЦР-продукт 5'sense\_F1 вставлен в плазмиду F1\_3'eRF1 для получения 5'eRF1sense\_F1\_3'eRF1 (табл. S2).

Использованные в работе мРНК получены из соответствующих ПЦР-продуктов, содержащих вносимые праймерами T7 промотор на 5'-конце и поли(A)<sub>50</sub>-хвост на 3'-конце, с помощью набора T7 RiboMAX™ Large Scale RNA Production System (“Promega”) с добавлением аналога кеп-структуры мРНК (ARCA, “NEB”) согласно протоколу производителя. Список мРНК и названия праймеров, использованных для их синтеза, приведены в дополнительных материалах (табл. S2, см. Дополнительные материалы).

**Выделение и очистка рекомбинантного eRF1.** Рекомбинантный eRF1 получали из ранее созданной нами конструкции retSUMO-eRF1, как описано для NSP1 [33], проводили дополнительную очистку препарата eRF1 с помощью ионообменной хроматографии на колонке HiTrap Q HP (“Cytiva”).

**In vivo и in vitro трансляция в бесклеточных системах трансляции и клетках HEK293.** Трансляцию репортерных мРНК осуществляли в 10 мкл смеси, содержащей 50% лизата HEK293, 20 мМ HEPES-KOH pH 7.5, 2 мМ DTT, 0.25 мМ спермидина, 0.6 мМ MgAc<sub>2</sub>, 16 мМ креатинфосфат, 0.06 ед/мкл креатинкиназы, 1 мМ АТФ, 0.6 мМ GTP, 60 мМ

КАс, 0.05 мМ каждой аминокислоты, 0.2 ед/мкл RiboLock, 0.5 мМ D-люциферин (или 1% NanoGlo® (“Promega”). eRF1 добавляли до конечной концентрации 0.6 мкМ. Интенсивность люминесценции измеряли при 30°C в планшетном ридере Tecan Infinite 200 Pro в течение 100 мин. Эффективность трансляции рассчитывали как максимальную производную растущего линейного участка кривой люминесценции (v<sub>0</sub>, RLU/мин). Сквозное прочтение стоп-кодона рассчитывали как отношение эффективности трансляции матрицы с ПСК к эффективности трансляции контрольной матрицы со смысловым кодоном в данном положении.

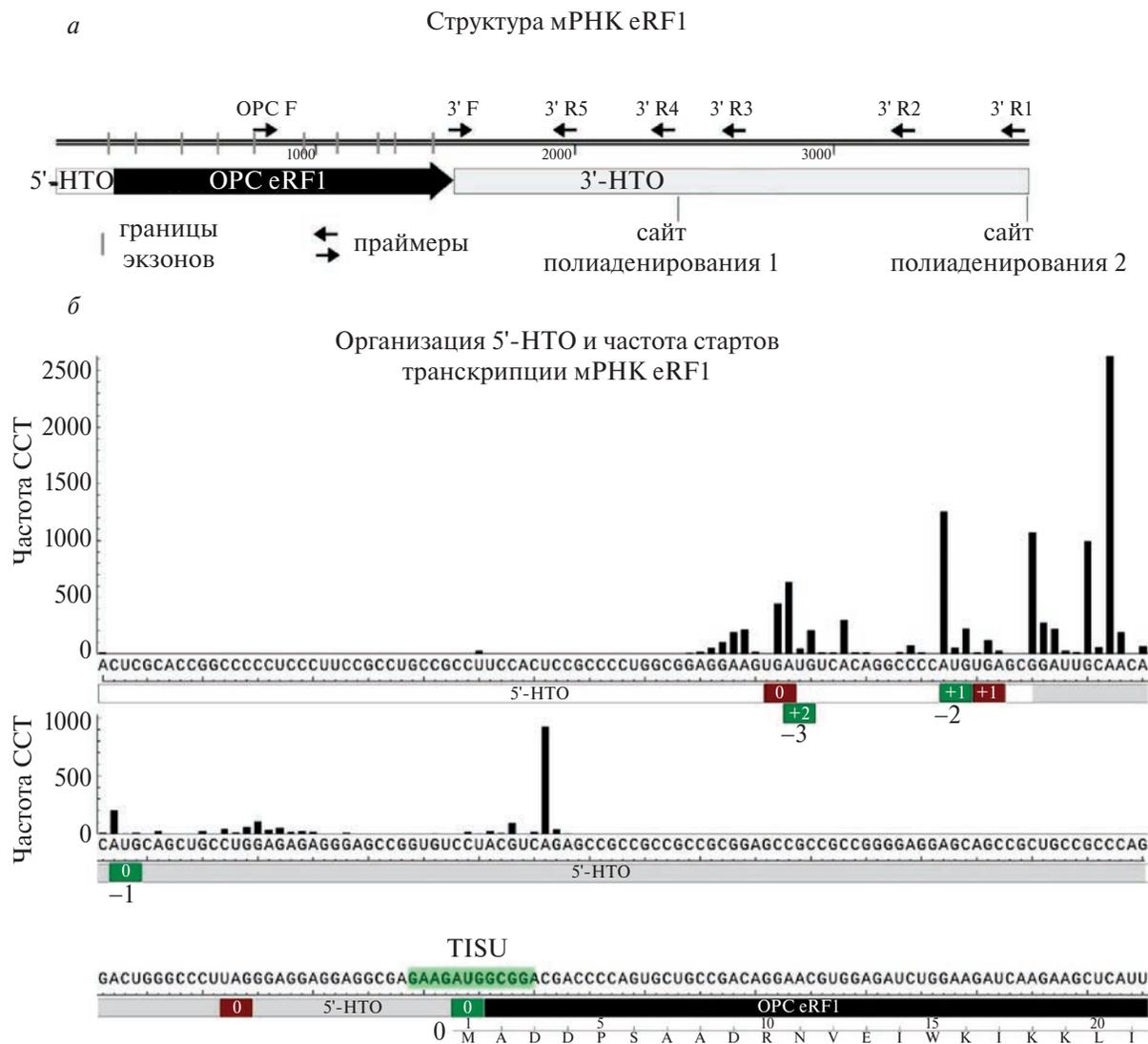
*In vivo* трансляцию репортеров осуществляли в клетках HEK293 в соответствии с ранее описанным методом FLERT (Fleeting mRNA Transfection technique) [38]. Клетки HEK293 культивировали в среде DMEM с 10% фетальной сыворотки крупного рогатого скота при 37°C и 5% CO<sub>2</sub>. За день до трансфекции экспоненциально растущие клетки HEK293 рассаживали на 24-луночные планшеты с плотностью покрытия 50%. Через 12–16 ч роста, когда плотность клеток достигала 70–90%, проводили трансфекцию с использованием GenJector™-U (“Molnec”, Россия) для РНК. Для этого 0.5 пмоль репортерной мРНК и 0.05 пмоль контрольной мРНК инкубировали с 0.2 мкл ре-агента для трансфекции в 60 мкл Opti-MEM (“Gibco”, 11058–021) в течение 10 мин, а затем добавляли в питательную среду к клеткам HEK293T в соответствующую лунку. Через 2 ч клетки собирали, активность люциферазы анализировали с помощью системы люциферазы Dual Luciferase® (“Promega”). Трансфекцию повторяли несколько раз в разных клеточных пассажах.

**Статистический анализ.** Эксперименты проводили в трех и более повторах. Данные представлены как среднее значение ± стандартная ошибка среднего (SE). Для сравнения средних значений между двумя группами использовали непарный двусторонний *t*-критерий. Множественные сравнения проводили с использованием поправки Холма на значение *p* [36]. Звездочками отмечены статистически значимые различия (\**p* < 0.05; \*\**p* < 0.01; \*\*\**p* < 0.0001).

## РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

### Организация мРНК eRF1

Референсная последовательность мРНК eRF1 человека в базе данных NCBI RefSeq имеет идентификационный номер NM\_004730. Схема этой мРНК с аннотированными в базе данных GenBank элементами представлена на рис. 1а. мРНК eRF1 имеет длинную 3'-НТО, расположенную в пределах одного экзона, т.е. после стоп-кодона OPC отсутствуют места соединения экзонов (EJC).

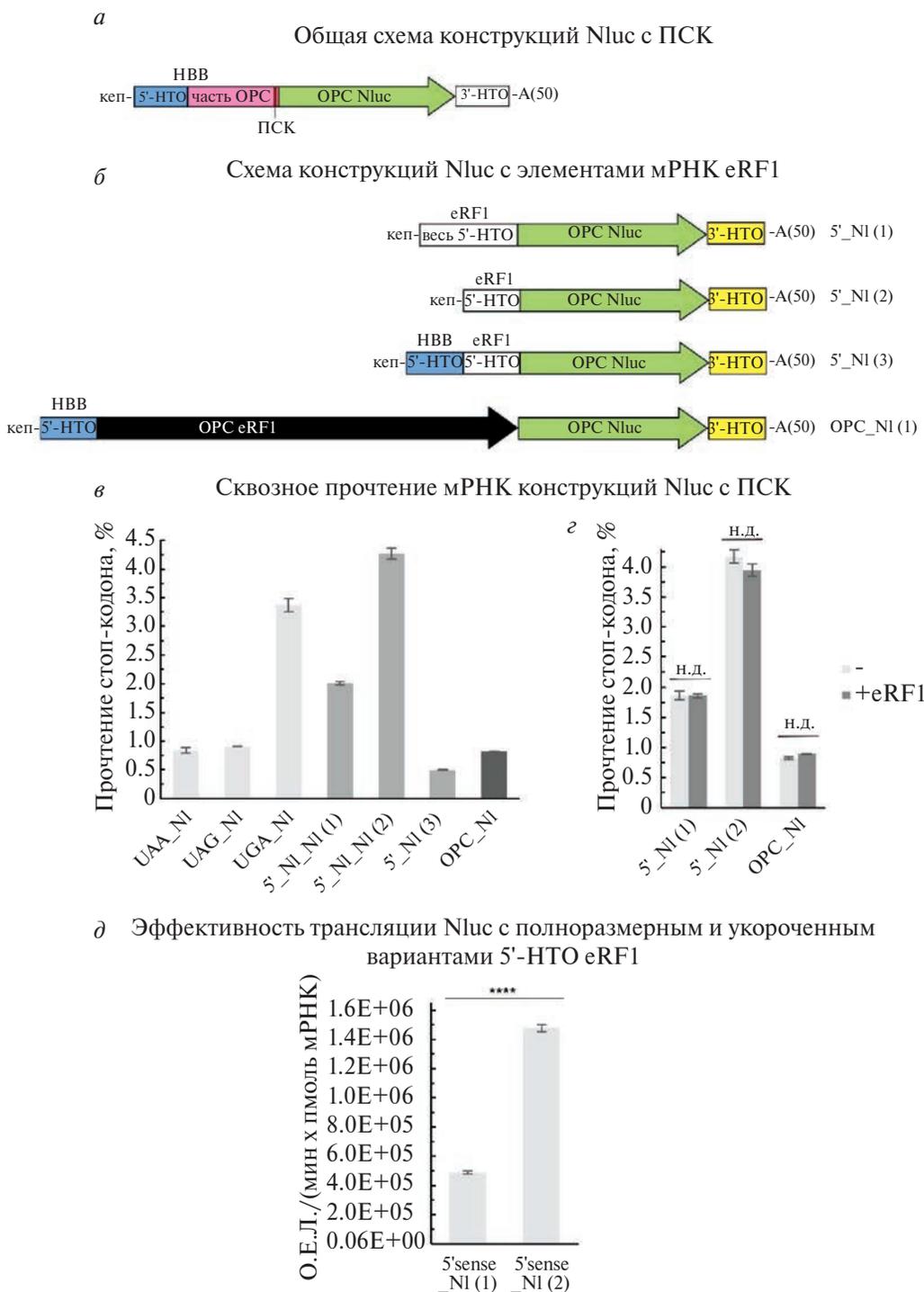


**Рис. 1.** Структура мРНК eRF1. *a* – Схема организации референсной мРНК eRF1 NM\_004730. Стрелками отмечены праймеры, использованные для получения 3'-НТО мРНК eRF1. *б* – Схема организации 5'-НТО eRF1 и частота использования сайтов старта транскрипции согласно базе DBTSS. Зелеными и красными прямоугольниками отмечены старт- и стоп-кодоны соответственно. Цифра внутри прямоугольников означает рамку считывания относительно основной ОПР. Цифра рядом с зелеными квадратами означает порядковый номер AUG относительно старт-кодона основной ОПР, принятый за 0. Светло-зеленым цветом выделена последовательность регуляторного элемента TISU. Серым обозначена укороченная последовательность 5'-НТО eRF1, изучаемая в работе, наряду с полноразмерной.

Таким образом, к eRF1 человека не применим механизм авторегуляции экспрессии за счет EJC-зависимого NMD.

3'-НТО мРНК eRF1 клонировали из клеточной линии НЕК293 эмбриональной почки человека. Из клеток НЕК293 была выделена тотальная мРНК (в трех биологических повторах), и на каждом из образцов синтезировали два вида кДНК с использованием олиго-dT или случайных праймеров. Из образцов кДНК, полученных с использованием олиго-dT и специфичного прямого праймера, синтезировались ПЦР-продукты длиной около 1600 п.н., которые клонировали в вектор. Секвенирование полученных последовательностей показало,

что они соответствуют области между прямым праймером и сайтом полиаденилирования 1, т.е. укороченному фрагменту 3'-НТО (рис. 1*a*). При этом ПЦР-продукты большего размера не детектировались. Затем из каждого полученного варианта кДНК мы попытались синтезировать ПЦР-продукты с прямым специфичным праймером и одним из обратных праймеров на 3'-НТО мРНК eRF1 (3'R1–5). При этом во всех случаях ПЦР-продукты получались только с обратными праймерами 3'R4–5 (рис. 1*a*). Секвенирование этих ПЦР-продуктов показало, что они соответствуют ранее клонированному укороченному 3'-НТО. Таким образом, в НЕК293 преимущественно присутствует



**Рис. 2.** Влияние на сквозное прочтение и трансляцию 5'-HTO и OPC eRF1, а также избытка eRF1. *а* – Общая схема конструкций с преждевременными стоп-кодонами перед OPC нанолуциферазы. Конструкция включает 5'-HTO и первые 13 кодонов OPC мРНК β-глобина с последующим стоп-/смысловым кодоном (UAA/AAA, UAG/UAU, UGA/UGU) и соответствующим 6-нуклеотидным 3'-контекстом (GGGCUG, GAUAAU, CUAGUA) для усиления сквозного прочтения, за которыми следует OPC Nluc без собственного старт-кодона в той же рамке считывания. 3'-HTO является искусственной и продуцируется из ~170 нуклеотидов после стоп-кодона в OPC Nluc в исходной плазмиде. *б* – Схема модельных конструкций для изучения сквозного прочтения стоп-кодонов мРНК eRF1. *в* – Измеренный в бесклеточных системах трансляции уровень сквозного прочтения модельных мРНК с репортером, содержащим нанолуциферазу. *г* – Влияние избытков eRF1 на сквозное прочтение стоп-кодонов в моделях с 5'-HTO и OPC собственной мРНК. *д* – Влияние длины 5'-HTO eRF1 на трансляцию. Сравнимые матрицы имеют идентичные последовательности после старта трансляции и отличаются только длиной 5'-HTO. О.Е.Л. – относительные единицы люминесценции, н.д. – различия не достоверны.

укороченный вариант 3'-НТО мРНК eRF1, который в дальнейшем использовали в экспериментах.

5'-НТО мРНК eRF1 содержит дополнительные старт- и стоп-кодоны в различных рамках считывания (рис. 1б). Интересно, что вторые старт- (-2) и стоп-кодоны организуют минимальную кОПС (AUGUGA), расположенную в другой рамке считывания относительно основной ОПС. При этом, третьи старт- и стоп-кодоны образуют еще одну кОПС (далее также 5'ОПС) в одной рамке с основной ОПС. Подробные данные о стартах транскрипции NM\_004730 (и размерах 5'-НТО) доступны в базе данных DBTSS (Data Base Transcriptional Start Sites, <https://dbtss.hgc.jp>) [37]. Согласно этим данным, полноразмерный 5'-НТО eRF1 в НЕК293 (и других доступных в базе образцах) синтезируется крайне редко, а старты транскрипции располагаются в широком диапазоне, преимущественно перед и после дополнительных старт-кодонов (рис. 1б). Нами получена полноразмерная 5'-НТО eRF1 (далее – “вся 5'-НТО” eRF1), а также укороченная, чаще транскрибируемая, версия 5'-НТО eRF1 (далее – “5'-НТО” eRF1), не включающая два первых дополнительных старт- и стоп-кодона (рис. 1б). Таким образом, мРНК eRF1 в клетках чаще всего укорочена с 5'- и 3'-концов относительно референсной последовательности, а размер ее 5'-НТО варьирует в широких пределах.

*Длина 5'-НТО eRF1 влияет на эффективность трансляции и сквозного прочтения. eRF1 не влияет на собственную трансляцию*

Для изучения влияния 5'-НТО eRF1 и белка eRF1 на трансляцию и сквозное прочтение стоп-кодонов мы создали ряд конструкций с репортером Nluc (рис. 2б). В данных конструкциях Nluc была лишена собственного старт-кодона. В конструкциях 5'\_N1 (1–3) репортер Nluc детектировал трансляцию кОПС 5'-НТО eRF1 (5'-ОПС). Для изучения сквозного прочтения были созданы соответствующие контрольные конструкции, в которых стоп-кодон UAG был заменен кодоном UCG, кодирующим серин. 5'\_N1 (1) содержала всю 5'-НТО eRF1 из референсной последовательности NM\_004730, а 5'\_N1 (2) – укороченный вариант (отмечен серым на рис. 1б). 5'\_N1 (3) содержала гибридную 5'-НТО, в которой непосредственно перед 5'-ОПС расположена последовательность 5'-НТО β-глобина (НВВ), не содержащая дополнительных старт- и стоп-кодонов. ОПС\_N1 перед репортером содержит 5'-НТО НВВ с последующими полной ОПС eRF1 и 9-нуклеотидным натуральным 3'-контекстом ее стоп-кодона.

Для оценки сквозного прочтения на разных стоп-кодонах созданы независимые конструкции, общая схема которых представлена на рис. 1а. В этих конструкциях перед ОПС Nluc расположены 5'-НТО и первые 13 кодонов ОПС НВВ с последующим стоп-/смысловым кодоном (UAA/AAA,

UAG/UAU, UGA/UGU) и соответствующим 6-нуклеотидным 3'-контекстом (GGGCUG, GAUAAU, CUAGUA) для усиления сквозного прочтения [35].

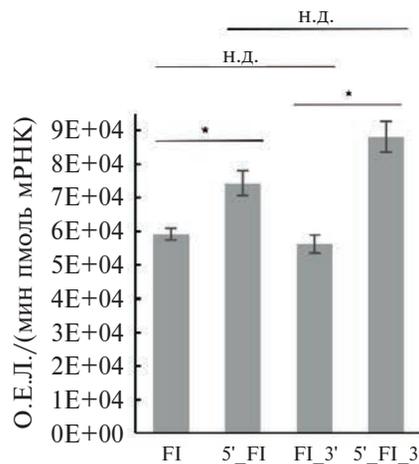
Изучение сквозного прочтения стоп-кодона UAG 5'-ОПС показывает, что на него влияет предшествующая последовательность: максимальная эффективность сквозного прочтения наблюдается на укороченной 5'-НТО eRF1 (более распространенной в клетке), чуть хуже прочитывается стоп-кодон с полноразмерной 5'-НТО eRF1 и хуже всего происходит сквозное прочтение стоп-кодона с 5'-НТО НВВ (рис. 2в). При этом сквозное прочтение UAG 5'-ОПС в присутствии 5'-НТО eRF1 выше, чем в модельной конструкции с преждевременным стоп-кодоном (ПСК) UAG\_N1, однако 5'-НТО НВВ снижает эффективность сквозного прочтения. Сквозное прочтение UAG ОПС eRF1 не отличается от сквозного прочтения ПСК UAG\_N1. Отсутствие повышенного сквозного прочтения основной ОПС eRF1 является дополнительным доводом против возможности описанной у растений и грибов авторегуляции трансляции eRF1 посредством NMD. Более того, добавление избытков eRF1 не влияет на сквозное прочтение стоп-кодонов в 5'-ОПС и основной ОПС eRF1 и на трансляцию соответствующих конструкций (рис. 2г, д). Эти данные противоречат возможности авторегуляции трансляции eRF1. Эффективность трансляции в присутствии полноразмерной 5'-НТО eRF1 ниже, чем в присутствии укороченной, что обусловлено, вероятнее всего, большим количеством дополнительных старт-кодонов (рис. 2д).

*5'-НТО и 3'-НТО eRF1 обеспечивают эффективную трансляцию основной рамки считывания*

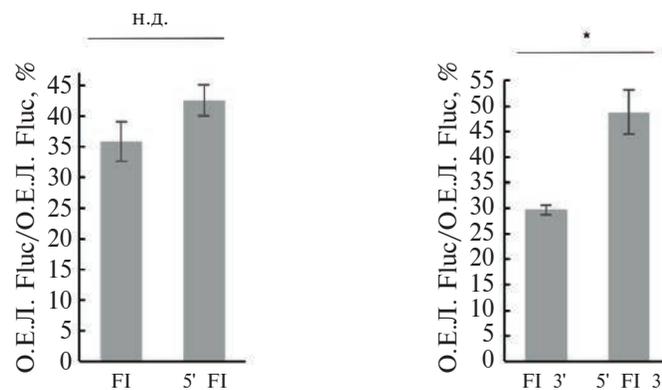
Для изучения влияния 5'- и 3'-НТО eRF1 на трансляцию фактора терминации созданы модельные мРНК на основе репортера Fluc (рис. 3а). В отличие от небольшой ОПС Nluc, ОПС Fluc имеет схожие с ОПС eRF1 длину и GC-состав. В качестве контроля использовали 5'-НТО НВВ и 3'-НТО GAPDH. Соответствующие белки представлены в клетках в высоких концентрациях, и НТО их мРНК должны способствовать эффективной трансляции. *In vitro* в лизате клеток НЕК293 5'-НТО eRF1 увеличивает эффективность трансляции даже относительно 5'-НТО НВВ, в то время как 3'-НТО eRF1 не улучшает трансляцию репортера Fluc (рис. 3б). 5'- и 3'-НТО eRF1 совместно обеспечивают максимальный уровень трансляции репортера. Этот же эксперимент мы повторили *in vivo* в клетках НЕК293, трансфицировав их теми же мРНК в соответствии с методом FLERT (Fleeting mRNA Transfection technique) [38]. *In vivo* 5'- и 3'-НТО eRF1 совместно достоверно улучшают трансляцию репортера Fluc, в то время как одна 3'-НТО eRF1 не влияет заметно на трансляцию



**б** Влияние 5'- и 3'-НТО мРНК eRF1 на трансляцию Fluc *in vitro*



**в** Влияние 5'- и 3'-НТО мРНК eRF1 на трансляцию Fluc *in vivo*



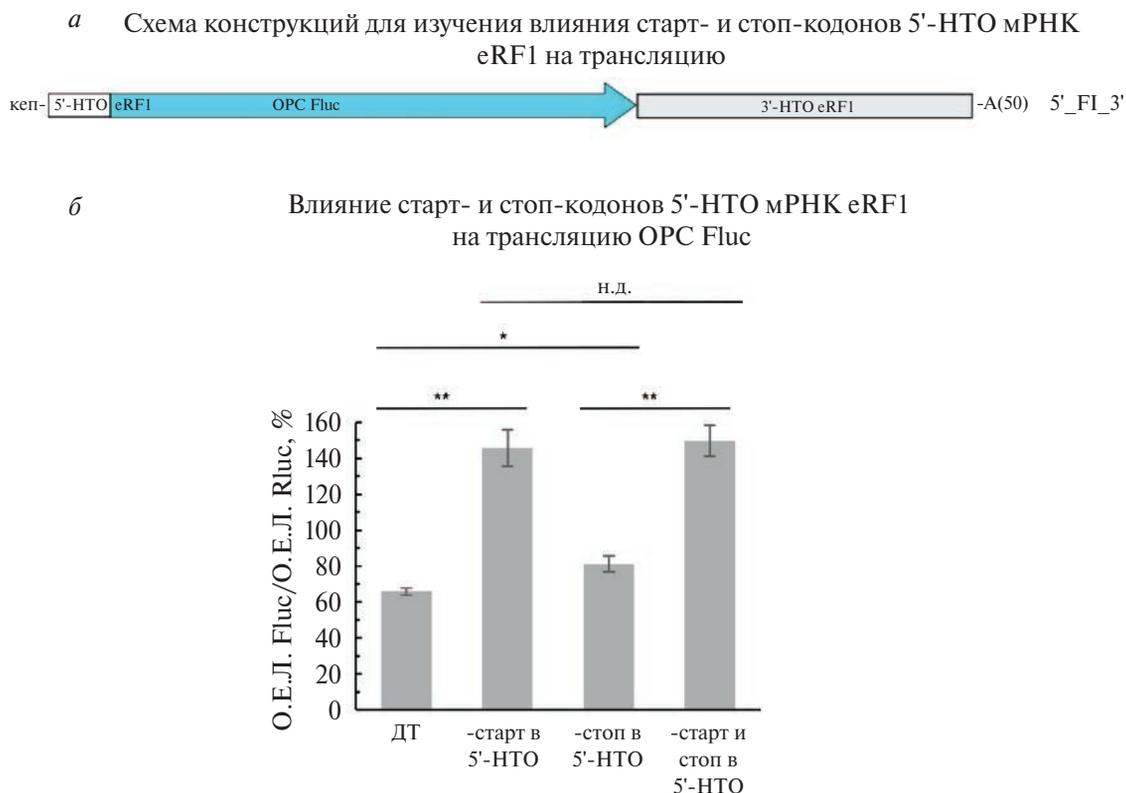
**Рис. 3.** Влияние 5'-НТО и 3'-НТО мРНК eRF1 на трансляцию. *а* – Схема конструкций, кодирующих люциферазу светлячка, использованных в экспериментах. *б* – Влияние 5'- и 3'-НТО мРНК eRF1 на трансляцию Fluc в лизате HEK293. *в* – Влияние 5'- и 3'-НТО мРНК eRF1 на трансляцию Fluc в культуре клеток HEK293.

(рис. 3в). Таким образом, 5'-НТО eRF1 сильнее всего влияет на трансляцию фактора терминации, тогда как эффект 3'-НТО eRF1 зависит от наличия 5'-НТО eRF1.

*Старт-кодон 5'-НТО eRF1 ингибирует трансляцию*

Влияние дополнительных старт- и стоп-кодонов, образующих kOPC в 5'-НТО eRF1, на трансляцию основной OPC мы изучали с использованием

созданных нами вариантов мРНК OPC Fluc с 5'- и 3'-НТО eRF1, в которых заменяли старт- и/или стоп-кодон на смысловой кодон (AAG и UCG соответственно). Эти конструкции использовали в экспериментах *in vivo* в культуре клеток HEK293 с помощью метода FLERT [38]. Удаление альтернативного старт-кодона из 5'-НТО eRF1 привело к более чем двукратному увеличению эффективности трансляции репортера, в то время как удаление стоп-кодона в 5'-НТО eRF1 практически не повлияло на



**Рис. 4.** Влияние старт- и стоп-кодонов в кОПС 5'-НТО мРНК eRF1. *a* – Общая схема конструкций, представляющая четыре варианта мРНК: “ДТ” – дикий тип 5'-НТО eRF1 не изменена и содержит старт- и стоп-кодон кОПС, “-старт в 5'-НТО” – старт-кодон кОПС заменен на кодон AAG, “-стоп в 5'-НТО” – стоп-кодон кОПС заменен на UCG, “-старт и стоп в 5'-НТО” – старт-кодон кОПС заменен на кодон AAG, а стоп-кодон – на UCG. *б* – Эффективность трансляции *in vivo* в клетках HEK293 репортерной мРНК с мутациями старт- и стоп-кодона кОПС 5'-НТО из мРНК eRF1.

трансляцию (рис. 4). Таким образом, наиболее заметным регулятором трансляции в мРНК eRF1 является старт-кодон кОПС в 5'-НТО eRF1.

#### Механизм регуляции трансляции мРНК eRF1 некодирующими областями

Полученные нами данные показывают, что наибольшее влияние на трансляцию оказывает 5'-НТО eRF1, а также ее сочетание с 3'-НТО eRF1, которые стимулируют трансляцию основной ОПС. При этом старт-кодон кОПС в 5'-НТО eRF1 ингибирует трансляцию, так как его удаление дополнительно увеличивает эффективность трансляции более чем в 2 раза, что в целом соответствует представлениям о регуляции трансляции кОПС [39–42]. Мы предполагаем, что трансляция eRF1 регулируется на двух уровнях: на уровне транскрипции и на уровне трансляции. Как видно из рис. 1б, мРНК eRF1 гетерогенна, имеет лидерные участки разной длины, при этом множественные сайты старта транскрипции концентрируются в области дополнительных старт-кодонов в 5'-НТО. Таким образом, за счет регуляции транскрипции может продуцироваться мРНК eRF1 с различным количеством дополнительных старт-кодонов и кОПС. В свою очередь,

дополнительные старт-кодоны и кОПС регулируют трансляцию по известным механизмам, ингибируя трансляцию основной ОПС за счет “захвата” факторов инициации трансляции. Таким образом, наибольшую эффективность трансляции eRF1 должны обеспечивать транскрипционные варианты мРНК с короткой 5'-НТО, не включающие дополнительные старт-кодоны. Следует отметить, что регуляция трансляции основной ОПС посредством кОПС активно изучается на других объектах, и полученные нами данные соответствуют общепринятому мнению о том, что кОПС подавляют трансляцию основной ОПС [39–42]. Кроме того, в ряде работ поднимается вопрос о регуляции трансляции за счет гетерогенности старта транскрипции некоторых мРНК, приводящей к возникновению в них набора 5'-НТО различной длины [43–45], что соответствует нашим выводам о транскрипции eRF1.

Также мы заметили, что старт-кодон основной ОПС eRF1 входит в состав трансляционно-транскрипционного регуляторного элемента TISU (Translation Initiator of Short 5'UTR). Ранее считалось, что трансляция первого старт-кодона мРНК, имеющей слишком короткую 5'-НТО, происходит неэффективно, с “утечкой” старта трансляции на

следующий AUG [46]. Однако позднее был описан регуляторный элемент TISU, способствующий эффективной трансляции с первого старт-кодона в матрицах мРНК, имеющих короткую 5'-НТО [47]. Локализация TISU частично перекрывается с последовательностью Козак. Кроме того, TISU описан как цис-регулятор транскрипции, способный связываться с транскрипционным фактором YY1 (Yin Yang 1 – Инь Ян 1) и эффективно стимулировать синтез мРНК с промоторов генов, имеющих высокую гетерогенность сайтов старта транскрипции [47]. Наличие TISU и его положение в мРНК eRF1 согласуются с нашей гипотезой о двухуровневой транскрипционно-трансляционной регуляции трансляции eRF1.

В целом, можно заключить, что регуляция трансляции основной OPC eRF1 человека принципиально отличается от регуляции трансляции мРНК eRF1 растений и грибов, у которых описано существование NMD-зависимого авторегуляторного контура экспрессии [31, 32]. Продукция фактора терминации eRF1 человека не зависит от его концентрации в клетке и регулируется собственными 5'- и 3'-НТО.

Исследование выполнено за счет гранта Российского научного фонда (№ 22-24-01019)

Настоящая работа выполнена без использования людей или животных в качестве объектов исследования.

Авторы сообщают об отсутствии конфликта интересов.

Дополнительные материалы доступны на сайте: [http://www.molecbio.ru/downloads/2024/4/supp\\_Shuvalov\\_rus.pdf](http://www.molecbio.ru/downloads/2024/4/supp_Shuvalov_rus.pdf).

## СПИСОК ЛИТЕРАТУРЫ

- Bertram G., Bell H.A., Ritchie D.W., Fullerton G., Stansfield I. (2000) Terminating eukaryote translation: domain 1 of release factor eRF1 functions in stop codon recognition. *RNA*. **6**, 1236–1247.
- Frolova L., Seit-Nebi A., Kisselev L. (2002) Highly conserved NIKS tetrapeptide is functionally essential in eukaryotic translation termination factor eRF1. *RNA*. **8**, 129–136.
- Bulygin K.N., Khairulina Y.S., Kolosov P.M., Ven'yaminova A.G., Graifer D.M., Vorobjev Y.N., Frolova L.Y., Kisselev L.L., Karpova G.G. (2010) Three distinct peptides from the N domain of translation termination factor eRF1 surround stop codon in the ribosome. *RNA*. **16**, 1902–1914.
- Bulygin K.N., Khairulina Y.S., Kolosov P.M., Ven'yaminova A.G., Graifer D.M., Vorobjev Y.N., Frolova L.Y., Karpova G.G. (2011) Adenine and guanine recognition of stop codon is mediated by different N domain conformations of translation termination factor eRF1. *Nucl. Acids Res.* **39**, 7134–7146.
- Kryuchkova P., Grishin A., Eliseev B., Karyagina A., Frolova L., Alkalaeva E. (2013) Two-step model of stop codon recognition by eukaryotic release factor eRF1. *Nucl. Acids Res.* **41**, 4573–4586.
- Brown A., Shao S., Murray J., Hegde R.S., Ramakrishnan V. (2015) Structural basis for stop codon recognition in eukaryotes. *Nature*. **524**, 493–496.
- Frolova L.Y., Tsivkovskii R.Y., Sivolobova G.F., Oparina N.Y., Serpinsky O.I., Blinov V.M., Tatkov S.I., Kisselev L.L. (1999) Mutations in the highly conserved GGQ motif of class 1 polypeptide release factors abolish ability of human eRF1 to trigger peptidyl-tRNA hydrolysis. *RNA*. **5**, 1014–1020.
- Zhouravleva G., Frolova L., Le Goff X., Le Guellec R., Inge-Vechtomov S., Kisselev L., Philippe M. (1995) Termination of translation in eukaryotes is governed by two interacting polypeptide chain release factors, eRF1 and eRF3. *EMBO J.* **14**, 4065–4072.
- Frolova L., Le Goff X., Zhouravleva G., Davydova E., Philippe M., Kisselev L. (1996) Eukaryotic polypeptide chain release factor eRF3 is an eRF1- and ribosome-dependent guanosine triphosphatase. *RNA*. **2**, 334–341.
- Stansfield I., Jones K.M., Kushnirov V.V., Dagkesamankayal A.R., Poznyakovski A., Paushkin S.V., Nierras C.R., Cox B.S., Ter-Avanesyan M.D., Tuite M.F. (1995) The products of the *SUP45* (eRF1) and *SUP35* genes interact to mediate translation termination in *Saccharomyces cerevisiae*. *EMBO J.* **14**, 4365–4373.
- Paushkin S.V., Kushnirov V.V., Smirnov V.N., Ter-Avanesyan M.D. (1997) Interaction between yeast Sup45p (eRF1) and Sup35p (eRF3) polypeptide chain release factors: implications for prion-dependent regulation. *Mol. Cell. Biol.* **17**, 2798–2805.
- Ito K., Ebihara K., Nakamura Y. (1998) The stretch of C-terminal acidic amino acids of translational release factor eRF1 is a primary binding site for eRF3 of fission yeast. *RNA*. **4**, S1355838298971874.
- Merkulova T.I., Frolova L.Y., Lazar M., Camonis J., Kisselev L.L. (1999) C-terminal domains of human translation termination factors eRF1 and eRF3 mediate their *in vivo* interaction. *FEBS Lett.* **443**, 41–47.
- Namy O., Hatin I., Rousset J. (2001) Impact of the six nucleotides downstream of the stop codon on translation termination. *EMBO Rep.* **2**, 787–793.
- Atkins J.F., Loughran G., Bhatt P.R., Firth A.E., Baranov P.V. (2016) Ribosomal frameshifting and transcriptional slippage: from genetic steganography and cryptography to adventitious use. *Nucl. Acids Res.* **44**, 7007–7078.
- Rodnina M.V., Korniy N., Klimova M., Karki P., Peng B. Z., Senyushkina T., Belardinelli R., Maracci C., Wohlgemuth I., Samatova E., Peske F. (2020) Translational recoding: canonical translation mechanisms reinterpreted. *Nucl. Acids Res.* **48**, 1056–1067.
- Brogna S., Wen J. (2009) Nonsense-mediated mRNA decay (NMD) mechanisms. *Nat. Struct. Mol. Biol.* **16**, 107–113.
- Шабельская С.В., Журавлева Г.А. (2010) Мутации в гене SUP35 нарушают процесс деградации мРНК, содержащих преждевременные стоп-кодона. *Молекуляр. биология*. **44**, 51–59.
- Schweingruber C., Rufener S.C., Zünd D., Yamashita A., Mühlemann O. (2013) Nonsense-mediated mRNA decay – Mechanisms of substrate mRNA recognition and degradation in mammalian cells. *Biochim. Biophys. Acta (BBA) – Gene Regulatory Mechanisms*. **1829**, 612–623.

20. Karousis E.D., Gurzeler L.-A., Annibaldis G., Dreos R., Mühlemann O. (2020) Human NMD ensues independently of stable ribosome stalling. *Nat. Commun.* **11**, 4134.
21. Yang Q., Yu C.-H., Zhao F., Dang Y., Wu C., Xie P., Sachs M.S., Liu Y. (2019) eRF1 mediates codon usage effects on mRNA translation efficiency through premature termination at rare codons. *Nucl. Acids Res.* **47**, 9243–9258.
22. Wada M., Ito K. (2019) Misdecoding of rare CGA codon by translation termination factors, eRF1/eRF3, suggests novel class of ribosome rescue pathway in *S. cerevisiae*. *FEBS J.* **286**, 788–802.
23. Dever T.E., Ivanov I.P., Sachs M.S. (2020) Conserved upstream open reading frame nascent peptides that control translation. *Annu. Rev. Genet.* **54**, 237–264.
24. Bidou L., Allamand V., Rousset J.-P., Namy O. (2012) Sense from nonsense: therapies for premature stop codon diseases. *Trends Mol. Med.* **18**, 679–688.
25. Janzen D.M. (2004) The effect of eukaryotic release factor depletion on translation termination in human cell lines. *Nucl. Acids Res.* **32**, 4491–4502.
26. Freitag J., Ast J., Böcker M. (2012) Cryptic peroxisomal targeting via alternative splicing and stop codon read-through in fungi. *Nature.* **485**, 522–525.
27. Schueren F., Lingner T., George R., Hofhuis J., Dickel C., Gärtner J., Thoms S. (2014) Peroxisomal lactate dehydrogenase is generated by translational readthrough in mammals. *ELife.* **3**, e03640
28. Hofhuis J., Schueren F., Nötzel C., Lingner T., Gärtner J., Jahn O., Thoms S. (2016) The functional readthrough extension of malate dehydrogenase reveals a modification of the genetic code. *Open Biol.* **6**, 160246.
29. Svidritskiy E., Demo G., Korostelev A.A. (2018) Mechanism of premature translation termination on a sense codon. *J. Biol. Chem.* **293**, 12472–12479.
30. Dubourg C., Toutain B., Le Gall J.-Y., Le Treut A., Guenet L. (2003) Promoter analysis of the human translation termination factor 1 gene. *Gene.* **316**, 91–101.
31. Nyikó T., Auber A., Szabadkai L., Benkovics A., Auth M., Mérai Z., Kerényi Z., Dinnyés A., Nagy F., Silhavy D. (2017) Expression of the eRF1 translation termination factor is controlled by an autoregulatory circuit involving readthrough and nonsense-mediated decay in plants. *Nucl. Acids Res.* **45**, 4174–4188.
32. Kurilla A., Szőke A., Auber A., Káldi K., Silhavy D. (2020) Expression of the translation termination factor eRF1 is autoregulated by translational readthrough and 3'UTR intron-mediated NMD in *Neurospora crassa*. *FEBS Lett.* **594**, 3504–3517.
33. Shuvalov A., Shuvalova E., Biziaev N., Sokolova E., Evmenov K., Pustogarov N., Arnautova A., Matrosova V., Egorova T., Alkalaeva E. (2021) Nsp1 of SARS-CoV-2 stimulates host translation termination. *RNA Biol.* **18**, 804–817.
34. Соколова Е.Е., Власов П.К., Егорова Т.В., Шува-лов А.В., Алкалаева Е.З. (2020) Влияние A/G-состава 3'-контекстов стоп-кодонов на терминацию трансляции у эукариот. *Молекуляр. биология.* **54**, 837–848.
35. Biziaev N., Sokolova E., Yanvarev D.V., Toropygin I.Y., Shuvalov A., Egorova T., Alkalaeva E. (2022) Recognition of 3' nucleotide context and stop codon readthrough are determined during mRNA translation elongation. *J. Biol. Chem.* **298**, 102133.
36. Holm S. (1979) A simple sequentially rejective multiple test procedure. *Scandinavian J. Statistics.* **6**, 65–70.
37. Suzuki A., Kawano S., Mitsuyama T., Suyama M., Kanai Y., Shirahige K., Sasaki H., Tokunaga K., Tsuchihara K., Sugano S., Nakai K., Suzuki Y. (2018) DBTSS/DBKERO for integrated analysis of transcriptional regulation. *Nucl. Acids Res.* **46**, D229–D238.
38. Akulich K.A., Andreev D.E., Terenin I.M., Smirnova V.V., Anisimova A.S., Makeeva D.S., Arkhipova V.I., Stolboush-kina E.A., Garber M.B., Prokofjeva M.M., Spirin P.V., Prassolov V.S., Shatsky I.N., Dmitriev S.E. (2016) Four translation initiation pathways employed by the leaderless mRNA in eukaryotes. *Sci. Rep.* **6**, 37905.
39. Lin Y., May G.E., Kready H., Nazzaro L., Mao M., Spealman P., Creeger Y., McManus C.J. (2019) Impacts of uORF codon identity and position on translation regulation. *Nucl. Acids Res.* **47**, 9358–9367.
40. Dever T.E., Ivanov I.P., Hinnebusch A.G. (2023) Translational regulation by uORFs and start codon selection stringency. *Genes Dev.* **37**, 474–489.
41. Starck S.R., Tsai J.C., Chen K., Shodiya M., Wang L., Yahiro K., Martins-Green M., Shastri N., Walter P. (2016) Translation from the 5' untranslated region shapes the integrated stress response. *Science.* **351**, aad3867
42. Akulich K.A., Sinitcyn P.G., Makeeva D.S., Andreev D.E., Terenin I.M., Anisimova A.S., Shatsky I.N., Dmitriev S.E. (2019) A novel uORF-based regulatory mechanism controls translation of the human MDM2 and eIF2D mRNAs during stress. *Biochimie.* **157**, 92–101.
43. Dieudonné F.-X., O'Connor P.B.F., Gubler-Jaquier P., Yasrebi H., Conne B., Nikolaev S., Antonarakis S., Baranov P.V., Curran J. (2015) The effect of heterogeneous transcription start sites (TSS) on the transcriptome: implications for the mammalian cellular phenotype. *BMC Genomics.* **16**, 986.
44. Wang X., Hou J., Quedenau C., Chen W. (2016) Pervasive isoform-specific translational regulation via alternative transcription start sites in mammals. *Mol. Systems Biol.* **12**, 875
45. Li H., Bai L., Li H., Li X., Kang Y., Zhang N., Sun J., Shao Z. (2019) Selective translational usage of TSS and core promoters revealed by transcriptome sequencing. *BMC Genomics.* **20**, 282.
46. Kozak M. (1991) A short leader sequence impairs the fidelity of initiation by eukaryotic ribosomes. *Gene Expression.* **1**, 111–115.
47. Elfakess R., Dikstein R. (2008) A translation initiation element specific to mRNAs with very short 5'UTR that also regulates transcription. *PLoS One.* **3**, e3094.

## Human eRF1 Translation Regulation

A. V. Shuvalov<sup>1, 2</sup>, A. A. Klishin<sup>1, 3</sup>, N. S. Biziaev<sup>1</sup>, E. Y. Shuvalova<sup>1, 2</sup>, E. Z. Alkalaeva<sup>1, 2, 3, \*</sup>

<sup>1</sup>*Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Moscow, 119991 Russia*

<sup>2</sup>*Center for Precision Genome Editing and Genetic Technologies for Biomedicine, Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Moscow, 119991 Russia*

<sup>3</sup>*Moscow Institute of Physics and Technology, Dolgoprudny, Moscow Region, 141700 Russia*

\*e-mail: [alkalaeva@imb.ru](mailto:alkalaeva@imb.ru)

Translation termination factor eRF1 is an important cellular protein that plays a key role in translation termination, nonsense-mediated mRNA decay (NMD), and stop-codons readthrough. An amount of eRF1 in the cell influences all these processes. The mechanism of eRF1 translation regulation through an autoregulatory NMD-dependent expression circuit has been described for plants and fungi, but the mechanisms of human eRF1 translation regulation have not yet been studied. Using reporter constructs, we studied the effect of eRF1 mRNA elements on its translation in cell-free translation systems and HEK293 cells. Our data do not support the presence of the NMD-dependent autoregulatory circuit of human eRF1 expression. We found that the 5'-untranslated region (5'-UTR) of eRF1 mRNA and the start codon of the upstream open reading frame (uORF) have the greatest influence on the translation of CDS. According to the DataBase of Transcriptional Start Sites (DBTSS), eRF1 mRNA has a high heterogeneity of transcription start sites and variable length of 5'-UTRs as a consequence. Moreover, the start codon of the eRF1 CDS is located within the known Translation Initiator of Short 5'UTR (TISU), which also stimulates mRNA transcription of genes with high transcription start heterogeneity. We hypothesize that regulation of eRF1 mRNA translation occurs at both the transcriptional and translational levels. At the transcription level, the length of the 5'-UTRs of eRF1 and the number of short open reading frames in it are regulated, which in turn regulate eRF1 production at the translational level.

**Keywords:** eRF1, translation, mRNA, 3'-untranslated region of mRNA, 5'-untranslated region of mRNA, short open reading frame, uORF, 5'-UTR, 3'-UTR

УДК 616.399,579.61

## МЕТАБОЛИЧЕСКИЙ ПРОФИЛЬ КИШЕЧНОЙ МИКРОБИОТЫ И СОДЕРЖАНИЕ ТРЕФОИЛОВЫХ ФАКТОРОВ У ВЗРОСЛЫХ С РАЗЛИЧНЫМИ МЕТАБОЛИЧЕСКИМИ ФЕНОТИПАМИ ОЖИРЕНИЯ

© 2024 г. И. М. Колесникова<sup>a, b, \*</sup>, Л. А. Ганенко<sup>c</sup>, И. Ю. Васильев<sup>d</sup>, Т. В. Григорьева<sup>d</sup>, Н. И. Волкова<sup>c</sup>, С. А. Румянцев<sup>a, b, e</sup>, А. В. Шестопапов<sup>a, b, e, f</sup>

<sup>a</sup>Российский национальный исследовательский медицинский университет им. Н. И. Пирогова Министерства здравоохранения Российской Федерации, Москва, 117997 Россия

<sup>b</sup>Национальный медицинский исследовательский центр эндокринологии Министерства здравоохранения Российской Федерации, Москва, 117292 Россия

<sup>c</sup>Ростовский государственный медицинский университет Министерства здравоохранения Российской Федерации, Ростов-на-Дону, 344022 Россия

<sup>d</sup>Казанский (Приволжский) федеральный университет, Казань, 420008 Россия

<sup>e</sup>Центр молекулярного здоровья, Москва, 117437 Россия

<sup>f</sup>Национальный медицинский исследовательский центр детской гематологии, онкологии и иммунологии им. Дмитрия Рогачева Министерства здравоохранения Российской Федерации, Москва, 117997 Россия

\*e-mail: ir.max.kolesnikova@gmail.com

Поступила в редакцию 17.11.2023 г.

После доработки 09.01.2024 г.

Принята к публикации 09.01.2024 г.

Ожирение ассоциировано с изменением состава кишечной микробиоты, а также с увеличением проницаемости кишечной стенки. У 130 здоровых волонтеров, 57 пациентов с метаболически здоровым ожирением (МЗО) и 76 пациентов с метаболически нездоровым ожирением (МНЗО) из образцов фекалий выделяли бактериальную ДНК с дальнейшим секвенированием гена 16S рРНК. Метаболический профиль микробиоты, предсказанный с использованием программного обеспечения PICRUSt2 (<https://huttenhower.sph.harvard.edu/picrust/>), был более изменен у пациентов с МНЗО, чем с МЗО. Ожирение, особенно МНЗО, сопровождалось увеличением способности кишечной микробиоты к катаболизму энергетических субстратов и продукции энергии в общих путях катаболизма, синтезу водорастворимых витаминов (В1, В6, В7), нуклеотидов, гема, ароматических аминокислот и ряда защитных структурных компонентов клеток. Подобные изменения могут быть следствием адаптации микробиоты к условиям, специфическим для МНЗО. По-видимому, таким образом формируется порочный круг, когда МНЗО способствуют истощению кишечного микробиома, а дальнейшее перерождение последнего вносит вклад в патогенез формирования метаболических нарушений. В сыворотке крови участников исследования определяли концентрацию треофиловых факторов, которая может отражать состояние мукозального барьера. У пациентов с МЗО и МНЗО было повышено содержание треофиловых факторов 2 и 3 типов, но значимых ассоциаций их уровня с метаболическим профилем кишечной микробиоты не зарегистрировано.

**Ключевые слова:** метаболически здоровое ожирение, метаболически нездоровое ожирение, кишечная микробиота, метаболический профиль микробиоты, TFF2, TFF3, PICRUSt2

DOI: 10.31857/S0026898424040105, EDN: IMMUOM

В последние десятилетия во многих исследованиях продемонстрирована тесная взаимосвязь между кишечной микробиотой и ожирением (см. обзоры [1, 2]). Большинство работ сосредоточено

на таксономическом составе кишечного микробиома, его изменениях при ожирении и взаимосвязи с метаболическими нарушениями и клинической картиной. Не так давно появились новые инструменты в изучении кишечного микробиома — методы реконструкции метаболического профиля микробного сообщества на основании результатов секвенирования гена 16S рРНК. Один из таких инструментов — программное обеспечение PICRUSt2 (<https://huttenhower.sph.harvard.edu/picrust/>), позволяющее прогнозировать функциональный

Сокращения: SCFAs (short-chain fatty acids) — короткоцепочечные жирные кислоты; TFF (trefoil factor family) — треофиловые факторы; TFF2/3 (trefoil factor family member 2/3) — треофиловый пептид типа 2/3; МЗО — метаболически здоровое ожирение; МНЗО — метаболически нездоровое ожирение; ЦТК — цикл трикарбоновых кислот.

потенциал микробного сообщества [3]. В ряде работ продемонстрировано влияние ожирения, его тяжести и сопутствующих метаболических нарушений на прогнозируемый функциональный профиль кишечной микробиоты [4–6]. Однако эти исследования проведены в основном на популяции жителей Китая. Но назрела необходимость выявления общих паттернов изменения метаболического профиля при ожирении – вне зависимости от особенностей диеты, этнических и географических факторов.

Кишечный микробиом не только тесно связан с метаболическими нарушениями при ожирении, но и влияет на состояние мукозального барьера, модулируя иммунную систему [7]. Ожирение негативно сказывается на состоянии стенки кишечника, приводя к увеличению проницаемости и всасывания патогенассоциированных молекулярных паттернов, что может усиливать провоспалительный статус пациентов с ожирением [8]. В связи с этим интерес представляет изучение уровня трефоиловых факторов (trefoil factor family, TFF) у пациентов с ожирением. Представители TFF секретируются железистым эпителием пищеварительной, респираторной и мочеполовой систем вместе с муцином и участвуют в процессах защиты и регенерации слизистых [9]. TFF усиливают миграцию и дифференцировку клеток, модулируют их апоптоз, а также регулируют синтез белков плотных контактов, что в том числе способствует поддержанию целостности стенки желудочно-кишечного тракта (ЖКТ) [10]. Синтез трефоилового пептида типа 2 (trefoil factor family member 2, TFF2) происходит в основном в желудке и двенадцатиперстной кишке, а трефоилового пептида типа 3 (TFF3) – в тонкой и толстой кишке [9, 11, 12]. Кроме участия в репарации кишечной стенки, TFF2 проявляет противовоспалительную активность, снижая продукцию провоспалительных цитокинов [13]. Таким образом, по уровню TFF можно судить о состоянии мукозального барьера. Следует отметить, что уровень системного воспаления, метаболических нарушений и изменений в кишечном микробиоме зависит от метаболического фенотипа ожирения [14–16].

Целью работы было изучение метаболического профиля кишечной микробиоты и его влияния на состояние мукозального барьера, что оценивали по уровню TFF у пациентов с разными метаболическими фенотипами ожирения.

## ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

**Группы исследования.** В 2019–2022 гг. было проведено когортное одномоментное исследование на базе ФГБОУ ВО РостГМУ Минздрава России, ООО “Центр молекулярного здоровья” и ФГАОУ ВО “Казанский (Приволжский) федеральный

университет”. Участниками исследования стали 263 человека старше 18 лет. Критериями включения в исследование были: отсутствие приема анти-/пре-/пробиотических препаратов в течение полугода до момента включения в исследование и подписанное добровольное информированное согласие. Критериями невключения были: тяжелые соматические заболевания (хроническая почечная недостаточность, хроническая печеночная недостаточность, хроническая сердечная недостаточность), заболевания ЖКТ, алкоголизм, беременность, депрессия, а также любое заболевание в острой форме.

Контрольную группу составили 130 здоровых волонтеров без ожирения (индекс массы тела (ИМТ) 18.5–24.9 кг/м<sup>2</sup>), без признаков метаболических нарушений и артериальной гипертензии. Пациенты с ожирением (n = 133): ИМТ > 30 кг/м<sup>2</sup> и окружность талии >102 см у мужчин и >88 см у женщин, – в зависимости от метаболического фенотипа ожирения по критериям NCEP-АТР III, были разделены на подгруппы с метаболически здоровым ожирением (МЗО) и с метаболически нездоровым ожирением (МНЗО) [17]. Ожирение считалось метаболически нездоровым, если для пациента были характерны два и более критериев: 1) триглицериды сыворотки ≥ 1.7 ммоль/л; 2) холестерин липопротеинов высокой плотности у мужчин <1.03 ммоль/л, у женщин <1.29 ммоль/л; 3) артериальное давление sys ≥ 130, dia ≥ 85 мм рт.ст.); 4) глюкоза натощак ≥ 6.1 ммоль/л. В подгруппу пациентов с МЗО вошло 57, а в подгруппу с МНЗО – 76 пациентов. В исследование были включены пациенты, проживающие на одной территории (Ростовская область и город Ростов-на-Дону) в летнее-осенний период, с целью минимизации влияния климатических условий, характера питания и этнических факторов на кишечный микробиом.

**Клинические образцы.** У пациентов собирали образцы фекалий и венозной крови, из которой впоследствии получали сыворотку.

**Иммуноферментный анализ (ИФА).** В образцах сыворотки определяли концентрацию TFF2 и TFF3 методом ИФА с использованием соответствующих наборов: ELISA Kit for Trefoil Factor 2 и ELISA Kit for Trefoil Factor 3 (“Cloud-Clone Corp.”, США) – по протоколу производителя.

**Выделение бактериальной ДНК, амплификация и секвенирование генов 16S рРНК.** Выделение бактериальной ДНК из образцов фекалий проводили с применением наборов QIAamp Fast DNA Stool Mini Kit (“QIAGEN GmbH”, Германия). Подготовку библиотек для секвенирования проводили согласно протоколу производителя: “16S Metagenomic Sequencing Library Preparation” (“Illumina, Inc.”, США). Бактериальную ДНК амплифицировали с использованием праймеров, специфичных к варибельному участку v3–v4 гена

16S рРНК, и набора с высокоточной полимеразой Q5® High-Fidelity DNA Polymerase (“NEB”, США) в стандартном режиме: 1) 30 с при 98°C; 2) 25 циклов: 30 с при 98°C, 20 с при 55°C, 20 с при 72°C; 3) 2 мин при 72°C; 4) длительно при 4°C. После очистки смеси от свободных праймеров и их димеров парамагнитными частицами AMPure XP (“Beckman Coulter”, США) ПЦР-продукты индексируют с использованием индекс-праймеров Nextera XT Index Kit (“Illumina, Inc.”). Повторно выполняли очистку смеси парамагнитными частицами, и сформированные библиотеки секвенировали на платформе MiSeq (“Illumina, Inc.”) по протоколу производителя. Полученные прочтения анализировали с использованием программы QIIME2 v.2020.8 [18] и референсной базы данных последовательностей генов 16S рРНК SILVA v.138 [19] с 97%-ным порогом кластеризации близких последовательностей.

**Биоинформатический анализ.** Для предсказания метаболических функций кишечной микробиоты использовано программное обеспечение PICRUSt2 [3], позволившее оценить представленность метаболических путей, нормализованную на число копий прочтений 16S рРНК.

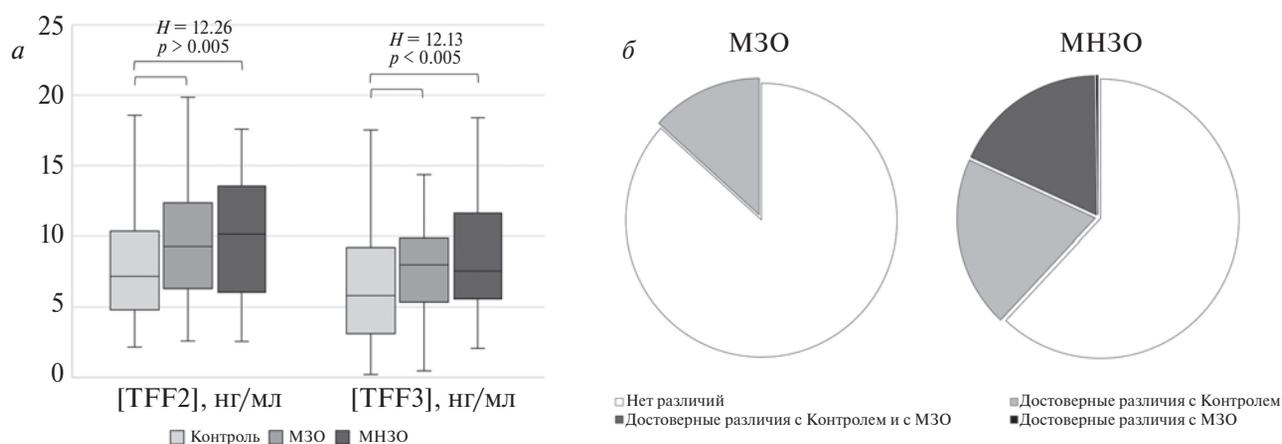
**Статистический анализ.** Статистическую обработку данных проводили с использованием программного обеспечения MedCalc v.20.210 (“MedCalc Software Ltd.”, Бельгия). Все полученные массивы данных были проверены на нормальность распределения с использованием критерия Шапиро-Уилка. Ввиду преобладания ненормального распределения данные были представлены в качестве медианы и квартилей (Me [Q1; Q3]). Сравнительный анализ выполняли с использованием непараметрического критерия Краскела-Уоллиса. При получении различий с уровнем значимости  $p < 0.05$  автоматически выполнялся

апостериорный тест по методу Conover для попарного сравнения групп. После получения результатов сравнительного анализа проводили контроль ложных открытий в соответствии с процедурой Бенджамини-Хохберга при частоте ложных открытий 5%. Кроме того, при необходимости сравнения частоты предсказания метаболических путей в кишечном микробиоме использовали анализ  $\chi^2$ . Выявленные различия считали статистически значимыми при уровне значимости  $p < 0.05$ . Корреляционный анализ проводили по методу Спирмена с последующей поправкой на множественную проверку гипотез процедурой Бенджамини-Хохберга при частоте ложных открытий 5%.

## РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

### Трефоилловые факторы

При исследовании уровня TFF в сыворотке крови выявлено, что у пациентов с МЗО и МНЗО, в сравнении с группой здоровых доноров, концентрация как TFF2, так и TFF3 повышена (рис. 1а). Высокий уровень TFF2 может быть связан с особенностями питания этих пациентов, так как пища с высоким содержанием жиров активирует экспрессию гена *TFF2* [20, 21]. У пациентов с ожирением вследствие избыточного питания также может происходить чрезмерное растяжение кишечника и увеличение площади его поверхности, что вносит вклад в избыточную продукцию TFF. Кроме того, повышенный уровень TFF может быть результатом повреждения стенки кишечника или формирования воспалительного отека слизистой оболочки у пациентов с ожирением – вне зависимости от метаболического фенотипа. Интересно, что при исследовании содержания TFF у детей с ожирением и с нормальным ИМТ различий не было выявлено [22]. Возможно, подобные различия обусловлены



**Рис. 1.** Анализ концентрации трефоилловых факторов в сыворотке крови (а) и метаболического профиля кишечной микробиоты (б) у пациентов с МЗО и МНЗО. Здесь и далее:  $H$  – значение  $H$ -критерия Краскела-Уоллиса,  $p$  – уровень значимости  $H$ -критерия Краскела-Уоллиса. Горизонтальные планки показывают различия между группами, выявленные с использованием апостериорного теста по методу Conover ( $p < 0.05$ ).

тем, что детское ожирение в большинстве случаев обратимо [23], а увеличение содержания TFF у взрослых можно рассматривать как маркер “не-обратимого” или прогрессирующего ожирения.

#### *Метаболический профиль кишечного микробиома*

Корреляционный анализ не выявил достоверных взаимосвязей между уровнем TFF в сыворотке и представленностью метаболических путей кишечного микробиома. По результатам биоинформатического анализа, в кишечном микробиоме пациентов, участвовавших в исследованиях, было предсказано 429 метаболических путей. Сравнительный анализ по методу Краскела–Уоллиса выявил, что изменение представленности при ожирении затронуло 163 метаболические пути. Как и ожидалось, наибольшее число изменений зарегистрировали у пациентов с МНЗО — 38% метаболических путей, тогда как при МЗО только у 13.30% (рис. 1б).

Для удобства описания и анализа метаболические пути были сгруппированы по видам обмена, в которых они принимают участие: энергетический обмен, синтез водорастворимых витаминов, синтез терпенов и хинонов, обмен нуклеотидов и процессинг нуклеиновых кислот, обмен железа и гема, обмен аминокислот, деградация ароматических соединений и синтез структурных компонентов клеток.

**Энергетический обмен.** И метаболически здоровый, и метаболически нездоровый фенотипы ожирения приводили к изменению в метаболических профилях распространенности путей, вовлеченных в энергетический обмен кишечной микробиоты (рис. 2). Большинство таких изменений регистрировали у пациентов с МНЗО, хотя снижение гликолиза V (*Pyrococcus*), характерного для архей, детектировали и у пациентов с МЗО.

Метаболически нездоровый фенотип ожирения сопровождался увеличением распространенности путей деградации различных энергетических субстратов: сульфоглюкозы, кетоглюконата, маннана, жирных кислот, этиленгликоля, глутаровой кислоты, гексита, глюкаратов и галакторатов – в метаболических профилях микробиоты. Кроме того, у таких пациентов наблюдали повышение различных вариантов цикла трикарбоновых кислот (ЦТК) — одного из основных метаболических путей, поставляющих восстановленные коферменты для окислительного фосфорилирования и, следовательно, для продукции АТФ. Эти результаты согласуются с данными Y. Wan и др. [24], которые показали, что увеличение концентрации интермедиатов ЦТК характерно для пациентов с ожирением. Кроме того, Hu J. с соавт. [4] обнаружили, что способность микробиоты к ЦТК VII ассоциирована с тяжестью ожирения. По-видимому, при

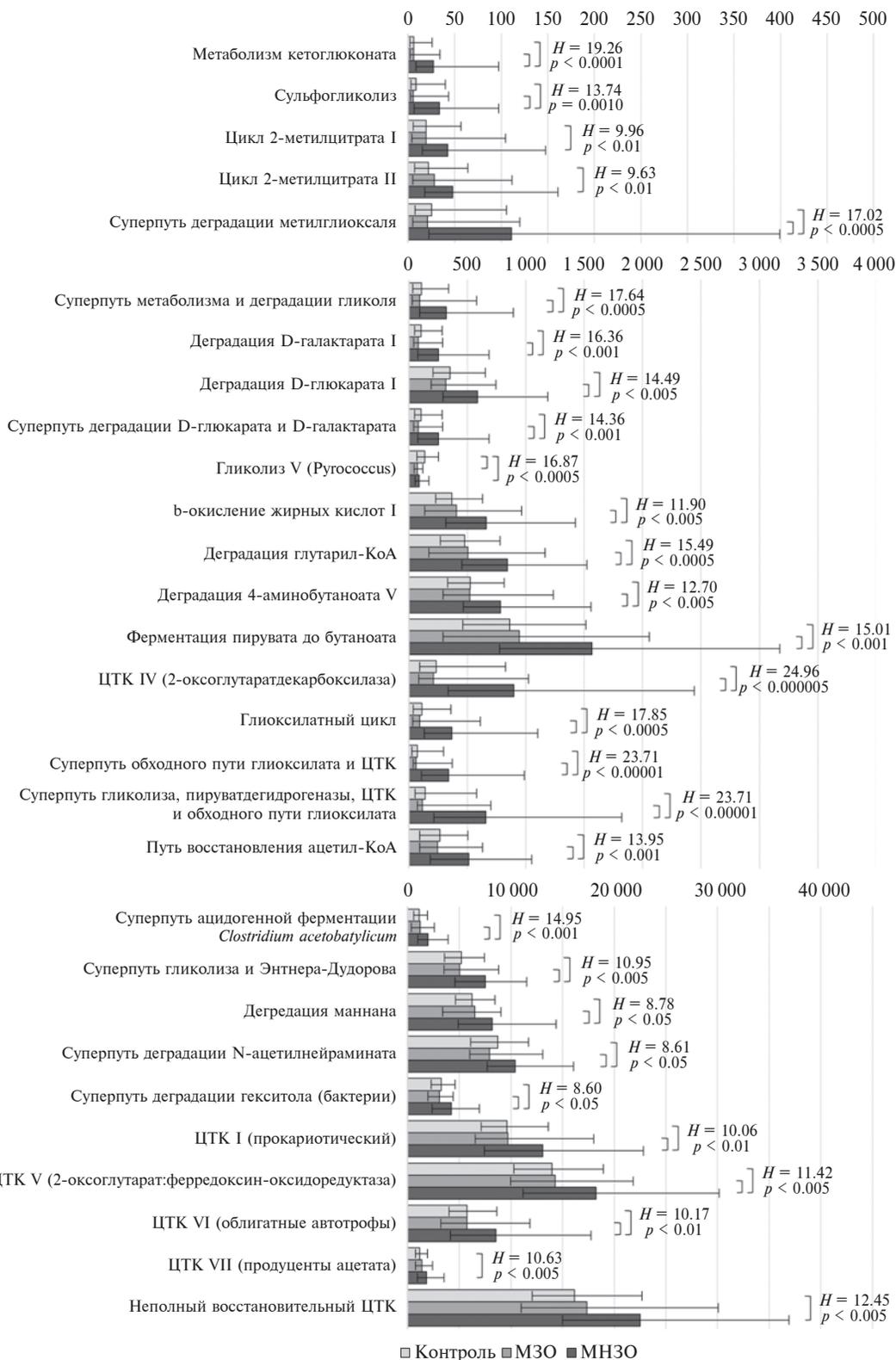
МНЗО микробиота более интенсивно деградирует энергетические субстраты, что способствует ее выживанию в условиях, специфичных для этого фенотипа ожирения.

Изменения метаболических профилей микробиоты при МНЗО затрагивают и процессы биосинтеза короткоцепочечных жирных кислот (SCFAs): ацетата, пропионата и бутирата. SCFAs, с одной стороны, значимые энергетические субстраты для организма хозяина, а с другой — медиаторы в системе взаимоотношений “микробиота–макроорганизм” [25–27]. В исследованиях по содержанию SCFAs в фекалиях показано, что ожирение ассоциировано с увеличением концентрации ацетата, пропионата и бутирата [28, 29]. Это может быть, с одной стороны, результатом диеты [30], а с другой — изменения метаболических возможностей микробиоты. Так, ацетат образуется в клетке на разных стадиях множества биохимических процессов, поэтому фактически невозможно отследить все метаболические пути его образования. Однако микробиота пациентов с МНЗО может активнее, чем микробиота пациентов с МЗО или лиц без ожирения, образовывать ацетат при окислении жирных кислот ( $\beta$ -окисление жирных кислот I) в ЦТК VII при деградации глутарил-КоА и N-ацетилнейрамата, а также образовывать бутират из пирувата и при деградации 4-аминобутаноата ( $\gamma$ -аминомасляной кислоты).

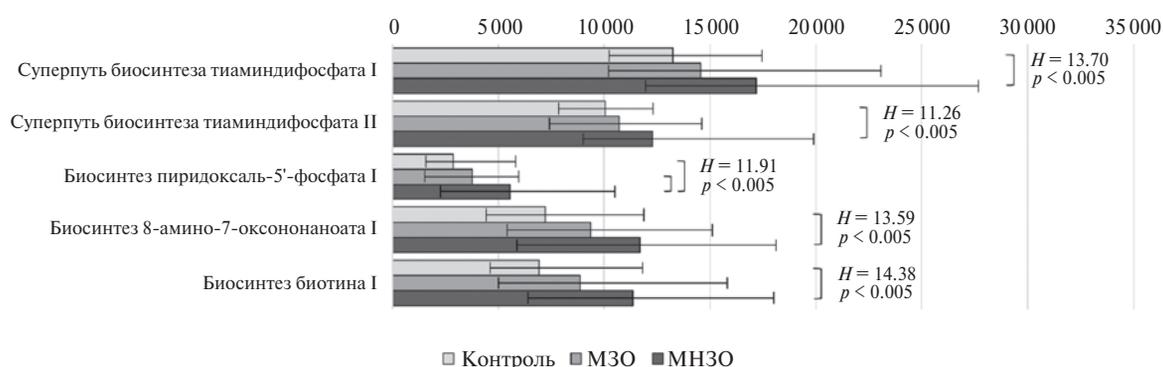
**Синтез водорастворимых витаминов.** Одна из ключевых функций микробиоты во взаимоотношениях с организмом-хозяином – продукция водорастворимых витаминов. У пациентов с ожирением в кишечной микробиоте чаще присутствовали таксоны, способные к продукции NAD из триптофана (биосинтез NAD II): 40.4% пациентов с МЗО, 44.7% пациентов с МНЗО против 26.2% здоровых доноров ( $p < 0.05$ ). Мы обнаружили, что метаболически нездоровый фенотип ожирения ассоциирован с изменением в метаболических профилях микробиоты распространенности путей синтеза ряда водорастворимых витаминов группы B (рис. 3).

Нами выявлено, что при МНЗО происходит интенсификация метаболических путей синтеза активных форм витаминов B1, B6 и B7. Эти результаты согласуются с данными M. Kim с соавт. [5] и M. Duan с соавт. [6], сообщавшими о повышенной продукции витаминов B1, B2, B5, B6, B9 и их активных форм в микробиоте пациентов с ожирением. Учитывая роль витаминов B1 и B7 в ЦТК, усиление их синтеза микробиотой согласуется с полученными данными по усилению энергетического обмена при МНЗО.

Ранее показано, что ожирение и связанные с ним метаболические нарушения могут быть ассоциированы с недостаточным или неоптимальным уровнем витаминов в сыворотке крови [31]. На этом фоне сохранение и увеличение



**Рис. 2.** Анализ путей энергетического обмена в метаболических профилях микробиома пациентов с МЗО и МНЗО. Здесь и далее: графики построены на основании медиан соответствующих метаболических процессов; в планках погрешностей – 25 и 75 перцентилей. ЦТК – цикл трикарбоновых кислот; КоА – коэнзим А. Здесь и далее: вертикальные планки показывают различия между группами, выявленные с использованием апостериорного теста по методу Conover ( $p < 0.05$ ).



**Рис. 3.** Анализ путей синтеза водорастворимых витаминов в метаболических профилях микробиома пациентов с МЗО и МНЗО.

распространенности путей синтеза водорастворимых витаминов микробиотой при ожирении, особенно при МНЗО, может носить позитивный характер для организма-хозяина. Кроме того, подобные изменения могут быть результатом разобщенности макроорганизма и микробиома, то есть потенциальными биомаркерами нарушения “метаболического симбиоза”.

Интересно, что ожирение, вне зависимости от метаболического фенотипа, сопровождалось увеличением частоты выявления в образцах фекалий таксонов, способных к биосинтезу *a*, *c*-диамида кобириновой II кислоты и аденозилкобаламина с поздним включением кобальта, т.е. продуцирующих витамин В12. По результатам биоинформатического анализа, наличие таких метаболических путей в кишечном микробиоме предсказано у 40.4% пациентов с МЗО, у 43.4% с МНЗО и только у 26.9% здоровых доноров без ожирения ( $p < 0.05$ ).

**Синтез терпенов и хинонов.** У пациентов с ожирением зарегистрирована активация в метаболических профилях микробиоты путей образования хинонов и их предшественников терпенов – наравне с витаминами они вовлечены в энергетический обмен и продукцию АТР (рис. 4).

Убихиноны (коэнзим Q), менахиноны и деметилменахинон служат промежуточными акцепторами электронов в электронтранспортных цепях и, следовательно, участвуют в генерации АТР у бактерий [32, 33]. Микробиота пациентов с МЗО и МНЗО характеризуется активацией путей синтеза менахинонов и деметилменахинона. Метаболически нездоровый фенотип ожирения ассоциирован и с увеличением распространенности путей синтеза убихинонов в метаболических профилях микробиоты. Менахиноны и деметилменахинон преимущественно участвуют в анаэробном дыхании бактерий, тогда как убихиноны — в аэробном [33]. Кроме того, менахиноны, известные как витамин К2, абсорбируются организмом-хозяином и используются в реакциях  $\gamma$ -карбоксилирования [32].

Метаболические профили обоих фенотипов ожирения характеризуются увеличением путей синтеза филлохинона и менахинонов, а МНЗО еще и путей синтеза убихинонов. В целом, полученные результаты согласуются с данными M. Duan с соавт. [6], которые продемонстрировали большую представленность путей синтеза хинонов при ожирении. Эти изменения могут быть связаны с увеличенной потребностью микробиоты пациентов с ожирением, особенно с МНЗО, в переносчиках электронов и косвенно с повышенной активностью электронтранспортных цепей и продукции АТР.

По сравнению со здоровыми донорами, у пациентов с МНЗО прогнозируется повышенная активность путей синтеза предшественников хинонов и терпенов: изопренов и таксадиенов. Ранее J. Hu с соавт. [4] тоже сообщали об изменении пути биосинтеза таксадиенов (инженерный) у пациентов с ожирением. Значение таксадиена для микробиоты на сегодняшний день остается неясным, хотя этот дитерпен используют для синтеза противоопухолевых препаратов [34].

**Обмен нуклеотидов и процессинг нуклеиновых кислот.** Изменения метаболического профиля микробиоты у пациентов с ожирением затронули и обмен нуклеотидов (рис. 5). У пациентов с МЗО этот паттерн не отличался значительно от здоровых доноров, в то время как при МНЗО регистрировали активацию путей обмена нуклеотидов. Кроме того, при МНЗО увеличивалась продукция уратов – логичное следствие повышенной продукции пуринов. Полученные данные согласуются с результатами M. Kim с соавт. [5], которые показали, что при МНЗО увеличена представленность путей синтеза нуклеотидов.

Кроме того, у пациентов с ожирением (и МЗО, и МНЗО) полностью отсутствовал суперпуть деградации пиримидиновых нуклеотидов, тогда как в когорте здоровых доноров он составлял 7.69% ( $p < 0.05$ ).

Нуклеотиды – ключевые компоненты процессов репликации, репарации, транскрипции

и трансляции, то есть основных механизмов функционирования клетки и ее деления [35]. Доступность нуклеотидов способствует увеличению разнообразия кишечной флоры, что было показано на модели мышей, склонных к ускоренному старению — Senescent Accelerated Mouse-Prone 8 (SAMP8) [36]. По-видимому, активация путей синтеза и реутилизации нуклеотидов в микробиоте пациентов с МНЗО способствует выживанию микроорганизмов в условиях, способствующих истощению кишечной флоры [5, 37].

Изменения метаболических возможностей микробиоты при ожирении затронули и ряд метаболических процессов, вовлеченных в процессинг нуклеиновых кислот (рис. 6).

У пациентов с ожирением зарегистрирована активация пути синтеза кевозина и его предшественника ргеQ0. Заметим, что продукция кевозина в микробиоте возрастала в следующем ряду: здоровые доноры → пациенты с МЗО → пациенты с МНЗО. Обогащение кишечной микробиоты кевозинпродуцирующими бактериями ранее обнаружили

у мышей с ожирением [38]. Кевозин — производное гуанозина, которое эукариоты не синтезируют самостоятельно, а получают либо с пищей, либо благодаря синтезу кишечной микробиотой. Модификации кевозином подвергаются тРНК аспарагиновой кислоты, аспарагина, тирозина и гистидина, защищая их от действия РНКаз [39]. Таким образом, доступность кевозина контролирует скорость трансляции белков в клетках эукариот, а его дефицит сопровождается дисфолдингом и развитием стресса эндоплазматического ретикулума [40].

Кроме того, в кишечном микробиоме пациентов и с МЗО, и с МНЗО выявлено увеличение представленности путей синтеза (p)ppGpp. Подобное влияние ожирения на синтез ppGpp было также продемонстрировано X. Nie с соавт. [41]. Синтез и накопление алармона (p)ppGpp — высококонсервативная реакция на стресс, которая способствует выживанию бактерий [42]. Таким образом, сохранение в кишечном микробиоме ppGpp-продуцирующих организмов может отражать агрессивные

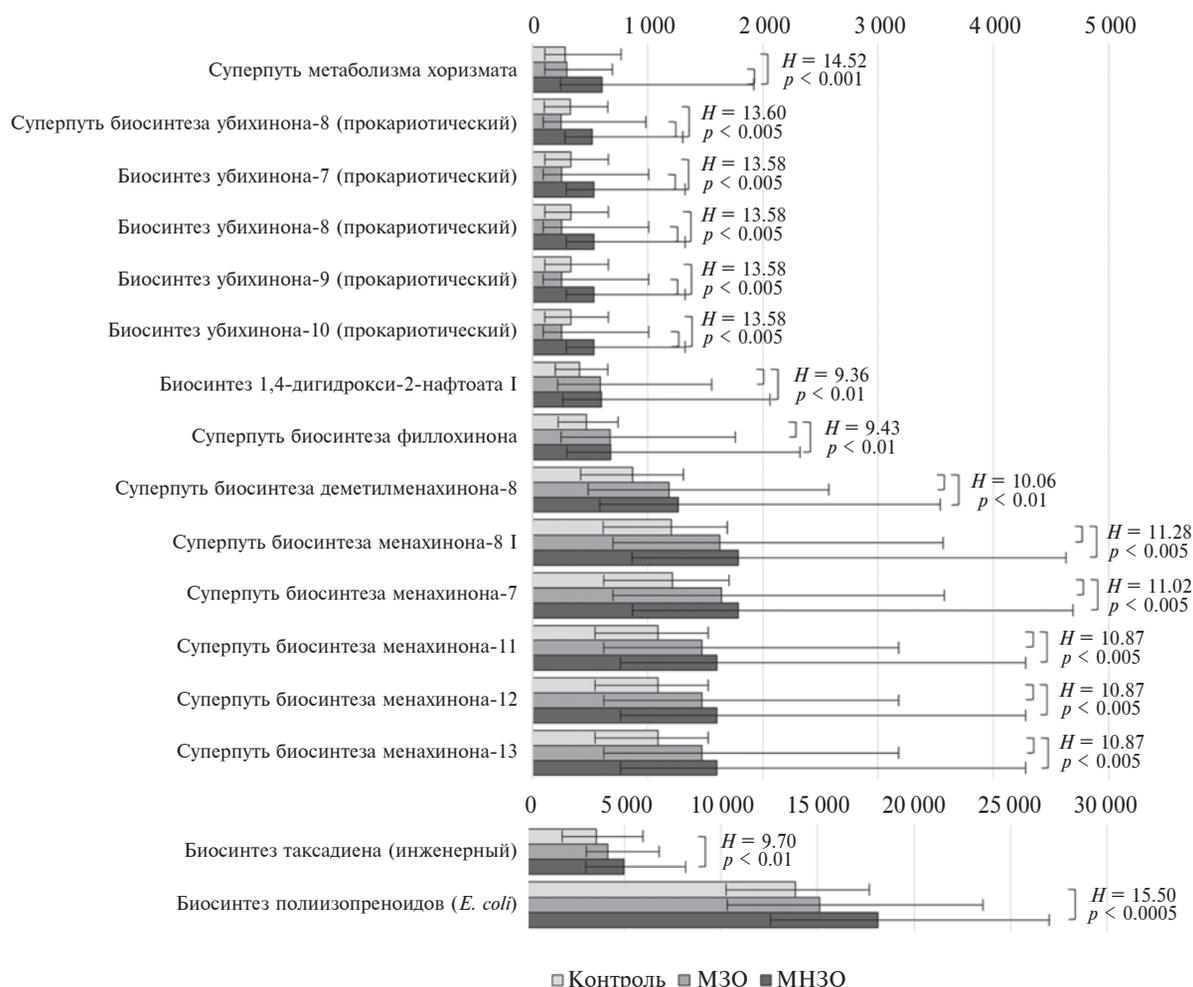


Рис. 4. Анализ путей биосинтеза терпенов и хинонов в метаболических профилях микробиома пациентов с МЗО и МНЗО.

условия, в которых находится микробиота кишечника при ожирении.

Также особенностью кишечного микробиома при МНЗО была бóльшая представленность пути процессинга тРНК по сравнению со здоровыми донорами и пациентами с МЗО. Следует отметить, что именно этот метаболический путь был выделен среди процессов, отличающих микробиом пациентов с сахарным диабетом II типа [43]. Подобное перерождение микробиоты может вносить свой вклад в формирование вялотекущего системного воспаления у пациентов с МНЗО. Бактерии способны продуцировать внеклеточные везикулы, содержащие в том числе и тРНК, которые могут участвовать в индукции воспаления [44, 45].

**Обмен железа и гема.** МЗО и МНЗО сопровождалась изменением в метаболических профилях микробиоты доли путей синтеза гема, а также аэробактина и энтеробактина — белков, способствующих захвату железа (рис. 7). Оба фенотипа ожирения приводили к увеличению метаболических путей синтеза гема, однако при МНЗО такие изменения были более выражены. Вместе с тем синтез аэробактина и энтеробактина был усилен только у пациентов с МНЗО.

Активизация метаболических путей биосинтеза гема в кишечном микробиоме может быть связана с повышенной активностью цепи переноса электронов в микробиоме при ожирении и, как следствие, с повышением потребности в геме. Как отмечено выше, об усилении процессов общего пути катаболизма и окислительного фосфорилирования при ожирении (особенно при МНЗО) можно судить по увеличению прогнозируемой представленности синтеза убихинонов и менахинонов — транспортеров цепи переноса электронов, — а также ЦТК.

Усиление синтеза гема может негативно сказываться на микробном сообществе. Ранее показано, что добавление гема в рацион мышей приводит к сокращению разнообразия кишечного микробиома и усугубляет течение колита за счет модуляции кишечной микробиоты [46]. Кроме того, гем вызывает повреждение эпителия, что повышает риск развития колоректального рака [47].

**Обмен аминокислот.** У пациентов с МНЗО наблюдалось изменение представленности ряда метаболических процессов обмена аминокислот (рис. 8), в то время как при МЗО это было менее выражено. В кишечной микробиоте пациентов с МЗО была повышена представленность пути

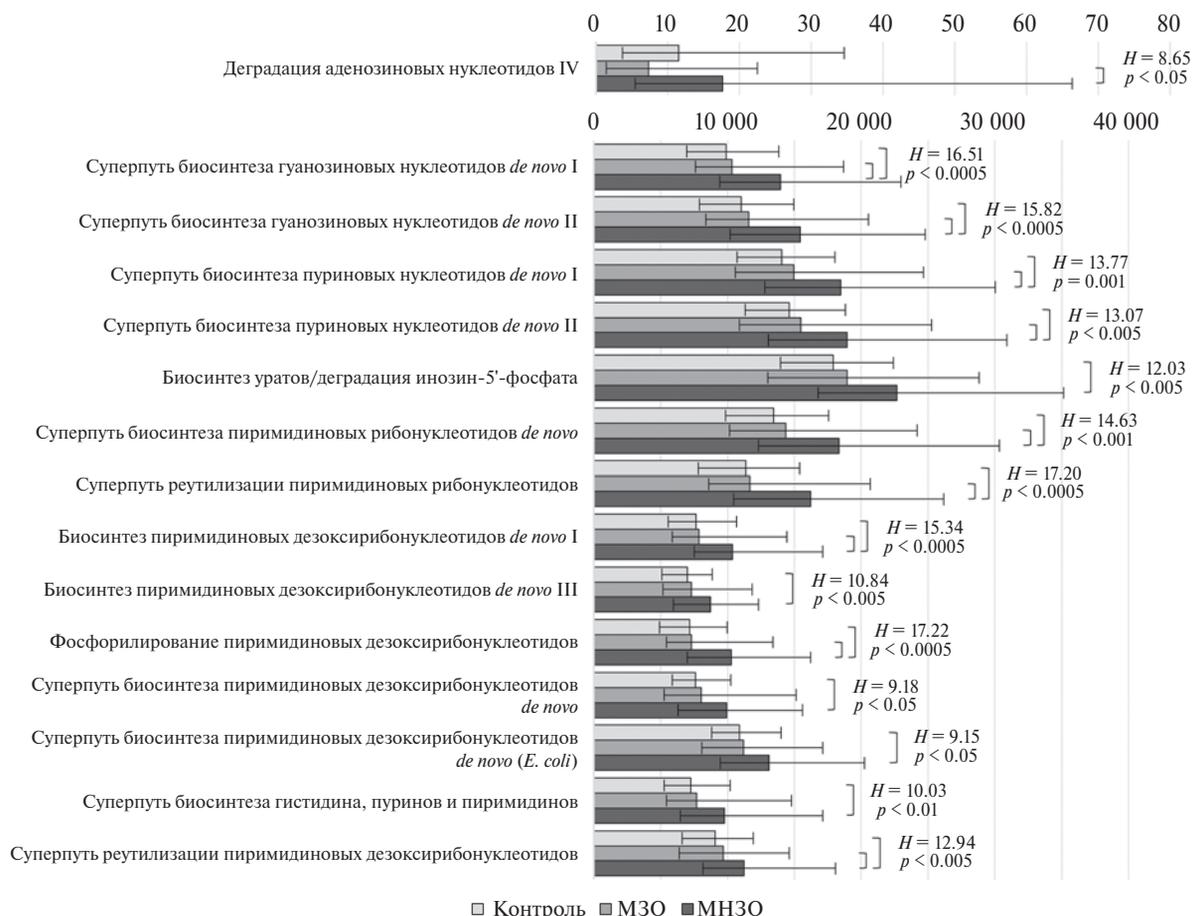


Рис. 5. Анализ путей метаболизма нуклеотидов в метаболических профилях микробиома пациентов с МЗО и МНЗО.

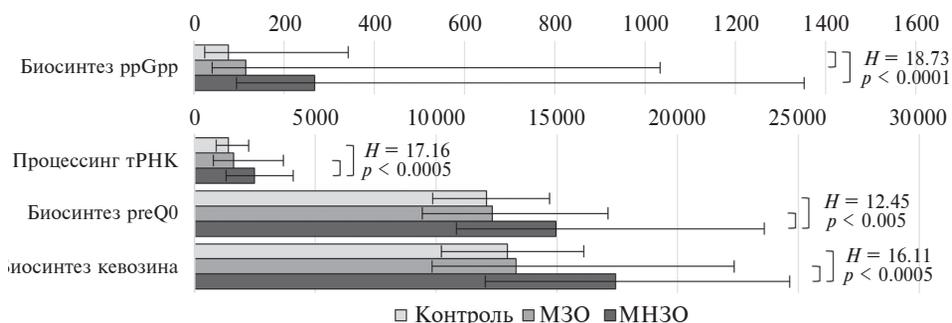


Рис. 6. Анализ путей процессинга нуклеиновых кислот в метаболических профилях микробиома пациентов с МЗО и МНЗО.

деградации тирозина I до ацетоацетата и фумарата, что отражает “повышенный спрос” на использование этой аминокислоты в качестве энергетического субстрата. Кроме того, у них чаще прогнозировали наличие различных путей деградации триптофана: деградации триптофана IX до ацетил-КоА у 22.80% пациентов, деградации триптофана до 2-амино-3-карбоксимуконата полуальдегида у 40.35% и триптофана XII (*Geobacillus*) до ацетил-КоА у 21.05% пациентов – по сравнению с соответственно 8.46, 26.15 и 8.46% у здоровых волонтеров без ожирения ( $p < 0.05$ ). Таким образом, в микробиоте кишечника пациентов с МЗО повышена возможность метаболической деградации ароматических аминокислот и, как следствие, использования продуктов их метаболизма в качестве энергетических субстратов.

В микробиоте пациентов с МНЗО также чаще прогнозировали метаболические пути деградации триптофана IX (21.05% пациентов), триптофана до 2-амино-3-карбоксимуконата полуальдегида (44.74% пациентов) и триптофана XII (*Geobacillus*) (21.05% пациентов), что достоверно отличалось от здоровых доноров без ожирения ( $p < 0.05$ ). Кроме триптофана, в их микробиоте потенциально

усилена деградация аминокислот тирозина, треонина, гистидина и аргинина.

На фоне повышенной деградации аминокислот, в частности аргинина, у пациентов с МНЗО выявлено увеличение представленности путей синтеза полиаминов: путресцина, спермидина и норспермидина. Ранее показано, что интенсивность продукции полиаминов зависит от степени ожирения [4]. Роль полиаминов в контексте ожирения неоднозначна. С одной стороны, показано, что введение спермидина и спермина мышам способствовало снижению массы тела, повышению толерантности к глюкозе и уменьшению стеатоза печени в модели ожирения, вызванного диетой с высоким содержанием жиров [48]. С другой стороны, известно, что полиамины участвуют в адипогенезе, а повышенный уровень путресцина при сахарном диабете коррелирует с концентрацией гликированного гемоглобина [49].

Следует отметить, что у пациентов с МНЗО повышенная деградация аминокислот была компенсирована увеличением представленности метаболических процессов их синтеза, в частности аргинина и ароматических аминокислот (фенилаланина, тирозина и триптофана). Это согласуется

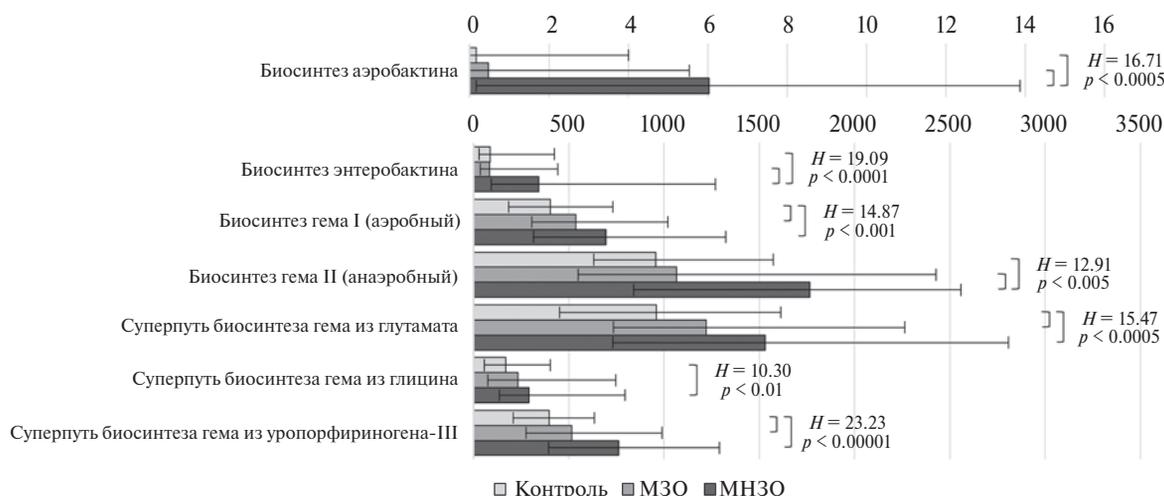


Рис. 7. Анализ путей обмена железа и гема в метаболических профилях микробиома пациентов с МЗО и МНЗО.

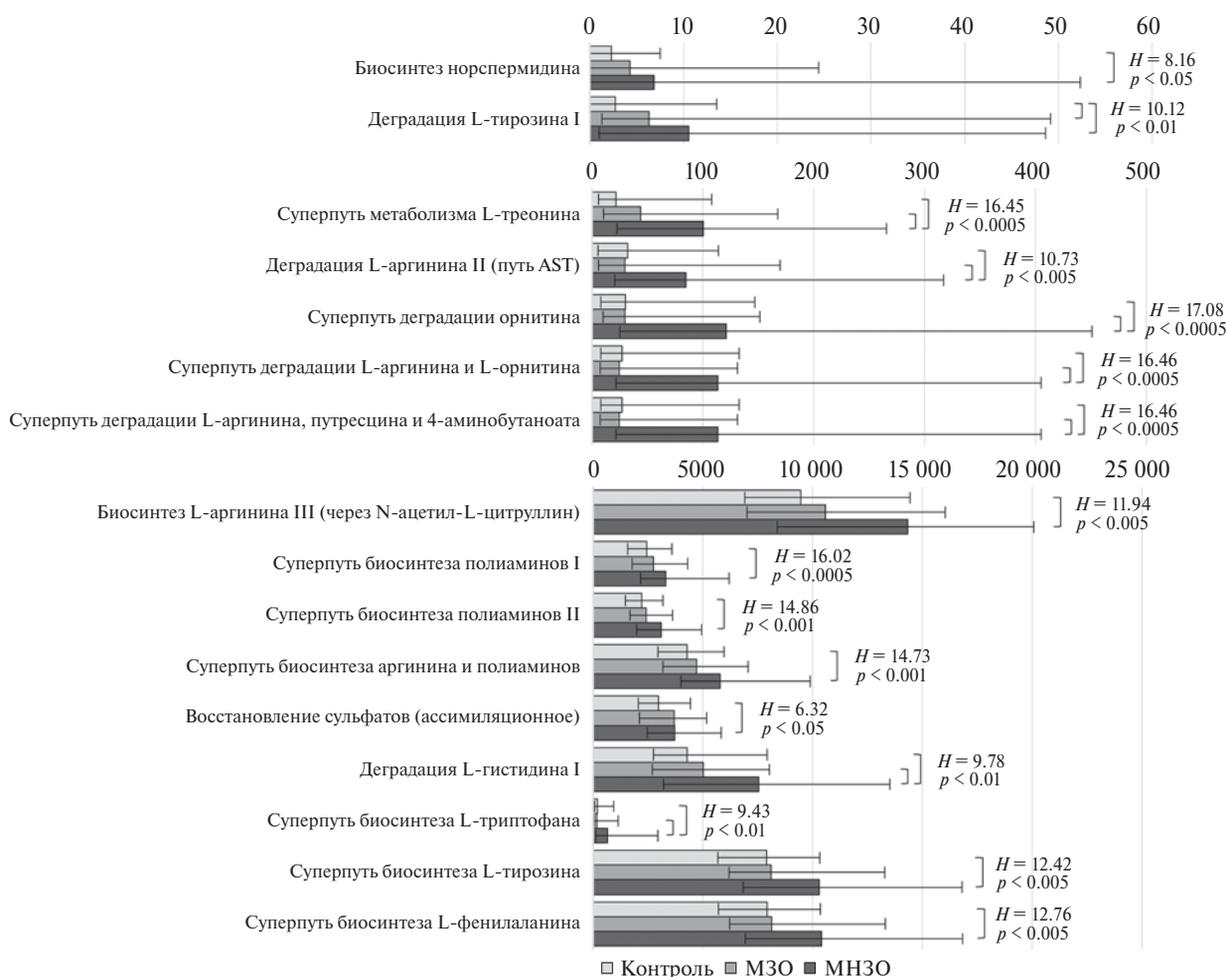


Рис. 8. Анализ путей обмена аминокислот в метаболических профилях микробиома пациентов с МЗО и МНЗО.

с ранее опубликованными данными по биосинтезу ароматических аминокислот в кишечной микробиоте пациентов с ожирением [4, 50]. Также при ожирении регистрировали снижение концентрации антраниловой кислоты в фекалиях пациентов с ожирением [51], что может быть ассоциировано с ее потреблением в качестве субстрата синтеза триптофана. Кроме того, у 35.53% пациентов с МНЗО мы прогнозировали наличие пути дегградации креатинина I, вовлеченного в биосинтез глицина, в то время как этот путь предсказали только для 12.31% здоровых волонтеров ( $p < 0.05$ ). Также при МНЗО нами отмечено увеличение представленности пути восстановления сульфатов, которые впоследствии могут включаться в синтез содержащих серу аминокислот: цистеина и метионина.

В целом, ожирение, особенно метаболически нездорового фенотипа, по-видимому, ассоциировано с повышенным использованием аминокислот в качестве энергетических субстратов в микробиоте кишечника.

**Дегградация ароматических соединений.** Мы обнаружили целый спектр метаболических путей

деградации ароматических соединений, представленность которых была изменена у пациентов с ожирением (табл. 1). Ряд таких метаболических путей был характерен менее чем для 25% обследованных (0.000 [0.000; 0.000] для всех групп), поэтому мы сравнили частоту их прогнозирования в кишечном микробиоме (рис. 9).

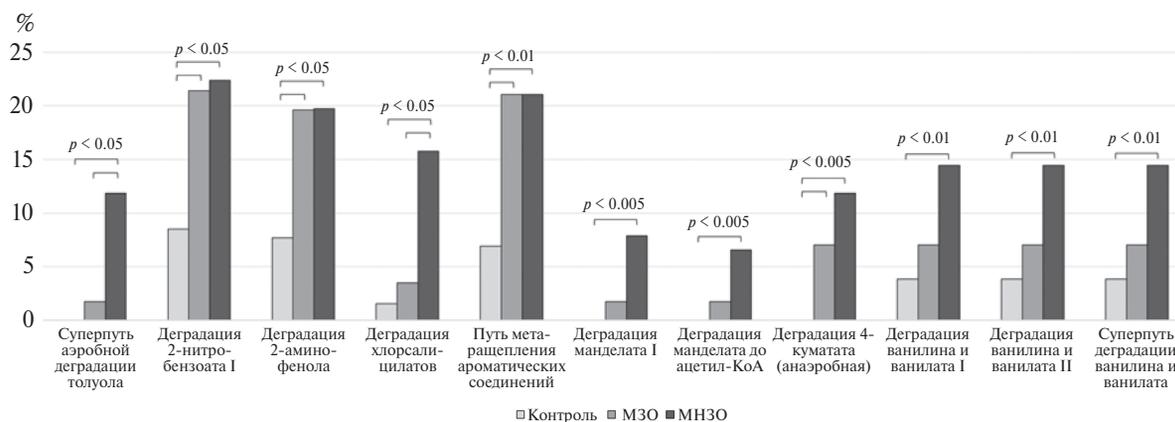
В микробиоме пациентов с ожирением, особенно с метаболически нездоровым фенотипом, обнаружена повышенная дегградация ароматических соединений по сравнению со здоровыми волонтерами. Фенилацетат, 4-гидроксифенилацетат, фенилпропионат, фенилэтиламин образуются при дегградации фенилаланина и тирозина. Увеличение представленности этих метаболических путей, по-видимому, является следствием усиления биосинтеза ароматических аминокислот, характерного для пациентов с МНЗО, и может отражать менее эффективное использование этих аминокислот в качестве пластических субстратов.

По результатам большого проспективного исследования, проведенного среди городских жителей Китая, повышенная дегградация ароматических

**Таблица 1.** Изменения прогнозируемой представленности путей деградации ароматических соединений у пациентов с разными фенотипами ожирения

Метаболический путь	Здоровые доноры	МЗО	МНЗО	Критерий Кра-скела-Уоллиса
Деградация 4-гидроксифенилацетата	2.58 [0.24; 21.99]	5.20 [1.29; 34.75]	17.03 [3.83; 81.51]**†	$H = 20.68$ $p < 0.00005$
Деградация фенилацетата I (аэробная)	3.21 [0.50; 36.69]	5.99 [1.95; 45.39]	24.14 [5.45; 114.47]**†	$H = 19.32$ $p < 0.0001$
Деградация 3-фенилпропионата и 3-(3-гидроксифенил)пропионата до 2-оксопент-4-еноата	3.26 [0.64; 24.42]	5.91 [1.16; 35.31]	24.44 [3.87; 64.74]**†	$H = 18.52$ $p < 0.0001$
Деградация 3-фенилпропионата и 3-(3-гидроксифенил)пропионата	8.07 [1.98; 49.59]	12.97 [3.98; 72.61]	47.93 [9.50; 124.04]**†	$H = 21.45$ $p < 0.00005$
Деградация 3-фенилпропионата	3.97 [0.61; 11.96]	3.76 [0.66; 13.76]	7.39 [2.00; 38.44]**†	$H = 9.10$ $p < 0.05$
Суперпуть деградации фенилэтиламина	2.64 [0.42; 31.77]	5.07 [1.84; 40.65]	20.67 [4.89; 100.86]**†	$H = 18.85$ $p < 0.0001$
Деградация толуола I (аэробная) (через <i>o</i> -крезол)	66.35 [30.01; 143.21]	82.88 [47.64; 221.89]*	189.30 [66.40; 320.54]**†	$H = 20.04$ $p < 0.0005$
Деградация толуола II (аэробная) (через 4-метилкатехол)	66.35 [30.01; 143.21]	82.88 [47.64; 221.89]	189.30 [66.40; 320.54]**†	$H = 20.04$ $p < 0.0005$
Деградация толуола III (аэробная) (через <i>n</i> -крезол)	0.43 [0.00; 2.94]	1.99 [0.00; 10.37]*	2.95 [0.00; 23.71]**†	$H = 10.94$ $p < 0.005$
Деградация толуола IV (аэробная) (через катехол)	0.00 [0.00; 0.76]	0.00 [0.00; 1.63]*	0.00 [0.00; 5.15]*	$H = 6.15$ $p < 0.05$
Деградация катехола I (путь <i>мета</i> -расщепления)	15.62 [4.85; 37.02]	16.41 [7.90; 42.82]*	31.00 [11.15; 79.39]*	$H = 6.15$ $p < 0.01$
Деградация катехола II (путь <i>мета</i> -расщепления)	0.00 [0.00; 0.000]	0.00 [0.00; 1.17]*	0.00 [0.00; 2.91]*	$H = 5.64$ $p < 0.005$
Деградация катехола III (путь <i>орто</i> -расщепления)	0.22 [0.00; 1.94]	1.17 [0.14; 7.30]*	1.81 [0.00; 22.02]*	$H = 11.91$ $p < 0.05$
Деградация катехола до $\beta$ -кетoadипата	0.12 [0.00; 1.70]	0.92 [0.00; 6.52]*	1.00 [0.00; 23.65]*	$H = 11.41$ $p < 0.005$
Деградация катехола до 2-оксопент-4-еноата II	0.00 [0.00; 0.00]	0.00 [0.00; 1.21]*	0.00 [0.00; 1.88]*	$H = 5.73$ $p < 0.005$
Деградация 4-метилкатехола ( <i>орто</i> -расщепление)	0.16 [0.00; 1.79]	1.33 [0.00; 7.66]*	1.85 [0.00; 16.51]*	$H = 13.26$ $p < 0.001$
Суперпуть деградации салицилата	0.19 [0.00; 1.67]	1.34 [0.09; 7.30]*	1.71 [0.00; 19.38]*	$H = 13.20$ $p < 0.001$
Деградация галлата I	0.00 [0.00; 1.31]	0.61 [0.00; 2.86]*	0.86 [0.00; 9.66]*	$H = 8.21$ $p < 0.05$
Деградация циннамата и 3-гидроксициннамата до 2-оксопент-4-еноата	3.26 [0.64; 24.42]	5.91 [1.16; 35.31]	24.44 [3.87; 64.74]**†	$H = 18.52$ $p < 0.0001$
Деградация ароматических соединений через $\beta$ -кетoadипата	0.22 [0.00; 1.94]	1.17 [0.14; 7.30]*	1.81 [0.00; 22.02]*	$H = 11.91$ $p < 0.005$
Деградация протокатехуата II (путь <i>орто</i> -расщепления)	2.23 [0.00; 11.21]	5.09 [0.54; 19.12]	7.08 [0.50; 61.77]*	$H = 8.45$ $p < 0.05$

\*Различия статистически значимы по сравнению со здоровыми донорами (апостериорный тест Conover,  $p < 0.05$ ); † различия статистически значимы по сравнению с пациентами с МЗО (апостериорный тест Conover,  $p < 0.05$ ).



**Рис. 9.** Анализ распространенности путей деградации ароматических соединений в метаболических профилях микробиома пациентов с МЗО и МНЗО.

соединений, таких как толуол или катехол, характерна для микробиома лиц, практикующих здоровое и разнообразное питание [52]. Толуол может поступать с пищевыми продуктами, например с газированными безалкогольными напитками [53]. Кроме того, толуол образуется при декарбоксилировании фенилацетата — продукта деградации фенилаланина [54] — или из бензойной кислоты — распространенного консерванта с антибактериальной активностью [55]. Производные галлата активно используют в пищевой промышленности в качестве антиоксиданта [56]. Салициловая, ванилиновая, коричневая, миндальная кислоты и их производные также присутствуют в пищевых продуктах. Интересно, что ванилин повышает чувствительность к инсулину, способствует снижению уровня провоспалительных цитокинов и препятствует снижению разнообразия кишечного микробиома при ожирении [57]. Таким образом, повышенная представленность путей деградации ванилатов в кишечном микробиоме пациентов с МНЗО можно рассматривать как потенциально неблагоприятный фактор. Кроме того, ранее мы обнаружили наличие путей деградации катехола в микробиоме здоровых детей и подростков, находившихся в младенческом возрасте на смешанном и искусственном вскармливании [58].

Таким образом, усиление метаболических возможностей деградации ароматических соединений можно объяснить особенностью питания таких пациентов — избыточного и гиперкалорийного.

**Синтез структурных компонентов клеток.** Среди метаболических путей, доля которых изменена у пациентов с ожирением, были процессы, вовлеченные в синтез структурных компонентов бактериальных клеток — углеводов и липидов (рис. 10).

Кишечная микробиота пациентов с МЗО и МНЗО обладала повышенным метаболическим потенциалом в продукции жирных кислот (пальмитиновой, стеариновой, пальмитолеиновой, олеиновой и миколовых кислот). Кроме того,

изменения затронули и ряд метаболических путей синтеза производных сахаров — компонентов липополисахаридов, пептидогликанов и других структур бактериальных клеток. Интересно, что представленность путей образования компонентов липополисахаридов липида IVA и CMP-3-дезоксид-манно-октулозоната (известного как Kdo), а также их объединение последовательно возрастала в ряду: здоровые доноры → пациенты с МЗО → пациенты с МНЗО. Кроме того, у пациентов с МНЗО повышена активность суперпути синтеза липополисахарида и резистентности к полимиксину. Подобные различия могут способствовать формированию вялотекущего системного воспаления и воспаления жировой ткани, более характерного для пациентов с МНЗО [14, 59].

Главная функция липополисахаридов, пептидогликанов и других компонентов оболочки бактериальных клеток заключается в защите бактерии [60, 61]. Увеличение потребности кишечной микробиоты в таких структурах может отражать агрессивные условия среды ее обитания, создаваемые ожирением, особенно метаболически нездорового фенотипа.

Ожирение у взрослых, вне зависимости от метаболического фенотипа, сопровождается повышением концентрации TFF2 и TFF3 в плазме крови, что отличает их от детей с ожирением. Мы не обнаружили значимых ассоциаций между уровнем TFF и метаболическими путями кишечной микробиоты, что оставляет открытым вопрос о влиянии метаболического профиля микробиоты на состояние кишечной стенки.

Функциональный профиль микробного сообщества ожидаемо был более изменен у пациентов с метаболически нездоровым фенотипом ожирения, чем у пациентов с метаболически здоровым. У пациентов с МНЗО кишечная микробиота отличается большими метаболическими возможностями катаболизма энергетических субстратов, в том числе аминокислот, продукции АТФ в цикле

Кребса и цепи переноса электронов, синтеза хинонов, витаминов В1, В6, В7, захвата железа и синтеза гема, синтеза нуклеотидов, ароматических аминокислот и деградации ароматических соединений. Кроме того, в микробиоте этих пациентов усилен синтез защитных компонентов клеточной стенки, в частности липополисахаридов.

Ранее нами показано [37, 62], что при МНЗО снижено разнообразие кишечной флоры, в то время как при МЗО увеличено. Возможно, при МЗО повышенная способность кишечной микробиоты к синтезу гема, хинонов способствует лучшему энергообеспечению бактериальных клеток и дает им метаболические преимущества, способствуя тем самым увеличению микробного разнообразия.

МНЗО в свою очередь, по-видимому, сопровождается формированием агрессивных условий для кишечного микробиома, что, с одной стороны, истощает кишечную флору, а с другой – выступает мощным фактором селекции микробиома. Изменение метаболического профиля микробиоты при МНЗО может быть залогом выживания бактерий. В таких условиях выживание микробных клеток зависит от их способности к пролиферации (например, интенсивный синтез нуклеотидов), продукции энергии и синтезу защитных компонентов клеток. Однако подобное перерождение кишечного микробиома может вносить вклад в развитие метаболических нарушений. Таким образом, может происходить формирование порочного круга: метаболически нездоровое ожирение способствует

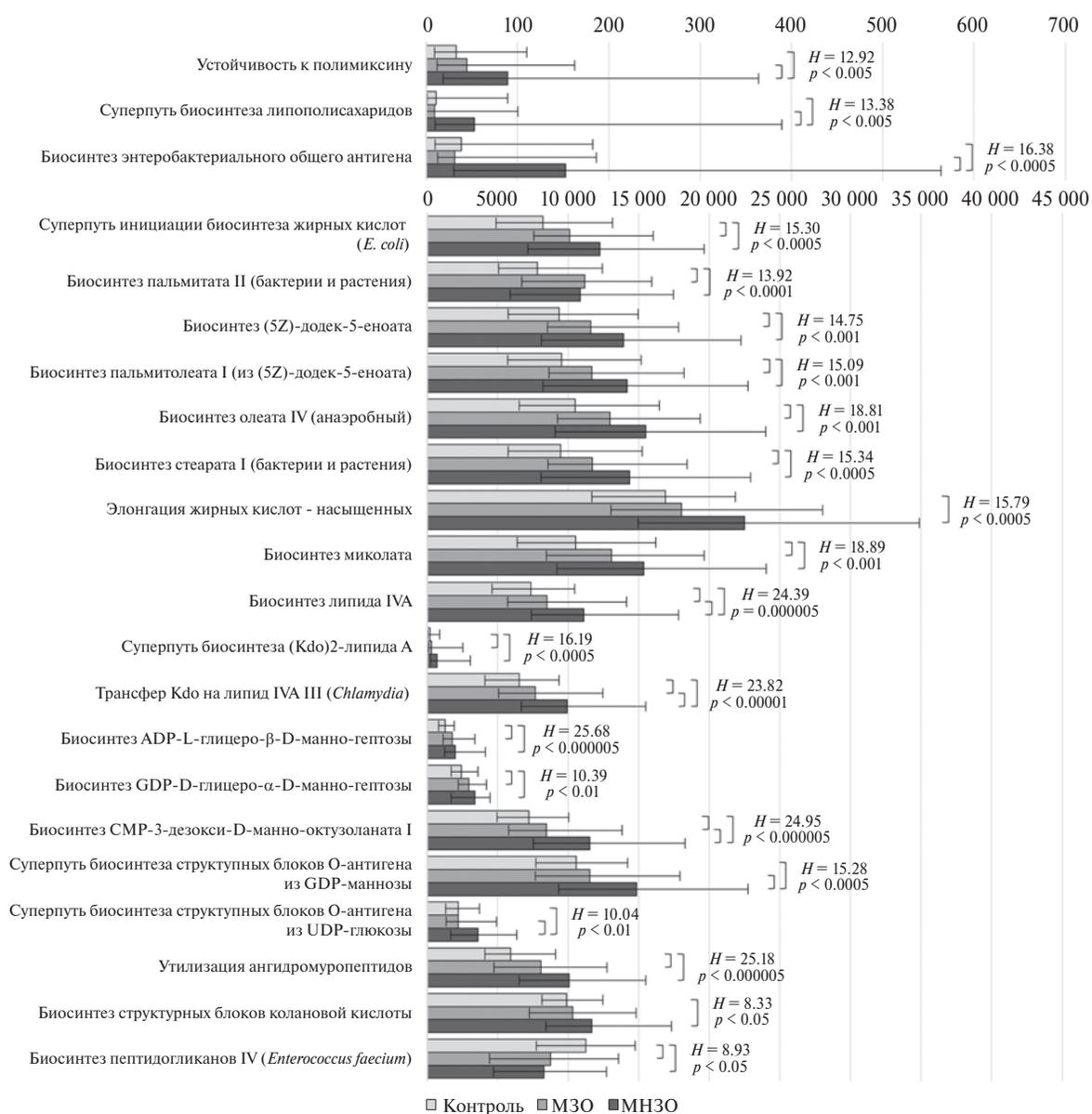


Рис. 10. Анализ путей синтеза структурных компонентов в метаболических профилях микробиома пациентов с МЗО и МНЗО.

истощению кишечного микробиома, дальнейшее перерождение которого вносит вклад в патогенез формирования метаболических нарушений у таких пациентов.

Статья подготовлена на основании результатов, полученных в ходе реализации «Соглашения о предоставлении гранта в форме субсидии из федерального бюджета на осуществление государственной поддержки создания и развития научных центров мирового уровня, выполняющих исследования и разработки по приоритетам научно-технологического развития» от 20 апреля 2022 года (№ 075-15-2022-310).

Все процедуры, выполненные в исследовании с участием людей, соответствуют этическим стандартам Национального комитета по исследовательской этике и Хельсинской декларации 1964 года и ее последующим изменениям или сопоставимым нормам этики. Проведение научно-исследовательской работы одобрено Локальным этическим комитетом (ЛЭК) ФГБОУ ВО РНИМУ им. Н.И. Пирогова Минздрава России (протокол № 186 от 26.06.2019) и Локальным независимым этическим комитетом (ЛНЭК) ФГБОУ ВО РостГМУ Минздрава России (протокол № 20/19 от 12.12.2019). От каждого из включенных в исследование участников было получено информированное добровольное согласие.

Авторы заявляют об отсутствии конфликта интересов.

## СПИСОК ЛИТЕРАТУРЫ

- Cheng Z., Zhang L., Yang L., Chu H. (2022) The critical role of gut microbiota in obesity. *Front. Endocrinol.* (Lausanne). **13**, 1025706. <https://doi.org/10.3389/FENDO.2022.1025706>
- Van Hul M., Cani P.D. (2023) The gut microbiota in obesity and weight management: microbes as friends or foe? *Nat. Rev. Endocrinol.* **19**(5), 258–271. <https://doi.org/10.1038/s41574-022-00794-0>
- Douglas G.M., Maffei V.J., Zaneveld J.R., Yurgel S.N., Brown J.R., Taylor C.M., Huttenhower C., Langille M.G.I. (2020) PICRUSt2 for prediction of metagenome functions. *Nat. Biotechnol.* **38**, 685–688. <https://doi.org/10.1038/s41587-020-0548-6>
- Hu J., Guo P., Mao R., Ren Z., Wen J., Yang Q., Yu J., Zhang T., Liu Y., Yan T. (2022) Gut microbiota signature of obese adults across different classifications. *Diabetes Metab. Syndr. Obes.* **15**, 3933–3947. <https://doi.org/10.2147/dms0.S387523>
- Kim M.H., Yun K.E., Kim J., Park E., Chang Y., Ryu S., Kim H.L., Kim H.N. (2020) Gut microbiota and metabolic health among overweight and obese individuals. *Sci. Rep.* **10**(1), 19417. <https://doi.org/10.1038/S41598-020-76474-8>
- Duan M., Wang Y., Zhang Q., Zou R., Guo M., Zheng H. (2021) Characteristics of gut microbiota in people with obesity. *PLoS One.* **16**(8), e0255446. <https://doi.org/10.1371/journal.pone.0255446>
- Takiishi T., Fenero C.I.M., Câmara N.O.S. (2017) Intestinal barrier and gut microbiota: shaping our immune responses throughout life. *Tissue Barriers.* **5**(4), e1373208. <https://doi.org/10.1080/21688370.2017.1373208>
- Portincasa P., Bonfrate L., Khalil M., De Angelis M., Calabrese F.M., D'amato M., Wang D.Q.-H., Di Ciaula A. (2022) Intestinal barrier and permeability in health, obesity and NAFLD. *Biomedicines.* **10**(1), 83. <https://doi.org/10.3390/biomedicines10010083>
- Braga Emidio N., Hoffmann W., Brierley S.M., Muttenthaler M. (2019) Trefoil factor family: unresolved questions and clinical perspectives. *Trends Biochem. Sci.* **44**(5), 387–390. <https://doi.org/10.1016/j.tibs.2019.01.004>
- Kjellev S. (2009) The trefoil factor family — small peptides with multiple functionalities. *Cell. Mol. Life Sci.* **66**(8), 1350–1369. <https://doi.org/10.1007/S00018-008-8646-5>
- Madsen J., Nielsen O., Tornøe I., Thim L., Holmskov U. (2007) Tissue localization of human trefoil factors 1, 2, and 3. *J. Histochem. Cytochem.* **55**(5), 505–513. <https://doi.org/10.1369/JHC.6A7100.2007>
- Шестопапов А.В., Дворников А.С., Борисенко О.В., Тутельян А.В. (2019) Трефоилловые факторы – новые маркеры мукозального барьера желудочно-кишечного тракта. *Инфекция и иммунитет.* **9**(1), 39–46. doi: 10.15789/2220-7619-2019-1-39-46
- Kurt-Jones E.A., Cao L.C., Sandor F., Rogers A.B., Whary M.T., Nambiar P.R., Cerny A., Bowen G., Yan J., Takaishi S., Chi A.L., Reed G., Houghton J.M., Fox J.G., Wang T.C. (2007) Trefoil family factor 2 is expressed in murine gastric and immune cells and controls both gastrointestinal inflammation and systemic immune responses. *Infect. Immun.* **75**(1), 471–480. <https://doi.org/10.1128/IAI.02039-05>
- Iacobini C., Pugliese G., Blasetti Fantauzzi C., Federici M., Menini S. (2019) Metabolically healthy versus metabolically unhealthy obesity. *Metabolism.* **92**, 51–60. <https://doi.org/10.1016/j.metabol.2018.11.009>
- Шестопапов А.В., Колесникова И.М., Гапонов А.М., Григорьева Т.В., Хуснутдинова Д.Р., Камальдинова Д.Р., Волкова Н.И., Макаров В.В., Юдин С.М., Румянцев А.Г., Румянцев С.А. (2022) Влияние метаболического типа ожирения на микробиом крови. *Вопросы биологической, медицинской и фармацевтической химии.* **25**(2), 35–41.
- Колесникова И.М., Карбышев М.С., Гапонов А.М., Хуснутдинова Д.Р., Григорьева Т.В., Камальдинова Д.Р., Борисенко О.В., Макаров В.В., Юдин С.М., Румянцев С.А., Шестопапов А.В. (2023) Особенности таксономической принадлежности бактериальной ДНК крови у пациентов с различными метаболическими фенотипами ожирения. *Бюллетень сибирской медицины.* **22**(2), 61–67. <https://doi.org/10.20538/1682-0363-2023-2-61-67>
- Expert panel on detection evaluation and treatment of high blood cholesterol in adults. (2001) Executive summary of the third report of the National Cholesterol Education Program (NCEP) expert panel on detection, evaluation, and treatment of high blood cholesterol in adults (Adult Treatment Panel III). *JAMA.* **285**(19), 2486–2497. <https://doi.org/10.1001/jama.285.19.2486>
- Bolyen E., Rideout J.R., Dillon M.R., Bokulich N.A., Abnet C.C., Al-Ghalith G.A., Alexander H., Alm E.J.,

- Arumugam M., Asnicar F., Bai Y., Bisanz J.E., Bittinger K., Brejnrod A., Brislaw C.J., Brown C.T., Callahan B.J., Caraballo-Rodríguez A.M., Chase J., Cope E.K., Da Silva R., Diener C., Dorrestein P.C., Douglas G.M., Durall D.M., Duvallet C., Edwardson C.F., Ernst M., Estaki M., Fouquier J., Gauglitz J.M., Gibbons S.M., Gibson D.L., Gonzalez A., Gorlick K., Guo J., Hillmann B., Holmes S., Hlste H., Huttenhower C., Huttley G.A., Janssen S., Jarmusch A.K., Jiang L., Kaehler B.D., Kang K. Bin., Keefe C.R., Keim P., Kelley S.T., Knights D., Koester I., Kosciolk T., Kreps J., Langille M.G.I., Lee J., Ley R., Liu Y.X., Loftfield E., Lozupone C., Maher M., Marotz C., Martin B.D., McDonald D., McIver L.J., Melnik A.V., Metcalf J.L., Morgan S.C., Morton J.T., Naimy A.T., Navas-Molina J.A., Nothias L.F., Orchanian S.B., Pearson T., Peoples S.L., Petras D., Preuss M.L., Pruesse E., Rasmussen L.B., Rivers A., Robeson M.S., Rosenthal P., Segata N., Shaffer M., Shiffer A., Sinha R., Song S.J., Spear J.R., Swafford A.D., Thompson L.R., Torres P.J., Trinh P., Tripathi A., Turnbaugh P.J., Ul-Hasan S., van der Hoof J.J.J., Vargas F., Vázquez-Baeza Y., Vogtmann E., von Hippel M., Walters W., Wan Y., Wang M., Warren J., Weber K.C., Williamson C.H.D., Willis A.D., Xu Z.Z., Zaneveld J.R., Zhang Y., Zhu Q., Knight R., Caporaso J.G. (2019) Reproducible, interactive, scalable and extensible microbiome data science using QIIME2. *Nat. Biotechnol.* **37**(8), 852–857. <https://doi.org/10.1038/S41587-019-0209-9>
19. Quast C., Pruesse E., Yilmaz P., Gerken J., Schweer T., Yarza P., Peplies J., Glöckner F.O. (2013) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* **41**(Database issue), D590–D596. <https://doi.org/10.1093/nar/gks1219>
20. Ghanemi A., Yoshioka M., St-Amand J. (2021) Trefoil factor family member 2: from a high-fat-induced gene to a potential obesity therapy target. *Metabolites.* **11**(8), 536. <https://doi.org/10.3390/metabo11080536>
21. Ghanemi A., Yoshioka M., St-Amand J. (2021) Trefoil factor family member 2 expression as an indicator of the severity of the high-fat diet-induced obesity. *Genes (Basel).* **12**(10), 1505. <https://doi.org/10.3390/genes12101505>
22. Shestopalov A.V., Kolesnikova I.M., Savchuk D.V., Teplyakova E.D., Shin V.A., Grigoryeva T.V., Naboka Yu.L., Gaponov A.M., Roumiantsev S.A. (2023) Effect of the infant feeding type on gut microbiome taxonomy and levels of trefoil factors in children and adolescents. *J. Evol. Biochem. Physiol.* **59**, 877–890. <https://doi.org/10.1134/S0022093023030201>
23. Коваленко Т.В., Ларионова М.А. (2019) Трекинг ожирения в детском возрасте. *Педиатрия.* **98**, 128–135. <https://doi.org/10.24110/0031-403X-2019-98-4-128-135>
24. Wan Y., Yuan J., Li J., Li H., Yin K., Wang F., Li D. (2020) Overweight and underweight status are linked to specific gut microbiota and intestinal tricarboxylic acid cycle intermediates. *Clin. Nutr.* **39**(10), 3189–3198. <https://doi.org/10.1016/j.clnu.2020.02.014>
25. Tan J., McKenzie C., Potamitis M., Thorburn A.N., Mackay C.R., Macia L. (2014) The role of short-chain fatty acids in health and disease. *Adv. Immunol.* **121**, 91–119. <https://doi.org/10.1016/B978-0-12-800100-4.00003-9>
26. Brahe L.K., Astrup A., Larsen L.H. (2013) Is butyrate the link between diet, intestinal microbiota and obesity-related metabolic diseases? *Obes. Rev.* **14**(12), 950–959. <https://doi.org/10.1111/OBR.12068>
27. Amabebe E., Robert F.O., Agbalalah T., Orubu E.S.F. (2020) Microbial dysbiosis-induced obesity: role of gut microbiota in homeostasis of energy metabolism. *Br. J. Nutr.* **123**(10), 1127–1137. <https://doi.org/10.1017/S0007114520000380>
28. Yang J., Keshavarzian A., Rose D.J. (2013) Impact of dietary fiber fermentation from cereal grains on metabolite production by the fecal microbiota from normal weight and obese individuals. *J. Med. Food.* **16**(9), 862–867. <https://doi.org/10.1089/JMF.2012.0292>
29. Martínez-Cuesta M.C., del Campo R., Garriga-García M., Peláez C., Requena T. (2021) Taxonomic characterization and short-chain fatty acids production of the obese microbiota. *Front. Cell Infect. Microbiol.* **11**, 598093. <https://doi.org/10.3389/FCIMB.2021.598093>
30. Krolenko E.V., Kupriyanova O.V., Nigmatullina L.S., Grigoryeva T.V., Roumiantsev S.A., Shestopalov A.V. (2024) Changes of the concentration of short-chain fatty acids in the intestines of mice with different types of obesity. *Bull. Exp. Biol. Med.* **176**(3), 347–353. doi: 10.1007/s10517-024-06022-1
31. Thomas-Valdés S., Tostes M. das G.V., Anunciação P.C., da Silva B.P., Sant’Ana H.M.P. (2017) Association between vitamin deficiency and metabolic disorders related to obesity. *Crit. Rev. Food Sci. Nutr.* **57**(15), 3332–3343. <https://doi.org/10.1080/10408398.2015.1117413>
32. Walther B., Philip Karl J., Booth S.L., Boyaval P. (2013) Menaquinones, bacteria, and the food supply: the relevance of dairy and fermented food products to vitamin K requirements. *Adv. Nutr.* **4**(4), 463–473. <https://doi.org/10.3945/AN.113.003855>
33. Aussel L., Pierrel F., Loiseau L., Lombard M., Fontecave M., Barras F. (2014) Biosynthesis and physiology of coenzyme Q in bacteria. *Biochim. Biophys. Acta.* **1837**(7), 1004–1011. <https://doi.org/10.1016/j.bbabi.2014.01.015>
34. Nowrouzi B., Li R.A., Walls L.E., d’Espaux L., Malci K., Liang L., Jonguitud-Borrego N., Lerma-Escalera A.I., Morones-Ramirez J.R., Keasling J.D., Rios-Solis L. (2020) Enhanced production of taxadiene in *Saccharomyces cerevisiae*. *Microb. Cell Fact.* **19**(1), 200. <https://doi.org/10.1186/S12934-020-01458-2>
35. Goncheva M.I., Chin D., Heinrichs D.E. (2022) Nucleotide biosynthesis: the base of bacterial pathogenesis. *Trends Microbiol.* **30**(8), 793–804. <https://doi.org/10.1016/J.TIM.2021.12.007>
36. Ding T., Xu M., Li Y. (2022) An overlooked prebiotic: beneficial effect of dietary nucleotide supplementation on gut microbiota and metabolites in senescence-accelerated Mouse prone-8 mice. *Front. Nutr.* **9**, 820799. <https://doi.org/10.3389/fnut.2022.820799>
37. Гапонов А.М., Волкова Н.И., Ганенко Л.А., Набока Ю.Л., Маркелова М.И., Сиягина М.Н., Харченко А.М., Хуснутдинова Д.Р., Румянцев С.А., Тугельян А.В., Макаров В.В., Юдин С.М., Шестопалов А.В. (2021) Особенности микробиома толстой кишки у пациентов с ожирением при его различных фенотипах

- (оригинальная статья). *Журн. Микробиол. Эпидемиол. Иммунобиологии*. **98**(2), 144–155.  
doi: 10.36233/0372-9311-66
38. Kesh K., Mendez R., Mateo-Victoriano B., Garrido V.T., Durden B., Gupta V.K., Oliveras Reyes A., Merchant N., Datta J., Banerjee S., Banerjee S. (2022) Obesity enriches for tumor protective microbial metabolites and treatment refractory cells to confer therapy resistance in PDAC. *Gut Microbes*. **14**(1), 2096328.  
https://doi.org/10.1080/19490976.2022.2096328
  39. Wang X., Matuszek Z., Huang Y., Parisien M., Dai Q., Clark W., Schwartz M.H., Pan T. (2018) Queuosine modification protects cognate tRNAs against ribonuclease cleavage. *RNA*. **24**(10), 1305–1313.  
https://doi.org/10.1261/RNA.067033.118/-/DC1
  40. Tuorto F., Legrand C., Cirzi C., Federico G., Liebers R., Müller M., Ehrenhofer-Murray A.E., Dittmar G., Gröne H., Lyko F. (2018) Queuosine-modified tRNAs confer nutritional control of protein translation. *EMBO J.* **37**(18), e99777.  
https://doi.org/10.15252/EMBJ.201899777
  41. Nie X., Chen J., Ma X., Ni Y., Shen Y., Yu H., Panagiotou G., Bao Y. (2020) A metagenome-wide association study of gut microbiome and visceral fat accumulation. *Comput. Struct. Biotechnol. J.* **18**, 2596–2609.  
https://doi.org/10.1016/J.CSBJ.2020.09.026
  42. Gaca A.O., Kajfasz J.K., Miller J.H., Liu K., Wang J.D., Abranches J., Lemos J.A. (2013) Basal levels of (p)ppGpp in *Enterococcus faecalis*: the magic beyond the stringent response. *mBio*. **4**(5), e00646–13.  
https://doi.org/10.1128/MBIO.00646-13
  43. Siptroth J., Moskalenko O., Krumbiegel C., Ackermann J., Koch I., Pospisil H. (2023) Investigation of metabolic pathways from gut microbiome analyses regarding type 2 diabetes mellitus using artificial neural networks. *Discov. Artif. Intell.* **3**, 19.  
https://doi.org/10.1007/S44163-023-00064-6
  44. Tosar J.P., Cayota A. (2020) Extracellular tRNAs and tRNA-derived fragments. *RNA Biol.* **17**(8), 1149–1167.  
https://doi.org/10.1080/15476286.2020.1729584
  45. Gutiérrez-Repiso C., Molina-Vega M., Bernal-López M.R., Garrido-Sánchez L., García-Almeida J.M., Sajoux I., Moreno-Indias I., Tinahones F.J. (2021) Different weight loss intervention approaches reveal a lack of a common pattern of gut microbiota changes. *J. Pers. Med.* **11**(2), 109.  
https://doi.org/10.3390/JPM11020109
  46. Constante M., Fragoso G., Calvé A., Samba-Mondonga M., Santos M.M. (2017) Dietary heme induces gut dysbiosis, aggravates colitis, and potentiates the development of adenomas in mice. *Front. Microbiol.* **8**, 1809.  
https://doi.org/10.3389/FMICB.2017.01809
  47. Ijssennagger N., Belzer C., Hooiveld G.J., Dekker J., Van Mil S.W.C., Müller M., Kleerebezem M., Van Der Meer R., Klaenhammer T.R. (2015) Gut microbiota facilitates dietary heme-induced epithelial hyperproliferation by opening the mucus barrier in colon. *Proc. Natl. Acad. Sci. USA*. **112**(32), 10038–10043.  
https://doi.org/10.1073/PNAS.1507645112
  48. Fernández Á.F., Bárcena C., Martínez-García G.G., Tamargo-Gómez I., Suárez M.F., Pietrocola F., Castoldi F., Esteban L., Sierra-Filardi E., Boya P., López-Otín C., Kroemer G., Mariño G. (2017) Autophagy counteracts weight gain, lipotoxicity and pancreatic  $\beta$ -cell death upon hypercaloric pro-diabetic regimens. *Cell Death Dis.* **8**(8), e2970. https://doi.org/10.1038/CDDIS.2017.373
  49. Ramos-Molina B., Queipo-Ortuño M.I., Lambertos A., Tinahones F.J., Peñafiel R. (2019) Dietary and gut microbiota polyamines in obesity- and age-related diseases. *Front. Nutr.* **6**, 24.  
https://doi.org/10.3389/FNUT.2019.00024
  50. Liu R., Hong J., Xu X., Feng Q., Zhang D., Gu Y., Shi J., Zhao S., Liu W., Wang X., Xia H., Liu Z., Cui B., Liang P., Xi L., Jin J., Ying X., Wang X., Zhao X., Li W., Jia H., Lan Z., Li F., Wang R., Sun Y., Yang M., Shen Y., Jie Z., Li J., Chen X., Zhong H., Xie H., Zhang Y., Gu W., Deng X., Shen B., Xu X., Yang H., Xu G., Bi Y., Lai S., Wang J., Qi L., Madsen L., Wang J., Ning G., Kristiansen K., Wang W. (2017) Gut microbiome and serum metabolome alterations in obesity and after weight-loss intervention. *Nat. Med.* **23**(7), 859–868.  
https://doi.org/10.1038/NM.4358
  51. Шатова О.П., Гапонов А.М., Григорьева Т.В., Васильев И.Ю., Столетова Л.С., Макаров В.В., Юдин С.М., Румянцев С.А., Шестопалов А.В. (2023) Катаболиты триптофана и гены ферментов микробиома кишечника. *Вестник РГМУ*. **4**, 41–59.  
doi: 10.24075/vrgmu.2023.027
  52. Yu D., Yang Y., Long J., Xu W., Cai Q., Wu J., Cai H., Zheng W., Shu X.O. (2021) Long-term diet quality and gut microbiome functionality: a prospective, shotgun metagenomic study among urban Chinese adults. *Curr. Dev. Nutr.* **5**(4), nzab026.  
https://doi.org/10.1093/CDN/NZAB026
  53. Fabietti F., Delise M., Piccioli Bocca A. (2001) Investigation into the benzene and toluene content of soft drinks. *Food Control*. **12**(8), 505–509.  
https://doi.org/10.1016/S0956-7135(01)00041-X
  54. Srain B.M., Pantoja-Gutiérrez S. (2022) Microbial production of toluene in oxygen minimum zone waters in the Humboldt Current System off Chile. *Sci. Rep.* **12**(1), 10669.  
https://doi.org/10.1038/s41598-022-14103-2
  55. Synowiec A., Żyła K., Gniewosz M., Kieliszek M. (2021) An effect of positional isomerism of benzoic acid derivatives on antibacterial activity against *Escherichia coli*. *Open Life Sci.* **16**(1), 594–601.  
https://doi.org/10.1515/biol-2021-0060
  56. Javaheri-Ghezeldizaj F., Alizadeh A.M., Dehghan P., Ez-zati Nazhad Dolatabadi J. (2023) Pharmacokinetic and toxicological overview of propyl gallate food additive. *Food Chem.* **423**, 135219.  
https://doi.org/10.1016/J.FOODCHEM.2022.135219
  57. Guo J., Han X., Zhan J., You Y., Huang W. (2018) Vanillin alleviates high fat diet-induced obesity and improves the gut microbiota composition. *Front. Microbiol.* **9**, 2733.  
https://doi.org/10.3389/FMICB.2018.02733
  58. Шестопалов А.В., Колесникова И.М., Савчук Д.В., Гапонов А.М., Теплякова Е.Д., Григорьева Т.В., Васильев И.Ю., Румянцев А.Г., Борисенко О.В., Румянцев С.А. (2023) Влияние типа вскармливания на первом году жизни на метаболические профили микробного сообщества кишечника детей и подростков с ожирением и нормальной массой тела, проживающих в Ростовской области. *Педиатрия им. Г.Н. Сперанского*. **102**(5), 90–102.  
doi: 10.24110/0031-403X-2023-102-5-90-102
  59. Hersoug L.G., Møller P., Loft S. (2018) Role of microbiota-derived lipopolysaccharide in adipose tissue inflammation,

- adipocyte size and pyroptosis during obesity. *Nutr. Res. Rev.* **31**(2), 153–163.  
<https://doi.org/10.1017/S0954422417000269>
60. Bertani B., Ruiz N. (2018) Function and biogenesis of lipopolysaccharides. *EcoSal Plus.* **8**(1), 10.1128/ecosalplus.ESP-0001-2018.  
<https://doi.org/10.1128/ECOSALPLUS.ESP-0001-2018>
61. Pazos M., Peters K. (2019) Peptidoglycan. *Subcell. Biochem.* **92**, 127–168.  
[https://doi.org/10.1007/978-3-030-18768-2\\_5](https://doi.org/10.1007/978-3-030-18768-2_5)
62. Колесникова И.М., Гапонов А.М., Румянцев С.А., Ганенко Л.А., Волкова Н.И., Григорьева Т.В., Лайков А.В., Макаров В.В., Юдин С.М., Шестопалов А.В. (2022) Взаимосвязь содержания нейротрофинов и кишечного микробиома при различных метаболических типах ожирения. *Журнал эволюционной биохимии и физиологии.* **58**(4), 43–56.  
<https://doi.org/10.31857/S0044452922040076>

## Metabolic Profile of Gut Microbiota and Levels of Trefoil Factors in Adults with Different Metabolic Phenotypes of Obesity

I. M. Kolesnikova<sup>1, 2, \*</sup>, L. A. Ganenko<sup>3</sup>, I. Yu. Vasilyev<sup>4</sup>, T. V. Grigoryeva<sup>4</sup>, N. I. Volkova<sup>3</sup>, S. A. Roumiantsev<sup>1, 2, 5</sup>, A. V. Shestopalov<sup>1, 2, 5, 6</sup>

<sup>1</sup>*N.I. Pirogov Russian National Research Medical University, Moscow, 117997 Russia*

<sup>2</sup>*The National Medical Research Center for Endocrinology, Moscow, 117292 Russia*

<sup>3</sup>*Rostov State Medical University, Rostov-on-Don, 344002 Russia*

<sup>4</sup>*Kazan (Volga region) Federal University, Kazan, 420008 Russia*

<sup>5</sup>*Center for Molecular Health, Moscow, 117437 Russia*

<sup>6</sup>*Dmitry Rogachev National Medical Research Center of Pediatric Hematology, Oncology and Immunology, Moscow, 117997 Russia*

\*e-mail: [ir.max.kolesnikova@gmail.com](mailto:ir.max.kolesnikova@gmail.com)

Obesity is associated with changes in the gut microbiota, as well as increased permeability of the intestinal wall. In 130 non-obese volunteers, 57 patients with metabolically healthy obesity (MHO), and 76 patients with metabolically unhealthy obesity (MUHO), bacterial DNA was isolated from stool samples, and the 16S rRNA gene was sequenced. The metabolic profile of the microbiota predicted by PICRUSt2 (<https://huttenhower.sph.harvard.edu/picrust/>) was more altered in patients with MUHO than MHO. Obesity, especially MUHO, was accompanied by an increase in the ability of the gut microbiota to degrade energy substrates, produce energy through oxidative phosphorylation, synthesize water-soluble vitamins (B1, B6, B7), nucleotides, heme, aromatic amino acids, and protective structural components of cells. Such changes may be a consequence of the microbiota adaptation to the MUHO-specific conditions. Thus, a vicious circle is formed, when MUHO promotes the depletion of gut microbiome, and further degeneration of the latter contributes to the pathogenesis of metabolic disorders. The concentration of the trefoil factor family (TFF) in the serum of the participants was also determined. In MHO and MUHO patients, TFF2 and TFF3 levels were increased, but we did not find significant associations of these changes with the metabolic profile of the gut microbiota.

**Keywords:** metabolically healthy obesity, metabolically unhealthy obesity, gut microbiota, metabolic profile, TFF2, TFF3, PICRUSt2

УДК: 571.89:616.98:575(66.961.34)

## ВИРУСОПОДОБНЫЕ ЧАСТИЦЫ, НЕСУЩИЕ Env ВИЧ-1 С МОДУЛИРОВАННЫМ СОСТАВОМ ГЛИКАНОВ

© 2024 г. Г. А. Каевичер<sup>а</sup>, Е. И. Самохвалов<sup>а</sup>, Д. В. Щебляков<sup>а</sup>,  
А. Л. Гинцбург<sup>а, б</sup>, А. Н. Взоров<sup>а, с, \*</sup>

<sup>а</sup>Федеральный национальный исследовательский центр эпидемиологии и микробиологии им. Н.Ф. Гамалеи  
Министерства здравоохранения России, Москва, 123098 Россия

<sup>б</sup>Кафедра инфектологии и вирусологии Первого Московского государственного медицинского университета  
им. И.М. Сеченова Министерства здравоохранения России, Москва, 123098 Россия

<sup>с</sup>Биологический факультет Московского государственного университета  
им. М.В. Ломоносова, Москва, 119234 Россия

\*e-mail: anvzorov@mail.ru

Поступила в редакцию 24.11.2023 г.

После доработки 14.02.2024 г.

Принята к публикации 26.02.2024 г.

Ранее получены высокоиммуногенные вирусоподобные частицы (VLP), содержащие белки оболочки (Env) ВИЧ-1, которые обладали способностью преодолевать природную резистентность поверхностных белков ВИЧ-1, связанную с их низким содержанием и труднодоступностью консервативных эпитопов для вируснейтрализующих антител. Разработанная технология ранее применена нами для получения VLP, несущих модифицированные тримеры Env штамма ZM53(T/F) ВИЧ-1. Для продукции VLP использовали рекомбинантные вирусы осповакцины, экспрессирующие Env и Gag-Pol (каркасный белок) ВИЧ-1/SIV, а также ген *hr* вируса коровьей оспы – для преодоления ограничений в репликации вируса осповакцины – в клетках CHO. Экспрессируемый белок Env содержал трансмембранный перекрывающий домен (TMS) вируса опухоли молочной железы мышей (MMTV), увеличивающий включение тримеров Env в VLP, и цитоплазматический домен (CT) с GCN4-последовательностью, влияющей на конформацию поверхностной субъединицы (SU). Для изучения состава гликанового паттерна Env и его влияния на эффективность формирования VLP использовали мутантную клеточную линию CHO Lec1, в которой нет фермента N-ацетилглюкозаминилтрансферазы T1 (GlcNAc-T1), вовлеченного в процессинг высокоманнозных углеводных цепей. Этот прием позволил модулировать состав гликанов на поверхности VLP, оптимизировать условия их формирования и наметить подходы к преодолению ограниченной иммуногенности ВИЧ-1, связанной с экранированием гликанами эпитопов Env.

**Ключевые слова:** ВИЧ-1, Env, гликопротеин, гликаны, вирусоподобные частицы, рекомбинантный вирус осповакцины, клетки CHO

DOI: 10.31857/S0026898424040113, EDN: IMLJGJ

Мишенью для индукции нейтрализующих антител у оболочечных РНК-вирусов служит поверхностный трансмембранный белок, выполняющий ключевую роль в проникновении вируса в клетку. На вирусной оболочке эти белки используют ряд механизмов для ускользания от иммунного ответа: плотность их распределения на поверхности вириона низкая, конформация метастабильная и их покрывает слой гликанов. Оболочечный гликопротеин Env ВИЧ-1 относится к трансмембранным белкам I типа и для проникновения в клетку

использует механизм мембранного слияния, в процессе которого конформация переходит в стабильную [1, 2]. Ранее мы детально изучали биологические характеристики различных штаммов ВИЧ-1/SIV (simian immunodeficiency virus – вирус иммунодефицита обезьян) и обнаружили, что структура и длина цитоплазматического домена (CT) трансмембранной субъединицы gp41 могут влиять на биологические свойства вируса (стабильность тримера, активность формирования синцитий и инфекционность) [3–7]. Более того, стабилизированные тримеры Env ВИЧ-1 вызывали образование антител с широкой нейтрализующей активностью и высокой авидностью, в связи с чем их рассматривают как потенциальные иммуногены для вакцин с повышенной эффективностью [3]. Иммуногены,

Сокращения: ВКО – вирус коровьей оспы; MMTV (mouse mammary tumor virus) – вирус опухоли молочной железы мышей; VLP (virus-like particles) – вирусоподобные частицы; rVV (recombinant vaccinia virus) – рекомбинантный вирус осповакцины.

которые имитируют важные процессы при естественной инфекции и инициируют соответствующие врожденные иммунные реакции, имеют преимущество перед рекомбинантными белками (растворимые белки Env) благодаря сохранности природной структуры.

#### *ВИЧ-1-нейтрализующие антитела и гликозилирование*

Нейтрализующие анти-ВИЧ-1-антитела широкого спектра действия (broadly neutralizing antibodies, bNAbs) находят у пациентов с хронической инфекцией ВИЧ-1. Консервативные эпитопы, которые они распознают, идентифицированы в определенных областях белка Env, включая рецептросвязывающий сайт и проксимальную внешнюю область gp41. При экспериментальных инфекциях приматы тоже индуцируют bNAbs. До сих пор вакцинация против ВИЧ-1 вызывала лишь низкие уровни нейтрализующих антител [8]. Одно из потенциальных препятствий к выработке таких антител – гликозилирование поверхностного белка ВИЧ-1. Трансмембранный белок Env gp120-gp41 нативного вириона ВИЧ-1 имеет тримерную структуру. Белок Env считается одним из наиболее высокогликозилированных поверхностных вирусных гликопротеинов. В ряде работ показано, что гликозилирование Env приводит к экранированию антигенных сайтов на тримере [9–13]. Позднее обнаружили, что сайты, распознаваемые нейтрализующими антителами, могут включать гликановые компоненты с высокоманнозными цепями, которые ковалентно связаны с остатком N332 Env [14]. Впоследствии показали, что bNAbs распознают N-связанные гликаны с высокоманнозными цепями и в отсутствие сайта гликозилирования при N332 [15]. Такие антитела могут иметь высокий терапевтический потенциал. Однако, в какой степени такие антитела могут быть вызваны вакцинами, предстоит продемонстрировать. Имеются доказательства, что гликозилирование белка Env изменяется при хронической ВИЧ-инфекции [16]. Это обстоятельство усложняет разработку вакцин, индуцирующих антитела, нацеленные на гликозилированные эпитопы.

#### *Влияние модификаций гликанов на иммунный ответ*

Влияние изменений гликанового слоя на инфекционность вируса (взаимодействие с клеточными рецепторами), а также на взаимодействие с антителами определяли для модифицированных по гликанам белков Env ВИЧ-1. В этих экспериментах использовали различные ингибиторы гликозилирования и клеточные линии, дефицитные по способности продуцировать комплексные углеводные цепи [17]. Модификация гликанов – один из способов преодоления низкой иммуногенности

экранированных ими эпитопов, причем с сохранением консервативных гликановых структур, входящих в состав эпитопа, распознаваемого bNAbs [18]. Тем не менее до сих пор нет четкого понимания, как те или иные модификации гликанов отражаются на иммуногенности.

Вирусоподобные частицы (virus-like particles, VLP) обладают преимуществами в качестве вакцинных антигенов. Преимущества заключаются в отсутствии векторных антигенов, что позволяет проводить буст-иммунизацию, а также в повышенной иммуногенности антигенов, представленных в виде частиц, по сравнению с их растворимыми формами [19]. Размер частиц способствует эффективному поглощению антигенпрезентирующими клетками и последующей перекрестной презентации эпитопов. На поверхности VLP антигены обеспечивают перекрестное связывание рецепторов В-клеток. VLP эффективно проникают в лимфатические узлы, что способствует усилению стимуляции В- и Т-клеток.

В представленной работе использованы высокоиммуногенные Env-VLP ВИЧ-1, а для их продукции – клеточные линии CHO, в которых пути гликозилирования белков генетически модифицированы. Полученные VLP с белками Env, содержащими модифицированный СТ-домен, открывают возможность для получения новой информации о составе гликанов на тримерах различной конформации и их возможном влиянии на иммуногенность.

## ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

**Конструирование рекомбинантных вирусов.** Для получения рекомбинантных вирусов гены ВИЧ-1 [3] в составе плазмидного вектора (pRB21+вставка) встраивали в геном мутантного вируса осповакцины (vaccinia virus) VRB12 посредством гомологичной рекомбинации в клетках CV-1 [20]. Получены рекомбинантные вирусы осповакцины (recombinant vaccinia virus, rVV), кодирующие поверхностные белки ВИЧ-1: (1) VVEnv-22hb (helical bundle –  $\alpha$ -спиральный узел), несущий SU (surface unit – поверхностная субъединица) штамма ZM53 и домен TMS (transmembrane spanning domain – трансмембранный перекрывающий домен) из 22 а.о. вируса опухоли молочной железы мышей (mouse mammary tumor virus, MMTV), к СТ-домену которого присоединена функциональная тримеризующая последовательность GCN4 дрожжевого фактора транскрипции; и (2) VVEnv-22, конструкция которого аналогична Env-22hb с GCN4-последовательностью, потерявшей функциональность после замены остатков изолейцина на аланин (табл. 1). Также использовали полученный ранее вирус VVGag-Pol, кодирующий Gag-Pol, необходимый для формирования VLP и менее

**Таблица 1.** Рекомбинантные вирусы осповакцины, использованные для экспрессии модифицированных белков Env ВИЧ-1

Вирус	Состав экспрессируемого белка Env или целевой ген	Источник
VVE <sub>env</sub> -22hb	SU-домен ZM53 ВИЧ-1	[3]
	TMS-домен MMTV	[3]
	СТ-домен MMTV	[3]
	GCN4 (дрожжевой фактор транскрипции)	[3]
VVE <sub>env</sub> -22	SU-домен ZM53 ВИЧ-1	[3]
	TMS-домен MMTV	[3]
	СТ-домен MMTV	[3]
	нефункциональный GCN4 (Ile→Ala)	[3]
VVE <sub>env</sub> IIIВ (положительный контроль)	Env IIIВ ВИЧ-1	R.W. Compans (Emory University, США)
VVSCII (отрицательный контроль)	нет генов ВИЧ-1	[4]
VVCP	ген <i>hr</i> ВКО, нет генов ВИЧ-1	B. Moss (NIH, США)

подверженный ингибированию коэкспрессией Env по сравнению с Gag [4]. Рекомбинантный вирус VVCP, несущий ген *hr* вируса коровьей оспы (ВКО; cowpox virus), необходимый для преодоления ограничений экспрессии описанных выше гVV в клетках CHO, любезно предоставлен Dr. B. Moss (National Institutes of Health, США).

**Клеточные линии и условия культивирования.** Для получения VLP использовали клетки Hep2 и две линии клеток CHO: родительскую Pro-5 (CHO) и мутантную Lec1, в которой отсутствует N-ацетилглюкозаминилтрансфераза-ТI (GlcNAc-TI). Клетки были любезно предоставлены Dr. R.W. Compans (Emory University, США). Клетки выращивали на среде DMEM или  $\alpha$ -MEM с добавлением 10% эмбриональной сыворотки крупного рогатого скота (FBS).

**Очистка клеточных мембран и VLP.** Клетки, зараженные гVV, снимали скрепером, осаждали на микроцентрифуге Microfuge E (“Beckman Coulter”, США) при 14 000 об/мин в течение 30 с. К осадку добавляли лизирующий буфер (25 мМ Трис-НСl, рН 7.5, 150 мМ NaCl, 0.5% Triton X-100, 0.1% RMSF), ресуспендировали на вортексе, инкубировали во льду 5 мин и центрифугировали на микроцентрифуге 5804R (“Eppendorf”, Германия) при 20 913 g и 4°C в течение 20 мин. Супернатант собирали и хранили при –80°C. Для приготовления VLP супернатант от зараженных гVV клеток собирали, осветляли центрифугированием при 3 500 об/мин на роторе А-4-44 (5804R), пропускали через 0.45 мкм PVDF-фильтр Millex-HV (“Merck Millipore Ltd.”, Ирландия). VLP осаждали на ультрацентрифуге Beckman Optima XPN-90 (“Beckman Coulter”) в роторе SW41 при 28 000 об/мин и 4°C в течение 1.5 ч. Осадок суспендировали в PBS,

содержащем 15% (w/v) сахарозы (“Диаэм”, Россия) и хранили при –80°C.

**Ферментативная обработка VLP.** Для сравнения паттернов N-гликозилирования VLP, полученных из клеток CHO или Lec1, проводили дегликозилирование N-эндогликозидазами PNGase F и Endo H. PNGase F удаляет с гликопротеинов все типы N-связанных олигосахаридов: комплексные, гибридные и высокоманнозные цепи, – тогда как Endo H расщепляет хитобиозное ядро высокоманнозных и гибридных N-связанных углеводных цепей, но не расщепляет комплексные цепи. В экспериментах использовали следующие ферменты: Endo H и PNGase F (“Promega Corporation”, США);  $\alpha$ -химотрипсин из поджелудочной железы крупного рогатого скота и трипсин из поджелудочной железы крупного рогатого скота (“AppliChem”, Германия); субтилизин А (“Sigma-Aldrich”, США), протеинкиназа К (“Millipore”, США). Концентрированные VLP инкубировали с ферментами при 37°C в течение 1 ч, а затем растворяли в 2× буфере Laemmli.

**Электрофорез Blue Native в полиакриламидном геле (BN-PAGE).** Для анализа Env в составе VLP при невозможности восстановления условий использовали модифицированный протокол BN-PAGE [21]. Для высвобождения Env VLP инкубировали с равным объемом буфера, содержащего 0.12% Triton X-100, 1 мМ EDTA, 1.5 М аминокaproновой кислоты и 1 мкл коктейля ингибиторов протеаз (#04693132001, “Sigma-Aldrich”) в течение 5 мин, после чего вносили равный объем буфера для образцов, содержащего 100 мМ морфолинпропансульфонатной кислоты (MOPS), 100 мМ Трис-НСl, рН 7.7, 40% глицерина и 0.1% Coomassie Brilliant Blue. Образцы загружали в гель Bis-Tris NuPAGE

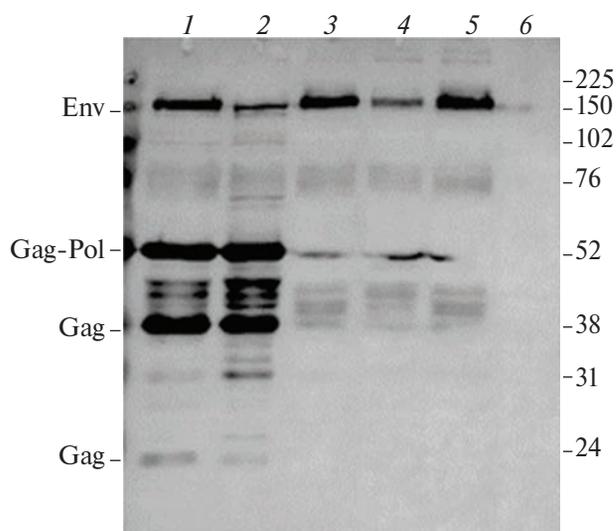
с концентрацией полиакриламида от 4 до 12% (“Invitrogen”, США). В качестве маркера использовали ферритин (“Amersham Biosciences”, США). Электрофорез проводили при 4°C в течение 3 ч при 100 В с использованием в качестве катодного буфера 50 мМ MOPS/50 мМ Трис, pH 7.7, 0.002% Coomassie Brilliant Blue, а в качестве анодного – тот же буфер без Coomassie. По окончании процесса белки с геля перенесли на PVDF-мембрану с использованием Trans-Blot® Turbo™ Transfer System (#1704150, “Bio-Rad Laboratories”, США).

**Иммуноблотинг.** При электрофорезе в 8, 10, 12, 4–15%-ном SDS-PAAG (“Bio-Rad Laboratories”) образцы денатурировали в 2× буфере Laemmli: 100 мМ, Трис-HCl (pH 6.8), 4% (w/v) SDS, 0.2% (w/v) Bromophenol blue, 20% (v/v) глицерин – без или с восстановителями (200 мМ дитиотреитол, 200 мМ β-меркаптоэтанол, 8 М мочевины), нагревали до 95°C или инкубировали при комнатной температуре в течение 5 мин. Перенос белков с геля на 0.45-мкм PVDF-мембрану (“GE Healthcare”, США) и на 0.45- или 0.22-мкм нитроцеллюлозную мембрану (“Bio-Rad Laboratories”) проводили в системе Trans-Blot® Turbo™ (“Bio-Rad Laboratories”). В качестве блокирующего буфера использовали 5%-ное сухое молоко (“PanReac AppliChem”, Германия) в PBS (“Эко-сервис”, Россия) с 0.1% Tween-20 (w/v) (“PanReac AppliChem”). Специфичные к ВИЧ-1 поликлональные антитела человека, полученные от инфицированных доноров (Россия и США), в разведении 1 : 1000 использовали в качестве первичных. Поликлональные козы антитела против IgG человека, конъюгированные с пероксидазой хрена (HRP) (“Sigma-Aldrich”), в разведении 1 : 80 000 использовали как вторичные. Все антитела разводили в вышеуказанном блокирующем буфере. В негативном контроле использовали незараженные клетки – для них не детектировали взаимодействия с анти-ВИЧ-1-антителами от инфицированных доноров. Люминесцентный сигнал от субстрата Clarity™ Western ECL (“Bio-Rad Laboratories”) детектировали на Amersham Imager 600 (“GE Healthcare”).

## РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ

### *Экспрессия модифицированных белков Env ВИЧ-1 рекомбинантными вирусами осповакцины в клетках Her2 и СНО*

В ходе исследования проанализирована способность гена *hr* ВКО преодолевать ограничение в репликации вируса осповакцины в СНО клетках (рис. 1, дорожки 2, 4, 6). Для сравнения использованы клетки Her2, которые перmissive для вируса осповакцины (рис. 1, дорожки 1, 3, 5). Клетки СНО заражали вирусами осповакцины, содержащими гены *env* и *gag-pol* ВИЧ-1, а также

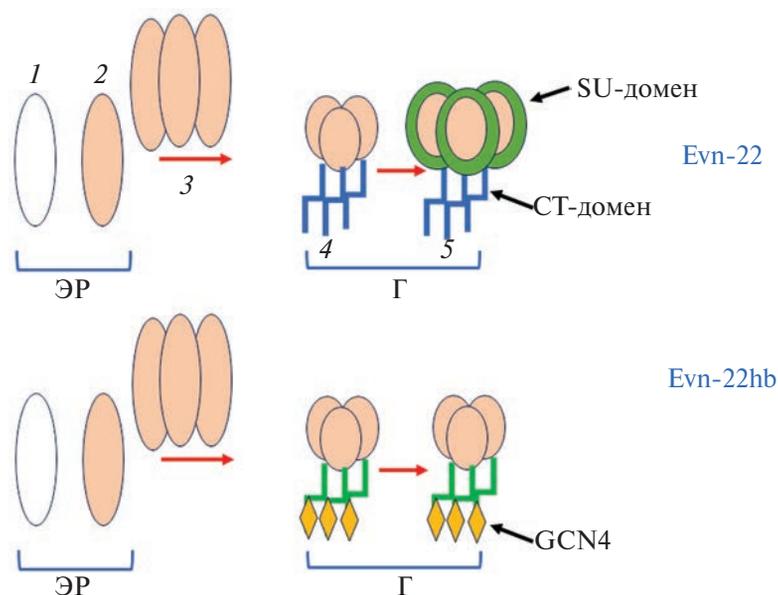


**Рис. 1.** Влияние экспрессии гена *hr* вируса коровьей оспы на синтез белков Env и Gag ВИЧ-1 в клетках СНО и Her2. Клетки Her2 (дорожки 1, 3, 5) или СНО (дорожки 2, 4, 6) коинфицированы VVEnvIIIВ (дорожки 1–6), VVGag-Pol (дорожки 1, 2), VVCP (дорожки 1–4) или VVSC11 (контрольный вирус без генов *hr* ВКО и ВИЧ-1) (дорожки 5, 6). Номера справа указывают на местоположение стандартов молекулярных масс белков (кДа).

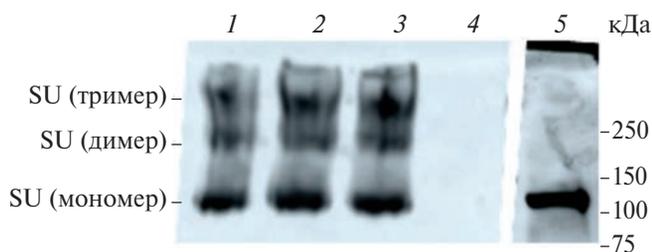
коинфицировали рекомбинантными вирусами, которые содержали интактный ген *hr* ВКО [22] (рис. 1, дорожки 1–4), или контрольным VVSC11 (рис. 1, дорожки 5, 6). Как видно из рис. 1, при коэкспрессии генов *env* ВИЧ-1 и *hr* ВКО в клетках СНО синтезировались белки Env и Gag (дорожки 2, 4), в то время как в отсутствие *hr* эти белки обнаружить не удалось (дорожка 6). При коинфицировании с VVGag-Pol в клетках Her2 и СНО наблюдались продукты нарезания белков Gag, что обусловлено присутствием в конструкции вирусной протеазы (дорожки 1, 2).

### *Формирование тримеров модифицированными белками Env ВИЧ-1*

Для изучения возможных изменений в составе гликанов белка Env ВИЧ-1 мы сравнили его стабилизированные и нестабилизированные конформации, для чего использовали перmissive клетки Her2. Трансляция белка Env сопровождается гликозилированием синтезированной полипептидной цепи и олигомеризацией в тримеры в эндоплазматическом ретикулуме клетки (рис. 2). Олигомеризация способствует перемещению тримеров в аппарат Гольджи, где происходит разрезание полипептидной цепи на SU и трансмембранную (transmembrane unit, TM) субъединицу и дальнейшее созревание углеводной цепи. Однако в составе тримера Env некоторые субстраты могут быть недоступны для ферментов [23].



**Рис. 2.** Трансляция белка Env ВИЧ-1 сопровождается гликозилированием синтезируемой полипептидной цепи и олигомеризацией в тримеры в эндоплазматическом ретикулуме (ЭР) клетки. Олигомеризация способствует перемещению тримеров в аппарат Гольджи (Г), где происходит нарезание полипептидной цепи на SU и трансмембранную субъединицу и дальнейшее созревание углеводной цепи. В составе тримера Env при наличии  $\alpha$ -спирали в СТ-домене некоторые субстраты могут быть недоступны для ферментов. (1) Env-предшественник; (2) олигоманнозные цепи (бежевый цвет), присоединенные по остатку аспарагина в сайте потенциального *N*-гликозилирования (Asn-X-Thr/Ser, где X  $\neq$  Pro) на мономере Env (gp160); (3) олигомеризация и перемещение тримеров Env; (4) протеолитическое расщепление gp160 клеточным фурином с образованием gp120 и gp41; (5) процессинг углеводных цепей и образование комплексных гликанов (зеленый цвет). Описание состава белков Env-22 и Env-22hb см. в табл. 1.



**Рис. 3.** Образование тримеров белком Env ВИЧ-1, экспрессируемым VVEnv-22. Клетки HeLa заражали вирусом с разной множественностью инфекции (MOI): 0.1 (дорожка 1), 0.2 (дорожка 2), 0.5 (дорожки 3, 5) БОЕ/клетка. Negативный контроль – неинфицированные клетки HeLa (дорожка 4). Через 48 ч клетки снимали и очищали фракцию плазматических мембран, содержащую белки ВИЧ-1. Образцы растворяли в буфере для нанесения образцов без восстановителей при комнатной температуре (дорожки 1–3) или с восстановителями (200 мМ дитиотреитол, 200 мМ  $\beta$ -меркаптоэтанол, 8 М мочевины) 5 мин при 96°C (дорожка 5). Белки разделяли в 4–15%-ном SDS-ПААГ и анализировали методом иммуноблоттинга.

Ранее установлено, что стабильность тримеров Env зависит от наличия  $\alpha$ -спирали в СТ-доме белка [3, 7, 24]. Учитывая этот факт, мы проанализировали способность рекомбинантного Env с лабильной конформацией формировать тримеры и процессинг его гликанов. С целью понять, как конформация белка Env, стабильная или более

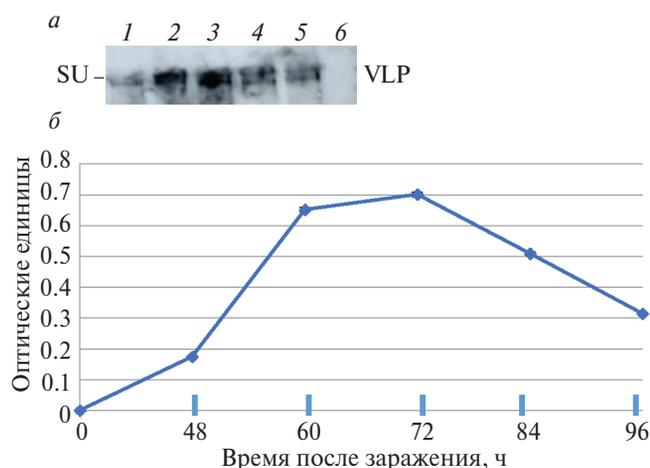
подвижная, влияет на формирование тримера и процессинг гликанов, мы использовали вариант VVEnv-22, который экспрессирует Env ВИЧ-1, содержащий TMS-домен MMTV и укороченный СТ-домен MMTV, несущий нефункциональную последовательность GCN4 (остатки Le заменены на Ala). В связи с тем что длина СТ-домена Env влияет на биологические свойства ВИЧ-1, нефункциональную последовательность GCN4 включили в состав СТ для соответствия его длины таковой в белке Env из VVEnv-22hb. Очищенную фракцию плазматических мембран клеток HeLa, зараженных VVEnv-22 в разных дозах (по числу бляшкообразующих единиц (БОЕ) на клетку), анализировали методом иммуноблоттинга (рис. 3, дорожки 1–3).

Как видно из результатов, приведенных на рис. 3, белок Env с укороченным СТ-доменом, без тримеризирующей последовательности GCN4, образовывал в клетках HeLa тримеры gp120SU, которые при одновременном воздействии высокой температуры и химических восстановителей распадались на мономеры (рис. 3, дорожка 5). Таким образом, продемонстрирована способность белка Env, не содержащего тримеризирующей последовательности в СТ-доме, формировать тримеры на поверхности мембраны клеток HeLa.

### Продукция VLP

При получении Env-VLP прежде всего определили оптимальные условия их формирования с максимальным включением модифицированных белков Env: во-первых, множественность заражения и, во-вторых, соотношение между тремя rVV. Ранее показано, что Env с транскрированным СТ более эффективно включается в VLP, чем полноразмерный белок, количество которого в частицах строго ограничено из-за длины СТ [3, 4, 25]. В связи с этим количество шипов в VLP, состоящих из укороченных бексов Env, можно увеличить за счет подбора соотношения вирусов VVCP, VVEnv-22 и VVGag-Pol. В ходе исследования кинетики продукции VLP в культуре клеток CHO, коинфицированной тремя rVV (VVCP + VVEnv-22 + VVGag-Pol), мы определили, что максимальный уровень VLP, несущих на поверхности Env ВИЧ-1, продуцируется через 72 ч после начала инфекции (рис. 4а, б), что коррелировало с данными, полученными ранее [4]. Выявлено, что модифицированный Env ВИЧ-1 включается в плазматические мембраны клеток в виде двух нековалентно связанных субъединиц: поверхностной и трансмембранной, — что также показано ранее для природных форм белка Env [23]. Мы проверили влияние изменений соотношений рекомбинантных вирусов на включаемость Env ВИЧ-1 в состав VLP.

Конфигирование тремя вирусами: VVCP + VVEnv-22 + VVGag-Pol — при одновременном возрастании значений MOI (от 0.1 до 0.5 БОЕ/клетка) для всех вирусов приводило к увеличению уровня Env-содержащих VLP. Однако при увеличении

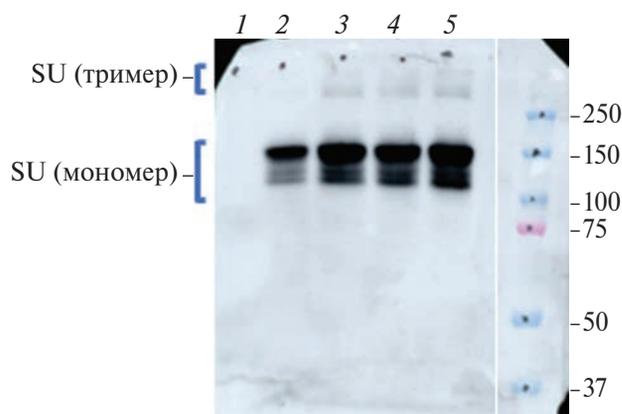


**Рис. 4.** Кинетика продукции VLP ВИЧ-1. а — Анализ продукции VLP методом иммуноблоттинга. Клетки CHO коинфицировали тремя вирусами: VVCP : VVEnv-22 : VVGag-Pol в соотношении 1 : 1 : 1 — с MOI 0.5 БОЕ/клетка; культуральную среду, содержащую VLP, собирали на 0 (дорожка б), 48 (1), 60 (2), 72 (3), 84 (4) и 96 (5) ч после начала заражения. б — Денситометрический анализ проведен с использованием программного обеспечения ImageJ. Данные представлены на основании двух независимых экспериментов.

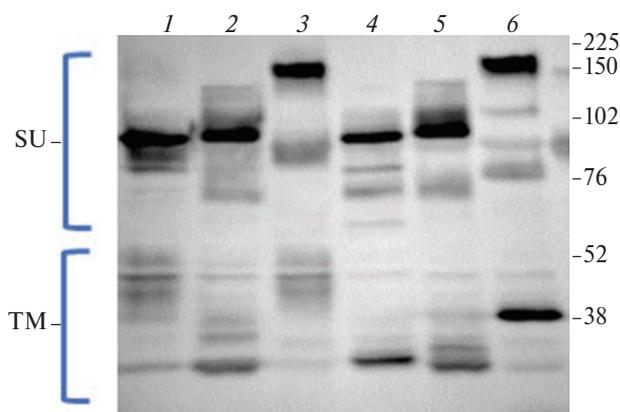
содержания VVCP или VVEnv-22 (при значениях MOI от 0.1 до 0.5 БОЕ/клетка) над количеством VVGag-Pol значительной разницы во включении Env в VLP не обнаружено, что, вероятно, связано с низкой множественностью заражения. При высоком значении MOI (от 5 до 25 БОЕ/клетка) включение тримеров Env в VLP увеличивалось, особенно в соотношении вирусов, где преобладала доля VVEnv-22 над VVGag-Pol (данные не приведены). По-видимому, это связано с тем, что при низких значениях MOI для трех вирусов (необходимых для продукции Env-VLP) одновременно заразить одну клетку сложнее. На основании этих результатов подобраны оптимальные условия для продукции в клетках CHO VLP с высоким уровнем включения тримеров Env.

### Паттерн гликозилирования модифицированных белков Env ВИЧ-1 в составе VLP

Мы изучали гликозилирование рекомбинантного белка Env ВИЧ-1 в клетках CHO. Для этого анализировали молекулярные массы гликановых форм белка Env ВИЧ-1, которые образуются в клетках CHO при заражении тремя вирусами: VVCP + VVEnv-22 + VVGag-Pol, — и сравнивали уровень их экспрессии (рис. 5). При увеличении содержания VVEnv-22 (рис. 5, дорожки 3–5) или VVCP (рис. 5, дорожки 3, 5) в смеси трех rVV для инокуляции в клетках CHO экспрессия Env повышалась, а уменьшение доли VVGag-Pol (VLP-формирующий вирус) не изменяло ее интенсивность (рис. 5, дорожка 5). В мембраны клеток включался Env ВИЧ-1 с разными молекулярными массами, преимущественно



**Рис. 5.** Иммуноблот-анализ молекулярной массы белков Env ВИЧ-1 на поверхности клеток CHO, коинфицированных смесью rVV. Содержание вирусов VVCP + VVEnv-22 + VVGag-Pol в инокуляте соответствовало 0.5 : 0.5 : 0.5 (дорожка 2), 1 : 1 : 1 (дорожка 3), 0.5 : 1.5 : 0.5 (дорожка 4), 1 : 1.5 : 0.5 (дорожка 5) БОЕ/клетка. Негативный контроль — клетки CHO, коинфицированные VVCP и VVGag-Pol в соотношении 1 : 1 БОЕ/клетка (дорожка 1). Белки разделяли в 8%-ном SDS-ПААГ и анализировали методом иммуноблоттинга, как описано в разделе “Экспериментальная часть”.



**Рис. 6.** Иммуноблоттинг белков Env ВИЧ-1, экспрессируемых клетками CHO (1–3) и CHO Lec1 (4–6), зараженных VVEnvШВ. Фракцию плазматических мембран обрабатывали Endo H (дорожки 1, 4), PNGase F (дорожки 2, 5) или не обрабатывали (дорожки 3, 6). Белки разделяли в 8%-ном SDS-ПААГ и анализировали методом иммуноблоттинга, как описано в разделе “Экспериментальная часть”.

соответствующими мономерной (120–150 кДа) и тримерной (360–450 кДа) форме SU-белка (рис. 5, дорожки 2–5). Машинерия гликозилирования в клетках CHO близка к таковой в клетках человека [26]. Вероятнее всего, диффузные полосы 120–150 и 360–450 кДа соответствуют различным стадиям процессинга углеводных цепей белка Env ВИЧ-1.

Также проведено сравнение белков Env, полученных в клетках CHO и в мутантных клетках CHO Lec1 (не образуют комплексных и гибридных N-гликозилированных углеводных цепей) до и после обработки гликозидазой (рис. 6).

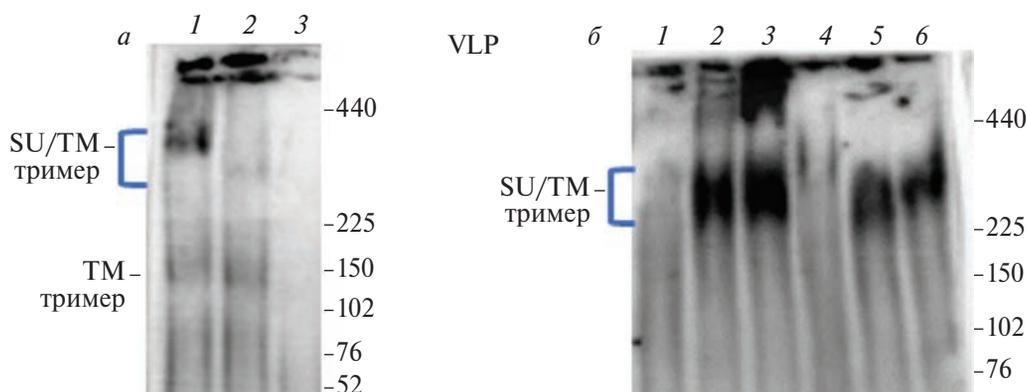
В не обработанной гликозидазами фракции плазматических мембран клеток CHO белки Env детектировали в виде диффузных полос в диапазоне от 41 до 160 кДа (рис. 6, дорожка 3), в то время

как в клетках Lec1 выявлены две более четкие полосы в районе 160 и 41 кДа (рис. 6, дорожка 6). Как видно на рис. 6, gp120 (SU) и gp41 (TM) высокогликозилированы в обоих клеточных линиях, преимущественно N-связанными олигосахаридами, чувствительными к обработке Endo H (3.2.1.96). Это высокоспецифичная эндогликозидаза, которая отщепляет от гликопротеинов Asn-связанные высокоманнозные углеводные цепи, но не комплексные – с высокой степенью процессинга. Производимые мономеры белков SU и TM интенсивно гликозилированы в обоих клеточных линиях, но в клетках Lec1 они более чувствительны к обработке Endo H, чем в клетках CHO дикого типа. Различная чувствительность клеток Lec1 и CHO к обработке Endo H связана с наличием комплексных и гибридных гликанов в CHO и доступностью для фермента высокоманнозных гликанов в Lec1.

Содержание модифицированных тримеров Env, которое оценивали по соотношению Gag/Env, находилось в диапазоне от 50 до 100 тримеров на VLP по сравнению с 10 тримерами на ВИЧ-вирионы первичных изолятов [3]. Более того, при изучении VLP с аутентичными белками Env представлены доказательства того, что частицы ВИЧ-1 могут содержать нефункциональные мономеры gp120/gp41 [21]. Для анализа полученных в клетках Lec1 VLP с белками Env22 или Env-22hb (SU/TM) в форме тримеров использовали BN-PAGE.

Чувствительность тримеров Env в составе VLP, полученных в клетках Lec1, к действию Endo H была избирательной (рис. 7а, дорожка 2; 7б, дорожки 2 и 5).

Как видно из рис. 7, тримеры Env-22 оказались не чувствительными к действию Endo H: молекулярные массы как не обработанных (рис. 7б, дорожка 3), так и обработанных (рис. 7б, дорожка 2) ферментом образцов не отличалась. Однако тримеры Env-22hb были чувствительны к действию Endo H: молекулярные массы необработанных (рис. 7б, дорожка 6) и



**Рис. 7.** BN-PAGE-анализ VLP, обработанных гликозидазами или протеазами. VLP, полученные в клетках Lec1, инфицированных с одинаковой множественностью VVGag-Pol и VVEnv-22 (а), с избытком VVEnv-22 или VVEnv-22hb по отношению к VVGag-Pol (б). VLP Env-22 (а, 1–3; б, 1–3); VLP Env-22hb (б, 4–6). Необработанные VLP (а, 1; б, 3 и 6), обработанные Endo H (а, 2; б, 2 и 5), смесью ферментов: химотрипсин, трипсин, субтилизин/протеиназа К (а, 3; б, 1 и 4).

обработанных (рис. 7б, дорожка 5) образцов отличались. На основании полученных результатов можно сделать вывод, что белок Env-22 содержит комплексные углеводные цепи, в то время как к Env-22hb прикреплены исключительно высокоманнозные гликаны.

## ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Ранее при разработке подходов к повышению иммуногенности белка Env установлено, что для этого необходимо выполнение, как минимум, двух условий: увеличения его содержания в несущей конструкции, то есть включаемости в VLP, и антигенной доступности, то есть экспонированности консервативных эпитопов, узнаваемых нейтрализующими антителами [3]. Замена TMS и СТ белка gp41 ВИЧ-1 на соответствующие последовательности из гликопротеина Env MMTV приводила к увеличению включения экспрессируемого модифицированного белка Env ВИЧ-1 в состав VLP. Скорее всего, это обусловлено тем, что плотность природного Env на вирионах MMTV гораздо выше, чем Env ВИЧ-1 на вирусной частице. Модификация СТ-домена включением в его состав тримеризирующей  $\alpha$ -спиральной последовательности GCN4 усиливала иммуногенность Env штамма ZM53 ВИЧ-1 – трансмиссивного штамма-основателя (transmitted founder; T/F) (этот вирус отличается от других ВИЧ-1 “компактностью” белка Env, что обусловлено более короткими петлями V1–V4 gp120 [27]). Эти модификации белка Env позволили достигнуть высоких уровней его включения в VLP [3].

Для продукции VLP с целью изучения состава гликанов поверхностных шипов ВИЧ-1 мы использовали клетки CHO. Этот выбор обусловлен тем, что эти клетки широко используют для сверхэкспрессии чужеродных белков и их суспензионные культуры легко масштабируемы, причем без сывороточной среды. Кроме того, клетки CHO относятся к очень немногим перевиваемым клеточным линиям, одобренным для экспрессии рекомбинантных белков в медицинских целях [22]. Мутантные варианты клеток CHO используют для изучения путей гликозилирования и идентификации генов белков, участвующих в этом процессе, для выяснения функциональной роли гликанов, а также в инженерии гликозилирования. Именно в клетках CHO, содержащих генетически измененные пути N-гликозилирования, удалось получить VLP, обогащенные Env. Это генетически сконструированная линия CHO Lec1, полученная из родительской CHO Pro-5 [28]. В ряде исследований показано, что значительное влияние на гликозилирование Env оказывает метаболическая активность клеточных линий, приводящая к различному содержанию N-гликанов в Env [17, 29, 30].

Мы обнаружили, что прикрепленные к белку Env без тримеризирующей последовательности высокоманнозные углеводные цепи подвергаются

процессингу с образованием комплексных цепей, в то время как Env с тримеризирующей последовательностью содержит исключительно высокоманнозные олигосахариды. По сравнению со стабилизированным нестабилизированный Env53-22 содержал гораздо более высокую долю комплексных гликанов. Различные факторы, такие как стабилизирующие мутации, скорость прохождения секреторного пути и другие, могут влиять на процессинг углеводных цепей гликопротеинов [31]. Как показано ранее, паттерн гликозилирования рекомбинантных Env – ключевой параметр при дизайне ВИЧ-вакцин [32]. Структура рекомбинантных растворимых тримеров Env SOSIP или других, среди которых есть клинические кандидаты, значительно отличается от аутентичных белков Env ВИЧ-1. У SOSIP нет СТ-домена. После транкирования СТ-домена на месте удаления остается открытая белковая поверхность, не защищенная гликанами. Эта белковая поверхность становится предпочтительной мишенью для антител, что изменяет антигенный паттерн и индуцирует нецелевой иммунный ответ или слабые штаммспецифические нейтрализующие антитела [32].

Иммуногенность полученных в представленной работе VLP, содержащих модифицированные белки Env с различной композицией гликанов, еще предстоит изучить. Однако давно показано, что участок, который модулирует взаимодействие поверхностной субъединицы с CD4 и корецептором, находится на петлях V1V2 вблизи  $\beta$ -листа gp120 [33, 34]. Более того, предполагается, что петля V4, расположенная на внешнем домене тримера gp120, при изменении ее размера влияет на упаковку и ориентацию гликанов вблизи CD4-связывающего сайта [27]. Таким образом, наличие компактной области V1–V4 в штамме ZM53 ВИЧ-1 имеет функциональное значение, по-видимому, благоприятствуя экспонированности CD4-связывающего домена и, как следствие, повышению иммуногенности [35–39]. Ранее продемонстрировано, что различия в аффинности к CD4 не ассоциированы с функциями Env (активностью слияния или инфекционностью вириона), в то время как первичная структура и длина домена СТ влияли на способность связываться с CD4. Кроме того, спиральная структура в домене СТ снижала способность связывания Env с CD4, но усиливала аффинность связывания с корецептором CCR5 [7, 23, 24, 40, 41].

Отличие VLP от растворимых белков состоит в том, что в составе частицы сохраняется природная структура тримера Env ВИЧ-1, которую можно изменять на другие конформации посредством модификации СТ-домена. Изменения конформации тримера Env сопровождаются изменениями в составе связанных с ним гликанов. Углеводные цепи на стабилизированных Env-тримерах ВИЧ-1 в высокоиммуногенных VLP [3] более однородны по составу и меньше по размеру, чем гликаны на тримерах Env с

подвижной конформацией. При дизайне VLP-иммуногена необходимо учитывать характеристики белка Env штаммов ВИЧ-1. Различия между белком Env штамма-основателя, имеющим “закрытую” конформацию и не вызывающим образования неспецифических антител, или Env штамма ВИЧ-1 при хронической инфекции — с “открытой” конформацией и высоким потенциалом изменений во внешнем и цитоплазматическом доменах — имеют значение при выборе стратегий, влияющих на демаскировку консервативных эпитопов. Надеемся, что с помощью сконструированных нами высокоиммуногенных VLP удастся расширить знания о влиянии гликанов на индукцию нейтрализующих антител, что позволит получить эффективные иммуногены ВИЧ-1.

Авторы благодарны проф. М.И. Букринскому за полезные дискуссии в ходе написания статьи.

Работа поддержана Государственным заданием Минобрнауки (Рег. № 122111700079-1).

Авторы заявляют об отсутствии конфликта интересов.

#### СПИСОК ЛИТЕРАТУРЫ

- Melikyan G.B., Markosyan R.M., Hemmati H., Delmedico M.K., Lambert D.M., Cohen F.S. (2000) Evidence that the transition of HIV-1 gp41 into a six-helix bundle, not the bundle configuration, induces membrane fusion. *J. Cell Biol.* **151**, 413–423.
- Rutten L., Lai Y.T., Blokland S., Truan D., Bisschop I.J.M., Strokappe N.M., Koornneef A., van Manen D., Chuang G.Y., Farney S.K., Schuitemaker H., Kwong P.D., Langedijk J.P.M. (2018) A universal approach to optimize the folding and stability of prefusion-closed HIV-1 envelope trimers. *Cell Rep.* **23**, 584–595.
- Vzorov A.N., Wang L., Wang B.Z., Compans R.W. (2016) Effects of modification of the HIV-1 Env cytoplasmic tail on immunogenicity of VLP vaccines. *Virology.* **489**, 141–150.
- Vzorov A.N., Compans R.W. (1996) Assembly and release of SIV env proteins with full-length or truncated cytoplasmic domains. *Virology.* **221**, 22–33.
- Vzorov A.N., Lea-Fox D., Compans R.W. (1999) Immunogenicity of full length and truncated SIV envelope proteins. *Viral Immunol.* **12**, 205–215.
- Vzorov A.N., Compans R.W. (2000) Effect of the cytoplasmic domain of the simian immunodeficiency virus envelope protein on incorporation of heterologous envelope proteins and sensitivity to neutralization. *J. Virol.* **74**, 8219–8225.
- Vzorov A.N., Compans R.W. (2011) Effects of stabilization of the gp41 cytoplasmic domain on fusion activity and infectivity of SIVmac239. *AIDS Res. Hum. Retroviruses.* **27**, 1213–1222.
- Haynes B.F., Wiehe K., Borrow P., Saunders K.O., Korber B., Wagh K., McMichael A.J., Kelsoe G., Hahn B.H., Alt F., Shaw G.M. (2023) Strategies for HIV-1 vaccines that induce broadly neutralizing antibodies. *Nat. Rev. Immunol.* **23**, 142–158.
- Back N.K., Smit L., De Jong J.J., Keulen W., Schutten M., Goudsmit J., Tersmette M. (1994) An N-glycan within the human immunodeficiency virus type 1 gp120 V3 loop affects virus neutralization. *Virology.* **199**, 431–438.
- Cole K.S., Steckbeck J.D., Rowles J.L., Desrosiers R.C., Montelaro R.C. (2004) Removal of N-linked glycosylation sites in the V1 region of simian immunodeficiency virus gp120 results in redirection of B-cell responses to V3. *J. Virology.* **78**, 1525–1539.
- Koch M., Pancera M., Kwong P.D., Kolchinsky P., Grundner C., Wang L., Hendrickson W.A., Sodroski J., Wyatt R. (2003) Structure-based, targeted deglycosylation of HIV-1 gp120 and effects on neutralization sensitivity and antibody recognition. *Virology.* **313**, 387–400.
- McCaffrey R.A., Saunders C., Hensel M., Stamatatos L. (2004) N-linked glycosylation of the V3 loop and the immunologically silent face of gp120 protects human immunodeficiency virus type 1 SF162 from neutralization by anti-gp120 and anti-gp41 antibodies. *J. Virology.* **78**, 3279–3295.
- Reitter J.N., Means R.E., Desrosiers R.C. (1998) A role for carbohydrates in immune evasion in AIDS. *Nat. Med.* **4**, 679–684.
- Julien J.P., Lee P.S., Wilson I.A. (2012) Structural insights into key sites of vulnerability on HIV-1 Env and influenza HA. *Immunol. Rev.* **250**, 180–198.
- Sok D., Doores K.J., Briney B., Le K.M., Saye-Francisco K.L., Ramos A., Kulp D.W., Julien J.P., Menis S., Wickramasinghe L., Seaman M.S., Schief W.R., Wilson I.A., Poirnard P., Burton D.R. (2014) Promiscuous glycan site recognition by antibodies to the high-mannose patch of gp120 broadens neutralization of HIV. *Sci. Transl. Med.* **6**, 236ra63.
- Lanteri M., Giordanengo V., Hiraoka N., Fuzibet J.G., Auberger P., Fukuda M., Baum L.G., Lefebvre J.C. (2003) Altered T cell surface glycosylation in HIV-1 infection results in increased susceptibility to galectin-1-induced cell death. *Glycobiology.* **13**, 909–918.
- Binley J.M., Ban Y.E., Crooks E.T., Eggink D., Osawa K., Schief W.R., Sanders R.W. (2010) Role of complex carbohydrates in human immunodeficiency virus type 1 infection and resistance to antibody neutralization. *J. Virol.* **84**, 5637–5655.
- Wagh K., Hahn B.H., Korber B. (2020) Hitting the sweet spot: exploiting HIV-1 glycan shield for induction of broadly neutralizing antibodies. *Curr. Opin. HIV AIDS.* **15**, 267–274.
- Lechner F., Jegerlehner A., Tissot A.C., Maurer P., Sebel P., Renner W.A., Jennings G.T., Bachmann M.F. (2002) Virus-like particles as a modular system for novel vaccines. *Intervirology.* **45**, 212–217.
- Blasco R., Moss B. (1995) Selection of recombinant vaccinia viruses on the basis of plaque formation. *Gene.* **158**, 157–162.
- Moore P.L., Crooks E.T., Porter L., Zhu P., Cayanan C.S., Grise H., Corcoran P., Zwick M.B., Franti M., Morris L., Roux K.H., Burton D.R., Binley J.M. (2006) Nature of nonfunctional envelope proteins on the surface of human immunodeficiency virus type 1. *J. Virol.* **80**, 2515–2528.
- Ramsey-Ewing A., Moss B. (1996) Recombinant protein synthesis in Chinese hamster ovary cells using a vaccinia virus/bacteriophage T7 hybrid expression system. *J. Biol. Chemistry.* **271**, 16962–16966.
- Checkley M.A., Lutge B.G., Freed E.O. (2011) HIV-1 envelope glycoprotein biosynthesis, trafficking, and incorporation. *J. Mol. Biol.* **410**, 582–608.
- Vzorov A.N., Yang C., Compans R.W. (2015) An amphipathic sequence in the cytoplasmic tail of HIV-1 Env alters cell tropism and modulates viral receptor specificity. *Acta Virol.* **59**, 209–220.
- Tedbury P.R., Novikova M., Ablan S.D., Freed E.O. (2016) Biochemical evidence of a role for matrix trimerization

- in HIV-1 envelope glycoprotein incorporation. *Proc. Natl. Acad. Sci. USA*. **113**(2), E182–E190. doi: 10.1073/pnas.1516618113
26. Nguyen N.T.B., Lin J., Tay S.J., Mariati, Yeo M., Nguyen-Khuong T., Yang Y. (2021) Multiplexed engineering glycosyltransferase genes in CHO cells via targeted integration for producing antibodies with diverse complex-type N-glycans. *Sci. Rep.* **11**, 12969.
  27. Derdeyn C.A., Decker J.M., Bibollet-Ruche F., Mokili J.L., Muldoon M., Denham S.A., Heil M.L., Kasolo F., Musonda R., Hahn B.H., Shaw G.M., Korber B.T., Allen S., Hunter E. (2004) Envelope-constrained neutralization-sensitive HIV-1 after heterosexual transmission. *Science*. **303**, 2019–2022.
  28. Patnaik S.K., Stanley P. (2006) Lectin-resistant CHO glycosylation mutants. *Methods Enzymol.* **416**, 159–182.
  29. Zhu X., Borchers C., Bienstock R.J., Tomer K.B. (2000) Mass spectrometric characterization of the glycosylation pattern of HIV-gp120 expressed in CHO cells. *Biochemistry*. **39**, 11194–11204.
  30. Raska M., Takahashi K., Czernekova L., Zachova K., Hall S., Moldoveanu Z., Elliott M.C., Wilson L., Brown R., Jancova D., Barnes S., Vrbkova J., Tomana M., Smith P.D., Mestecky J., Renfrow M.B., Novak J. (2010) Glycosylation patterns of HIV-1 gp120 depend on the type of expressing cells and affect antibody recognition. *J. Biol. Chem.* **285**, 20860–20869.
  31. Srinivas R.V., Compans R.W. (1983) Glycosylation and intracellular transport of spleen focus-forming virus glycoproteins. *Virology*. **125**, 274–286.
  32. Cao L., Pauthner M., Andrabi R., Rantalainen K., Berndsen Z., Diedrich J.K., Menis S., Sok D., Bastidas R., Park S.R., Delahunty C.M., He L., Guenaga J., Wyatt R.T., Schief W.R., Ward A.B., Yates J.R. 3rd., Burton D.R., Paulson J.C. (2018) Differential processing of HIV envelope glycans on the virus and soluble recombinant trimer. *Nat. Commun.* **9**, 3693.
  33. Kwong P.D., Wyatt R., Robinson J., Sweet R.W., Sodroski J., Hendrickson W.A. (1998) Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature*. **393**, 648–659.
  34. Rizzuto C.D., Wyatt R., Hernández-Ramos N., Sun Y., Kwong P.D., Hendrickson W.A., Sodroski J. (1998) A conserved HIV gp120 glycoprotein structure involved in chemokine receptor binding. *Science*. **280**, 1949–1953.
  35. Kolchinsky P., Kiprilov E., Sodroski J. (2001) Increased neutralization sensitivity of CD4-independent human immunodeficiency virus variants. *J. Virol.* **75**, 2041–2050.
  36. Puffer B.A., Pöhlmann S., Edinger A.L., Carlin D., Sanchez M.D., Reitter J., Watry D.D., Fox H.S., Desrosiers R.C., Doms R.W. (2002) CD4 independence of simian immunodeficiency virus Envs is associated with macrophage tropism, neutralization sensitivity, and attenuated pathogenicity. *J. Virol.* **76**, 2595–2605.
  37. Edwards T.G., Hoffman T.L., Baribaud F., Wyss S., LaBranche C.C., Romano J., Adkinson J., Sharron M., Hoxie J.A., Doms R.W. (2001) Relationships between CD4 independence, neutralization sensitivity, and exposure of a CD4-induced epitope in a human immunodeficiency virus type 1 envelope protein. *J. Virol.* **75**, 5230–5239.
  38. Johnson W.E., Morgan J., Reitter J., Puffer B.A., Czajak S., Doms R.W., Desrosiers R.C. (2002) A replication-competent, neutralization-sensitive variant of simian immunodeficiency virus lacking 100 amino acids of envelope. *J. Virol.* **76**, 2075–2086.
  39. Center R.J., Earl P.L., Lebowitz J., Schuck P., Moss B. (2000) The human immunodeficiency virus type 1 gp120 V2 domain mediates gp41-independent intersubunit contacts. *J. Virol.* **74**, 4448–4455.
  40. Vzorov A.N., Compans R.W. (2016) Cytoplasmic domain effects on exposure of co-receptor-binding sites of HIV-1 Env. *Arch. Virol.* **161**, 3011–3018.
  41. Liao H.X., Tsao C.Y., Alam S.M., Muldoon M., Vandergrift N., Ma B.J., Lu X., Sutherland L.L., Scearce R.M., Bowman C., Parks R., Chen H., Blinn J.H., Lapedes A., Watson S., Xia S.M., Foulger A., Hahn B.H., Shaw G.M., Swanstrom R., Montefiori D.C., Gao F., Haynes B.F., Korber B. (2013) Antigenicity and immunogenicity of transmitted/founder, consensus, and chronic envelope glycoproteins of human immunodeficiency virus type 1. *J. Virol.* **87**, 4185–4201.

## VLPs Carrying HIV-1 Env with Modulated Glycan Composition

G. A. Kaevitser<sup>1</sup>, E. I. Samokhvalov<sup>1</sup>, D. V. Scheblyakov<sup>1</sup>, A. L. Gintsburg<sup>1,2</sup>, A. N. Vzorov<sup>1,3,\*</sup>

<sup>1</sup>Gamaleya Federal Research Center of Epidemiology and Microbiology, Moscow, 123098 Russia

<sup>2</sup>Sechenov First Moscow State Medical University, Ministry of Health of the Russian Federation, Moscow, 123098 Russia

<sup>3</sup>Department of Immunology, Faculty of Biology, Lomonosov Moscow State University, Moscow, 119234 Russia

\*e-mail: anvzorov@mail.ru

Previously obtained highly immunogenic Env-VLPs ensure overcoming the natural resistance of HIV-1 surface proteins associated with their low level of incorporation and inaccessibility of conserved epitopes to induce neutralizing antibodies. We also adopted this technology to modify Env trimers of ZM53(T/F) strain to produce Env-VLPs by recombinant vaccinia viruses (rVVs). These rVVs expressing Env, Gag-Pol (HIV-1/SIV), as well as the cowpox virus *hr* gene allowing to avoid the restriction of vaccinia virus replication in CHO cells were used for VLP production. The CHO Lec1 engineered cell line lacking GlcNAc-TI was used for generating VLPs with Env proteins containing a cytoplasmic domain (CT) affecting on surface subunit (SU) conformation. This has created the opportunity to modulate the glycan composition, and refine the conditions for their production, and optimize approaches to overcoming HIV-1 resistance associated with abundant glycosylation.

**Keywords:** HIV-1, Env, glycoprotein, glycans, virus-like particles, recombinant vaccinia virus, CHO cells

## СТРУКТУРНО-ФУНКЦИОНАЛЬНЫЙ АНАЛИЗ БИОПОЛИМЕРОВ И ИХ КОМПЛЕКСОВ

УДК 577.218, 577.32

### МОЛЕКУЛЯРНЫЕ БЛОКАТОРЫ ИОННЫХ КАНАЛОВ ВИРУСОВ ГРИППА А И SARS-CoV-2

© 2024 г. Ю. Н. Воробьев<sup>а</sup>, \*

<sup>а</sup>Институт химической биологии и фундаментальной медицины Сибирского отделения Российской академии наук,  
Новосибирск, 630090 Россия

\*e-mail: ynvorob@niboch.nsc.ru

Поступила в редакцию 14.11.2023 г.

После доработки 07.02.2024 г.

Принята к публикации 15.02.2024 г.

Предложены молекулы, блокирующие функциональный цикл вирусов гриппа А и SARS-CoV-2. Молекулы-блокаторы, эффективно связывающиеся в каналах М2 и Е вирусов гриппа А и SARS-CoV-2 и блокирующие диффузию ионов  $H^+/K^+$ , разрушают жизненный цикл вирусов. Предложено семейство положительно заряженных, +2 е.у., молекул-блокаторов диффузии ионов  $H^+/K^+$  через М2- и Е-каналы. Изучено сродство молекул-блокаторов, производных диазабициклооктана, к связыванию белков каналов М2 и Е. Проведено моделирование тепловой динамики нативных и мутантных структур каналов и связывание молекул-блокаторов методом исчерпывающего докинга. Расчеты энергии связывания показали, что белки М2- и Е-каналов имеют внутриканальные, блокирующие и внеканальные сайты связывания. Предложены молекулы-блокаторы с преимущественным сродством к блокирующим модам связывания. Рассмотрены наиболее вероятные мутации аминокислот белка М2 и Е каналов, проанализирована эффективность блокирования каналов и предложены оптимальные структуры молекул-блокаторов.

**Ключевые слова:** вирусы гриппа А и SARS-CoV-2, белки ионных каналов М2 и Е, молекулярная динамика, блокаторы ионных каналов М2 и Е, производные диазабициклооктана, связывание и блокирование каналов М2 и Е

DOI: 10.31857/S0026898424040125, EDN: IMBPHX

#### ВВЕДЕНИЕ

Последнее столетие (1918–2019 гг.) известно серьезными эпидемиями гриппа А (<https://www.cdc.gov/flu/pandemic-resources/1918-commemoration/1918-pandemic-history.html>) (испанка: 1918–1922 гг.) и атипичной пневмонии (ковид-19: 2019–2023 гг.), вызванные вирусами H1N1 [1] и SARS-CoV-2 (<https://www.who.int/news-room/questions-and-answers/item/SARS-CoV-2-evolution>) соответственно. Поэтому разработка эффективных противовирусных препаратов остается задачей большой важности из-за способности этих вирусов быстро инфицировать большие популяции, преодолевая межвидовые барьеры путем быстрых мутаций. Известно значительное количество мутантных форм мембранного белка  $M^{pro}$ , формирующего поверхностные шипы вирусов гриппа А H1N1 [1, 2], и белка  $M^{pro}$  вируса SARS-CoV-2 [3, 4]. Лекарственный препарат, эффективный в том числе против мутантных штаммов, должен воздействовать на функционально значимые, но структурно консервативные элементы вируса. Консервативным структурным функционально значимым элементом вирусов являются

ионные каналы – М2 вируса гриппа А и Е-канал вируса SARS-CoV-2. Известны атомные структуры высокого разрешения ионных каналов вирусов гриппа А, канал М2 (код PDB2KAD) [5, 6] и канал Е вирусов SARS-CoV-2 (PDB код 5X29), что дает возможность конструировать молекулы блокаторов этих каналов [7, 8]. Функциональная значимость каналов М2 и Е состоит в транспорте ионов  $H^+/K^+$  через липидную мембрану вирусов гриппа А и SARS-CoV-2. Диаметр М2-канала вируса гриппа А составляет 4–6 Å, тогда как диаметр Е-канала SARS-CoV-2 варьирует в пределах 4–9 Å.

Важность функциональности ионных каналов для вирулентности показана на нескольких вирусах. Например, активность протонного канала белка М2 вируса гриппа А имеет решающее значение для проникновения вируса через клеточную мембрану и инфицирования клетки [9]. Активность канала белка Е коронавируса SARS-CoV-2 необходима для инфицирования живой клетки [10]. А патогенность вируса инфекционного бронхита значительно снижается, когда трансмембранный сегмент белка Е заменяли гетерологичным доменом, у которого отсутствовала активность ионного

канала [10]. Важность E-каналов для вирулентности и способности вызывать заболевание показана на нескольких вирусах: после эндоцитоза E-канал облегчает перенос ионов  $H^+$  и  $K^+$  во внутренний объем вируса [11], позволяя вирусной РНК иницировать репликацию. Показано, что канальная активность белка E имеет решающее значение для инфекционности РНК коронавирусов [12]. Оказалось также, что вирусы SARS-CoV-1, у которых заблокирована активность белка E, были гораздо менее заразными [11, 12]. Белок E участвует в сборке, высвобождении и патогенезе вируса [13, 14]. В целом, каналы M2 и E это небольшие гидрофобные вирусные белки, которые встраиваются в клеточную мембрану, облегчая тем самым высвобождение вируса из инфицированных клеток. Показано, что активность E-каналов нескольких коронавирусов, включая SARS-CoV-1 [15–17], коронавирус MERS [18], коронавирус HCoV-225 человека, вирус гепатита мыши и вирус инфекционного бронхита [19], может быть заблокирована гексаметиленаминоридом [20, 21]. Активность M2- или E-каналов вирусов гриппа A и SARS-CoV-2 имеет решающее значение для процесса заражения живой клетки [22–24]. Диффузия протонов  $H^+$  через M2- и E-каналы может быть стерически заблокирована связыванием электронеutralной молекулы блокатора внутри ионного канала [25, 26]. Более эффективное блокирование может быть достигнуто путем связывания положительно заряженной молекулы, создающей как стерическое блокирование, так и дополнительный электростатический барьер для положительно заряженных ионов  $H^+$  и  $K^+$  [27–29].

В настоящей работе суммированы результаты наших работ [6, 27–29], в которых предложена серия молекул, блокирующих ионные каналы нативных и наивероятных мутантных белков M2 вируса гриппа A и белка E вирусов SARS-CoV-2. Показано, что молекулы, производные 1,4-диазабикло-[2.2.2]октана (DABCO), эффективно связываются внутри M2- и E-каналов вирусов гриппа A и SARS-CoV-2, блокируют проницаемость для протонов и ионов  $K^+$  как стерически, так и электростатически, нарушая транспорт ионов и жизнеспособность вирусов.

## ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

**Метод моделирования.** Атомную структуру и тепловую динамику белков M2 и E при физиологической температуре (308 K) моделировали, исходя из известных экспериментальных данных. Моделирование равновесной структуры нативных белков M2 и E, в также их наивероятных мутантных форм выполнено методом молекулярной динамики (МД).

Моделирование равновесных тепловых флуктуаций трехмерных структур выполнено программным комплексом BioPASED [30–32] с неявной

моделью среды, описываемой обобщенной корректированной моделью Борна GB-MSR6c [33, 34].

Энергия биополимера в растворе U:

$$U = E_{\text{bond}} + E_{\text{angel}} + E_{\text{tors}} + E^{14}_{\text{vdw}} + E_{\text{vdw}} + E_{\text{elect}} + E_{\text{Hbond}} + E_{\text{solv}},$$

где  $E_{\text{bond}}$ ,  $E_{\text{angel}}$  и  $E_{\text{tors}}$  описывают энергию деформации валентных связей;  $E^{14}_{\text{vdw}}$  и  $E_{\text{vdw}}$  описывают парные Ван-дер-Ваальсовы (ВДВ) взаимодействия атомов, разделенных тремя и более валентными связями; энергии электростатических взаимодействий ( $E_{\text{elec}}$ ) пар атомов, разделенных тремя и более валентными связями, отличные от нуля в радиусе 15 Å, рассчитаны методом обобщенного поля реакции; энергия водородных связей  $E_{\text{Hbond}}$  рассчитывается как  $X-H \cdots Y$  между атомом водорода H, присоединенным к электроотрицательному атому X-H, и электроотрицательным атомом-акцептором Y. Функциональная форма включает зависимость от угла между X-H и H-Y, уменьшаясь при искривлении водородной связи; энергию сольватации ( $E_{\text{solv}}$ ) вычисляют по модели гауссовских гидратных оболочек, вычислительно эффективной и разумно точной.

Структура M2-канала с молекулой амантадина, связанной в канале, получена методом ЯМР, код 2KAD, при  $T = 243\text{K}$ , pH 7.5. Молекулу амантадина удалили из ионного канала M2, затем структуру канала оптимизировали методом симуляции отжига, т.е. выполняли нескольких циклов медленного разогрева системы до 400K и охлаждения до 300K.

Динамическая структура M2-канала получена как серия структур вдоль равновесной МД траектории длительностью 25 нс с временным интервалом 100 пс. Эта серия структур позволяет моделировать динамическое связывание молекул блокаторов флуктуирующей структурой M2-канала. Отметим, что при комнатных температурах позиционные флуктуации атомов белка более 1Å, как это следует из серии структур, полученных методом ЯМР, код 2KAD. Равновесные МД траектории при постоянном нейтральном pH 6.5 моделировали с периодическим уточнением ионизационных состояний ионизируемых аминокислотных остатков молекулы белка M2, содержащей четыре остатка гистидина,  $pK_0$  которых близка к физиологическому интервалу величин pH организма человека.

Структура E-канала вируса SARS-CoV-2, PDB код 5X29, получена методом ЯМР при температуре 308K, pH 5.5 – 16 моделей, которые отражают тепловую динамику среднеквадратичным отклонением атомов в пределах 2.6 Å, т.е. умеренную тепловую динамику молекулы при физиологических температурах. Динамическая структура E-канала для моделирования взаимодействия с молекулами блокаторов получена как серия структур вдоль

равновесной траектории метода МД длительно-стью 100 нс.

Расчеты значений констант ионизации  $pK_a$  и степени ионизации в зависимости от pH растворителя выполнены методами FAMBE и FAMBE-pH [31, 32]. Расчет степени ионизации аминокислот белков в растворе основан на минимизации энергии молекулы белка в данной конформации, зависимой от степени ионизации, по ее величине. Достоверность метода расчета  $pK_a$  в пределах 1.0 ед. показана для 297 ионизируемых групп в 32 белках [34].

Расчеты энергии межатомных взаимодействий выполнены в силовом поле AMBER94 и AMBER-GAFF [35–37]. Выполнен слепой исчерпывающий докинг молекул, который состоит из нескольких этапов [38]: (1) предварительный поиск всех возможных сайтов связывания на поверхности белка, доступной для молекулы лиганда; (2) селекция этих сайтов по контактному рангу, числу атомов белка, находящихся в контакте с атомами лиганда; (3) для всех сайтов с большим контактным рангом выполняется глобальная оптимизация полной энергии связывания лиганда по положению, ориентации и конформации лиганда, локальной структуре белка. Глобальная оптимизация структуры белка со связанным лигандом достигается стартом из 72 ориентаций лиганда, равномерно распределенных в пространстве ориентаций, методом симуляции отжига, т.е. выполняется МД комплекса белок-лиганд с несколькими циклами медленного нагревания до 400–500К и охлаждения до целевой температуры.

Белки, образующие M2- и E-каналы, встроены в липидную мембрану клетки, моделированную эффективной средой. Экспериментальные данные показывают, что диэлектрическая проницаемость липидной мембраны (т.е. вне объема белка E или M2),  $D_{out}$  равна  $\sim 30$  [39]. При моделировании электростатических взаимодействий атомов структур каналов внутренний объем каналов M2 и E рассматривается как объем белка с диэлектрической проницаемостью  $D_{in}$ , равной 12 [40–43].

**Ингибиторы ионных каналов белков M2 и E.** Трехмерные структуры белков M2 (код PDB2KAD) и E-каналов (код PDB5X29) определены методом ЯМР и содержат все атомы водорода. Структура M2-канала образована четырьмя  $\alpha$ -спиралями – остатки 22–46 каждой из цепей A, B, C, D [5]. Наиболее известными блокаторами ионного канала M2 являются амантадин и ремантадин [44, 45]. Существуют блокаторы ионного канала с более сложными структурами, например, производные пикротоксина [46], производные адамантана и спиро-адамантиламин [47]. МД-моделирование показало, что профиль свободной энергии иона  $H^+$  при движении вдоль ионного канала M2 имеет энергетический барьер  $\sim 5$  ккал/моль [48]. Моделирование методом МД потенциальной энергии

связывания амантадина с неявной моделью растворителя и окружающей среды, т.е. объема клеточной мембраны, предсказывает величину энергии связывания  $\sim 4$  ккал/моль, что удовлетворительно соответствует экспериментальным данным [49].

Новый класс молекулярных ингибиторов M2- и E-каналов сконструирован на основе базовой молекулы диазабициклооктана (DABCO), рис. 1а, которая служит первичным соединением [24]. Предложенные модификации DABCO значительно увеличивают вероятность связывания внутри M2- и E-каналов вирусов гриппа А и SARS-CoV-2.

**Полициклические производные диазабициклооктана.** Полициклические производные молекулы DABCO, рис. 1а, состоят из большего числа атомов, чем исходная молекула, поэтому они могут иметь более сильные Ван-дер-Ваальсовы и электростатические взаимодействия с внутренней поверхностью каналов M2 и E. Рассмотрен набор молекул блокаторов, состоящих из большего числа атомов, чем в молекуле DABCO. Молекулы диазабициклооктанбензола (DABCOB), рис. 1б, и диазабициклооктаннафталина (DABCON), рис. 1г. Более объемные молекулы, диазабициклооктан-3-бензол (DABCO3B) и диазабициклооктан-3-толуол (DABCO3N), рис. 1в и рис. 1д, получены модификацией циклической структуры корневой молекулы DABCO в трех симметричных направлениях.

Следует отметить, что конструирование ингибитора M2- и E-каналов как жесткой полициклической структуры более предпочтительно, по сравнению с гибкой структурой, поскольку при связывании и дополнительном ограничении конформационной подвижности молекулы лекарственного препарата наблюдается меньшая потеря конформационной энтропии по сравнению с подвижностью свободной молекулы.

## РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

### *Связывание DABCO и его производных с нативным M2 каналом*

Ионный канал M2 образован сегментами аминокислотных остатков 24–46 каждой из четырех  $\alpha$ -спиралей (A, B, C, D) с центральным продольным каналом [46, 47]. Вероятность транспорта протона через канал M2 незначительно меняется в области значений pH 5.4–7.0 [48]. Структурные данные показывают, что аминокислотные остатки Val27, Ala30, Ser31 и Gly34, важные для транспорта протона и взаимодействия с молекулой амантадина, блокирующей канал M2, лежат на внешней стороне каждой из  $\alpha$ -спиралей (A, B, C, D) [50, 51]. Мутация этих остатков с большой вероятностью будет влиять на транспорт протона через канал M2.

Известны несколько атомных структур пустого канала M2 (PDB коды 2H95, 2L0J, 4QKL, 4QKM

и 2KAD), полученных методами ЯМР и рентгеноструктурного анализа. ЯМР-структура (PDB код 2KAD), полученная при 243K и pH 7.5 [52–54], включает все атомы водорода, т.е. дает полную информацию об ионизационном состоянии остатков и типе таутомера нейтральных остатков гистидина. Вход в ионный канал M2 образован четырьмя остатками серина, Ser31A, Ser31B, Ser31C и Ser31D, которые контролируют диффузию ионов  $H^+$  через канал. Четыре остатка гистидина – His37A, His37B, His37C и His37D – лежат в одной плоскости, перпендикулярной к оси ионных каналов M2, и находятся в форме нейтрального таутомера HE2 (или HE, как принято в структурном моделировании). Средняя потенциальная энергия белка M2 вдоль равновесной МД траектории для ионизированного и нейтральных состояний четырех остатков гистидина приведена в табл. 1–3. Видно, что различия средних потенциальных энергий  $\langle E_{pot} \rangle_{MD}$  белка M2 для ионизированного состояния  $HIS^+$  и нейтрального состояния HE составляет  $\sim 6$  ккал/моль. Средняя величина тепловых флуктуаций энергии таутомера HE по равновесной МД траектории длительностью 25 нс, равна  $\sim 12$  ккал/моль.

Заселенности заряженных и нейтральных состояний четырех остатков гистидина заметно флуктуируют вдоль МД траектории (рис. 2).

Однако средние значения долей заселенностей ионизированного и нейтральных форм остатков His37A, His37B, His37C и His37D равны в пределах 0.04–0.06 (табл. 3). Ионизированное состояние остатков гистидина,  $HIS^+$  и нейтральный таутомер HE при pH 6.5, реализуются с вероятностями 0.64 и 0.32 соответственно (табл. 3). Средняя потенциальная энергия  $\langle E_{pot} \rangle_{MD}$  белка M2 с нейтральным таутомером HE значительно больше (на  $\sim 50$  ккал/моль) энергии состояний остатков гистидина  $HIS^+$  и HE (табл. 3), а вероятность состояний HE мала  $\sim 0.02$ – $0.04$ . Средняя пространственная структура нейтрального состояния канала M2 незначительно отличается, в пределах амплитуды тепловых флуктуаций  $\sim 0.9$  Å, от экспериментальной структуры канала M2 (PDB2KAD). Можно отметить, что среднеквадратичное отклонение (RMSD) для положений атомов  $O\gamma$  остатков Ser31A, Ser31B, Ser31C, Ser31D для структуры канала M2 вдоль траектории МД не превышало 1.1 Å. Например, расстояния между атомами  $O\gamma$  остатков Ser31A/Ser31C и Ser31B/Ser31D

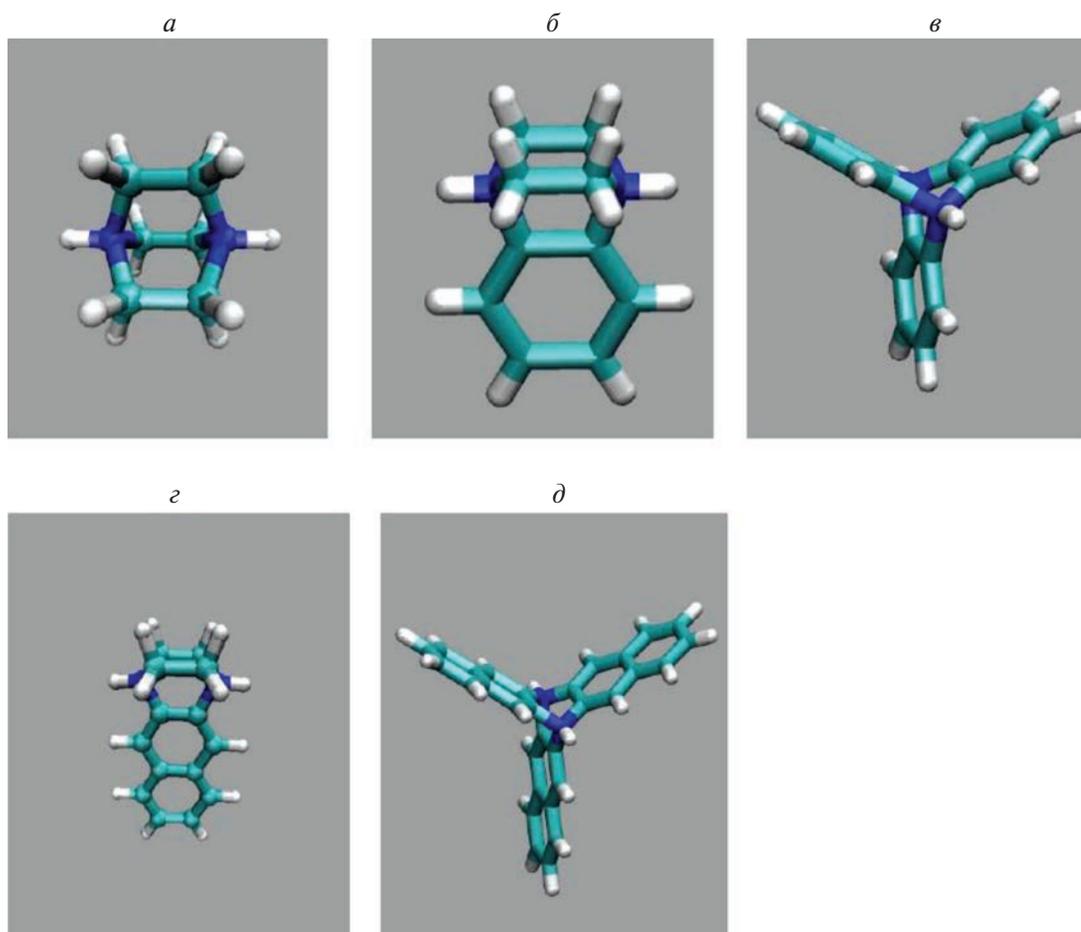


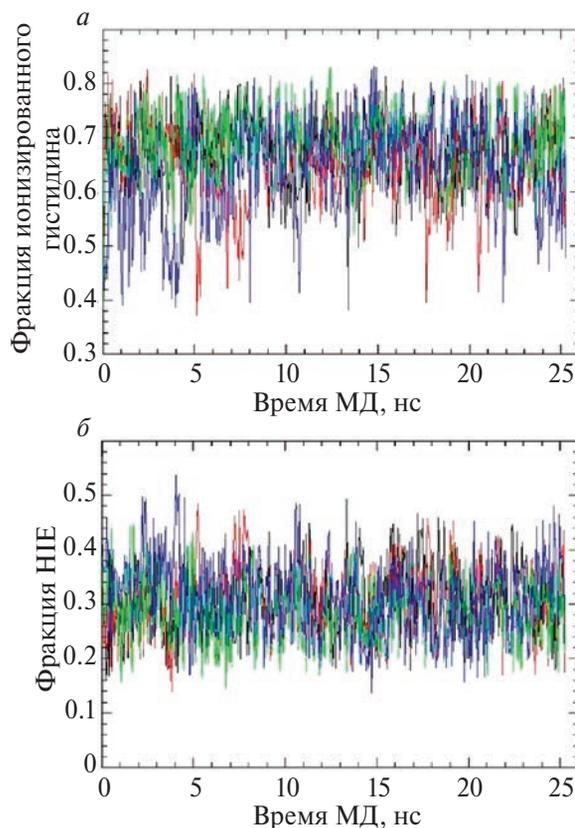
Рис. 1. Структура молекул блокаторов E- и M2- каналов. а – DABCO; б – DABCOB; в – DABCO3B; г – DABCON; д – DABCO3N.

**Таблица 1.** Средняя заселенность ионизационных состояний остатков гистидина (His34A, His34B, His34C, His34D) белка M2

pH	Ионизационное состояние		
	HIS <sup>+</sup>	HIE	HID
6.0	0.83 (0.03)	0.16 (0.02)	0.01 (0.01)
6.25	0.76 (0.03)	0.22 (0.02)	0.02 (0.01)
6.50	0.65 (0.04)	0.33 (0.03)	0.02 (0.01)
6.75	0.57 (0.04)	0.40 (0.03)	0.03 (0.01)
7.0	0.46 (0.04)	0.50 (0.030)	0.03 (0.01)
7.25	0.35 (0.03)	0.60 (0.02)	0.05 (0.02)
7.5	0.26 (0.03)	0.67 (0.02)	0.07 (0.02)

Примечание. Приведена средняя заселенность ионизационных состояний остатков гистидина; в скобках – величины среднеквадратичных (RMSD) флуктуаций.

колеблются в пределах 8.4–9.5 Å. Выход из канала M2 контролируется остатками гистидина His37A, His37B, His37C, His37D и характеризуется флуктуациями попарных расстояний между ближайшими атомами C<sub>ε1</sub> и N<sub>ε2</sub> в парах His37A–His37B, His37B–His37C, His37C–His37D, His37D–His37A, в диапазоне 4.5–5.4 Å. Средние значения pK<sub>a</sub> (табл. 3) остатков гистидина смещены на 0.3 единицы в сторону высоких значений pH, <pK<sub>a</sub>> ~6.8, (pK<sub>0</sub> = 6.5). Следовательно, остатки гистидина имеют незначительный отрицательный электростатический потенциал, который стабилизирует ионизированное состояние H<sup>+</sup> на ~8 ккал/моль (табл. 2). Значительные величины тепловых флуктуации энергий состояний HIS<sup>+</sup>

**Рис. 2.** Ионизированная HIS<sup>+</sup>-фракция остатков гистидина (а) и фракции нейтрального таутомера HIE (б). His37A (черный цвет), His37B (красный), His37C (синий), His37D (зеленый), белка M2 при pH 6.5, T=300K, вдоль равновесной МД траектории в 25 нс.**Таблица 2.** Энергии молекулы белка M2 при ионизированном и нейтральном состоянии остатков гистидина, T = 300K

Состояние ионизации	<E <sub>pot</sub> > <sub>MD</sub>	<E <sub>VDW</sub> > <sub>MD</sub>	<E <sub>COL</sub> > <sub>MD</sub>	<E <sub>Hbond</sub> > <sub>MD</sub>	<E <sub>Solv</sub> > <sub>MD</sub>
HIS <sup>+</sup>	-1344.5(15.2)	-423.7(13.6)	-495.1(6.6)	-331.2(7.0)	-98.3(7.1)
HIE	-1336.5(12.1)	-416.8(10.9)	-489.5(4.2)	-323.2(5.6)	-106.1(8.3)
HID	-1289.7(10.9)	-413.7(9.3)	-477.2(4.4)	-294.7(5.5)	-133.2(7.2)

E<sub>pot</sub> – Средняя потенциальная энергия вдоль равновесной МД-траектории, ккал/моль.

E<sub>VDW</sub> – Энергия Ван-дер-Ваальсовых взаимодействий.

E<sub>COL</sub> – Энергия электростатических взаимодействий.

E<sub>Hbond</sub> – Энергия водородных связей.

E<sub>Solv</sub> – Энергия взаимодействия с окружающей средой.

**Таблица 3.** Средние значения pK<sub>a</sub> и доли фракций ионизированных и нейтральных таутомеров остатков гистидина белка M2 вдоль равновесной МД траектории при pH 6.5, T = 300K

Остаток His	Ионизационное состояние			pK <sub>a</sub>
	HIS <sup>+</sup>	HID	HIE	
His37A	0.65(0.04) <sup>d</sup>	0.03(0.01) <sup>d</sup>	0.32(0.03)	6.80(0.14)
His37B	0.63(0.06)	0.04(0.02)	0.33(0.03)	6.80(0.09)
His37C	0.66(0.03)	0.05(0.01)	0.29(0.02)	6.82(0.16)
His37D	0.62(0.05)	0.05(0.03)	0.32(0.04)	6.78(0.18)
Доля	0.64	0.04	0.32	6.80

Примечание. Приведены средние значения; в скобках – величины RMSD флуктуаций.

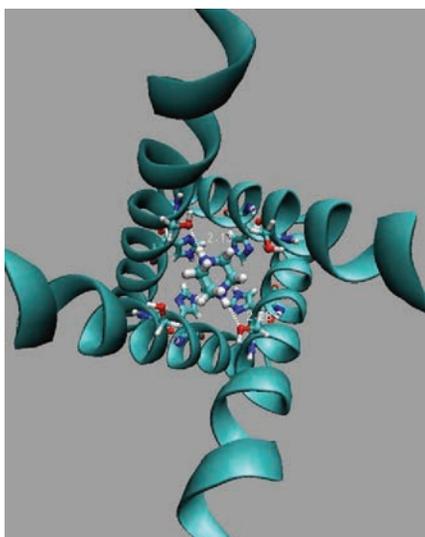
**Таблица 4.** Энергия оптимальных сайтов связывания производных молекулы-блокатора DABCO с нативным M2-каналом

Молекула	ePL	eVDW	eCoul	eHb	Y...HX	Расстояние, Å	eSolv	in(out)
DABCO <sup>a</sup>	-9.6 -9.4 -9.2	-8.2 -7.9 -11.0	4.7 4.3 1.2	-9.2 -9.1 -6.0	SER31A-O <sub>γ</sub> ...HN1	2.12	-8.1 -8.4 -4.6	<i>in</i> <i>in</i> <i>out</i>
					SER31C-O <sub>γ</sub> ...HN2	2.28		
					HIE37A-ND1...HN1	2.21		
					HIE37C-ND1...HN2	2.26		
					LEU36A-O...HN1	2.34		
					LEU46B-O...HN2	2.38		
DABCOB	-18.3 -18.2 -18.1	-14.1 -26.8 -16.3	1.2 3.7 2.5	-9.0 -7.6 -9.6	SER31A-O <sub>γ</sub> ...HN1	2.13	-7.9 0.3 -3.4	<i>in</i> <i>in</i> <i>out</i>
					SER31C-O <sub>γ</sub> ...HN2	2.29		
					HIE37A-ND1...HN1	2.28 2.34		
					HIE37C-ND1...HN2	2.19		
					LEU36A-O...HN1	2.17		
					LEU46B-OXT...HN2	2.17		
DABCON	-27.4 -21.3	-21.5 -19.7	-5.3 1.5	-9.8 -8.7	ASP44A-OD1...HN1	2.17	-5.1 -7.6	<i>out</i> <i>in</i>
					LYU46C-OXT...HN2	2.15		
					SER31A-O <sub>γ</sub> ...HN1	2.24 2.26		
					SER31C-O <sub>γ</sub> ...HN2	2.24 2.26		

Примечание. ePL — полная энергия связывания с белком M2, ккал/моль;  
eVDW — энергия Ван-дер-Ваальсовых взаимодействий;  
eCoul — энергия электростатических взаимодействий;  
eHb — энергия водородных связей;  
Y — акцептор водородной связи молекулы блокатора с атомами M2-канала;  
HX — донор водородной связи;  
eSolv — взаимодействие с растворителем;  
*in(out)* — позиция связывания лиганда внутри/вне M2-канала.

и HIE, ~15 и 12 ккал/моль (табл. 2), способствуют процессу переноса протона через “гистидиновый затвор” — квартет остатков His37A, His37B, His37C, His37D ионного канала M2. Параметры связывания молекул блокаторов, производных DABCOB, с ионным каналом M2, приведены в табл. 4.

Оптимальная мода связывания, блокирующая M2-канал, приведена на рис. 3. Эти данные

**Рис. 3.** Основной сайт связывания молекул DABCOB с нативным M2-каналом.

показывают, что оптимальная молекула-блокатор DABCOB имеет высокую энергию связывания внутри M2-канала с образованием двух сильных водородных связей (рис. 3) остатками SER31A(C) либо HIE37A(D). Связывание вне канала M2 может происходить с вероятностью ~ 0.32.

Оценочная вероятность мутации остатков SER31A(B, C, D) канала M2 в отличный от SER остаток, влияющий на свойства канала, достаточно высока ~ 0.14 [57]. Мутации остатков SER31 в четырех альфа-цепях белка M2, образующих внутреннюю поверхность канала M2, в остатки THR, ASN, ALA, GLU, CYS, случаются с вероятностями 38, 34, 28, 17 и  $16 \times 10^{-4}$  соответственно [55]. Расчеты энергий связывания молекул-блокаторов с мутантными M2-каналами приведены в табл. 5. Видно, что оптимальным сайтом связывания остается блокирующая позиция внутри канала M2, аналогичная позиции в нативном M2-канале. Энергии связывания незначительно (~ 1.5–2.0 ккал/моль) выше, чем в нативном M2- канале (табл. 5).

#### Связывание DABCO и его производных с E-каналом

Подготовка набора оптимизированных низкоэнергетических белковых структур E-канала к моделированию связывания молекул блокаторов выполнена методом, описанным, в разделе *Метод моделирования*.

**Таблица 5.** Энергии связывания оптимального блокатора DABCOB с нативными и мутантными структурами M2-канала

Мутант Ser31	ePL	eVDW	eCoul	eHb	Y...HX	Расстояние, Å	eSolv	in/out	вр
Mut1/Thr31	-19.8 -17.9	-29.5 -7.6	3.4 -7.4	-7.4 -4.8	HIE37A-ND1...HN1	1.92	1.5 -9.4	in out	38
					HIE37C-ND1...HN2	2.03			
					ASP44A-OD1...HN1	2.02			
					ASP44C-OD1...HN2	1.91			
Mut2/Asn31	-19.6 -17.1	-28.1 -12.6	3.4 -3.9	-6.8 -5.0	HIE37A-ND1...HN1	1.99	0.7 -7.5	in out	34
					HIE37C-ND1...HN2	2.01			
					LEU46A-OXT...HN1	2.12			
					LEU46C-OXT...HN2	2.16			
Mut3/Ala31	-19.8 -16.9	-28.5 -11.9	4.1 -5.6	v7.1 -5.0	HIE37A-ND1...HN1	1.97	-0.1 -6.4	in out	28
					HIE37C-ND1...HN2	2.01			
					OXT-LEU46A...HN1	2.17			
					OXT-LEU46C...HN2	2.21			
Mut4/Gly31	-18.4 -14.4	-28.8 -13.9	4.5 -1.1	-7.2 5.0	HIE37A-ND1...HN1	1.95	1.0 -6.8	in out	16
					HIE37C-ND1...HN2	2.01			
					ASP44A-OD1...HN1	2.02			
					ASP44A-OD1...HN2	1.94			
Mut5/Cys31	-20.0 -8.9	-27.2 -21.4	2.9 6.7	-7.9 -4.7	HIE37A-ND1...HN1	1.94	-0.1 -2.3	in in	11
					HIE37C-ND1...HN2	2.03			
					CYS31A-O...HN1	2.12			
					CYS31C-O...HN2	2.08			

Примечание. ePL – полная энергия связывания лиганда с белком M2, ккал/моль;

eVDW – энергия Ван-дер-Ваальсовых взаимодействий;

eCoul – энергия электростатических взаимодействий;

eHb – энергия водородных связей;

Y – акцептор водородной связи молекулы-блокатора с атомами M2-канала;

HX – донор водородной связи;

eSolv – взаимодействие с растворителем;

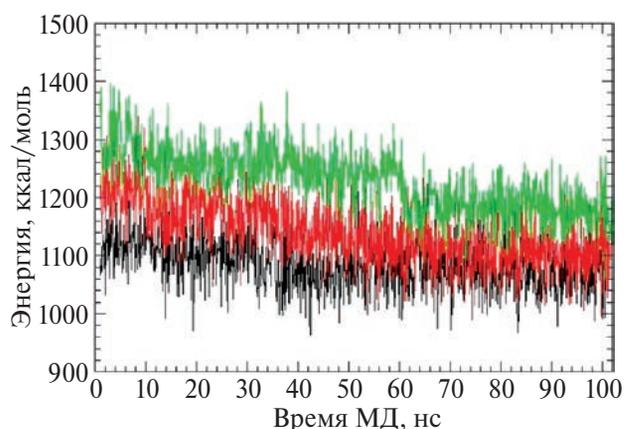
in/out – позиция связывания лиганда внутри/вне M2-канала;

вр – вероятность мутации  $\times 10^4$ .

Рассмотрены нативная и две мутантные структуры белка E, нативная, *mut0*, содержащая GLU8A(B, C, D, E) и ASN15A(B, C, D, E), *mut1* с остатками GLU8A(B, C, D, E), ASP15A(B, C, D, E) и *mut2* с остатками GLN8A(B, C, D, E) и ASN15A(B, C, D, E). Мутантные структуры позволили выяснить влияние кислотных, отрицательно заряженных остатков на структуру входа в E-канал. Равновесные МД траектории при комнатной температуре, 308K, длительностью 100 нс для нативной и мутантных форм E-канала показаны на рис. 4. В табл. 6 приведены значения средних энергий и амплитуды флуктуаций потенциальной энергии структур. Средние амплитуды флуктуаций находятся в пределах 53.2 (*mut0*), 50.1 (*mut1*), 44.3 (*mut2*), ккал/моль для нативной и двух мутантных структур. Видно, что нативная структура, *mut0*, с GLU8A(B, C, D, E) и ASN15A(B, C, D, E) имеет минимальную потенциальную энергию вдоль равновесной МД траектории по сравнению с мутантными *mut1* с GLU8A(B, C, D, E), ASP15A(B, C, D, E) и *mut2* с остатками GLN8A(B, C, D, E) и ASN15A(B, C, D, E). Средние значения энергии

с мутантных структур *mut1* и *mut2* больше на 30.1 и 94.4 ккал/моль соответственно.

Флуктуирующие потенциальные энергии нативной структуры *mut0* и мутантных структур *mut1*



**Рис. 4.** Флуктуации потенциальной энергии нативной и мутантных структур E-канала при  $T=308K$ , pH 6.5; черный – нативная структура, красный – *mut1*, зеленый – *mut2*.

**Таблица 6.** Средние энергии нативной и мутантных структур E-белка при  $T=308\text{K}$ 

Мутантный белок	eP	eVDW	eCoul	eHb	eDef	delta	eRMSD
Нативный/mut0	1075.5	-1078.4	-1441.5	-1073.5	5194.1	0.0	53.2
mut1	1106.2	-1082.1	-1429.2	-1042.8	5170.5	30.7	52.1
mut2	1171.1	-1074.1	-1391.6	-1066.7	5205.6	94.4	49.3

Примечание. eP— средняя энергия нативного E-белка и его мутантных форм, ккал/моль;  
 eVDW— энергия Ван-дер-Ваальсовых взаимодействий;  
 eCoul— энергия электростатических взаимодействий;  
 eHb— энергия водородных связей;  
 eDef— энергия тепловых деформаций структуры;  
 delta— энергия относительно нативной структуры;  
 eRMSD— энергия тепловых флуктуаций.

**Таблица 7.** Сайты связывания производных DABCO с нативным, *mut0*, E-белком вируса SARS-CoV-2 [Glu8<sup>-</sup>Asn15(A, B, C, D, E)]

Молекула-блокатор	ePL	eVDW	eCoul	eHb	Y...HX	Расстояние, Å	eSolv	<i>in(out)</i>
DABCO <sup>a</sup>	-24.3 -23.8 -23.5	-8.1 -10.2 -11.8	-2.9 -0.3 4.9	-9.9 -9.8 -9.7	GLU8A-OE2...HN1	1.91	-4.4 -3.5 -5.2	<i>in</i> <i>out</i> <i>in</i>
					GLU8E-OE1...HN2	2.00		
					VAL24C-O...HN2	2.09		
					LEU65-OXT...HN1	1.94		
					GLU8A.-OE1...HN1	1.92		
					ASN15A-OD1...HN2	2.06		
DABCOB	-34.4 -33.4 -33.4	-13.9 -16.4 -13.8	-6.9 -1.1 -7.0	-9.9 -9.8 -9.9	LEU65B-O...HN1	2.09	-3.7 -6.4 -2.7	<i>out</i> <i>in</i> <i>out</i>
					VAL47C-O...HN2	2.03		
					ASN15C-OD1...HN1	2.14		
					GLU8E-OE1...HN2	2.08		
					LEU9C-OXT...HN2	1.97		
					VAL47C-O...HN1	1.98		
DABCO3B	-46.5 -44.9 -44.4	-28.4 -31.0 -28.6	-4.9 -3.7 -5.3	-8.9 -9.9 -9.6	GLU8E-OE1...HN2	2.25	-3.2 -0.2 -0.9	<i>in</i> <i>out</i> <i>out</i>
					ASN15C-OD1...HN1	2.12		
					LEU65A-OXT...HN1	2.10		
					LEU21C-O...HN2	2.18		
					LEU65B-OXT...HN1	2.01		
					VAL47C-O...HN2	1.98		
DABCON	-38.8 -35.9 -35.9	-16.0 -17.3 -24.1	-3.2 -4.4 2.9	-9.7 -9.9 -9.7	GLU8A-OE2...HN1	2.06	-9.9 -4.3 -4.9	<i>in</i> <i>out</i> <i>out</i>
					ASN15C-OD1...HN2	2.14		
					PHE20D-O...HN1	2.07		
					LEU65C-OXT...HN2	1.93		
					SER50B-O...HN1	2.12		
					TYNB-OH...HN2	2.08		
DABCO3N	-58.1 -52.3 -51.9	-34.1 -38.1 -40.1	-6.8 -3.4 -0.2	-9.1 -5.0 -4.8	ASN15C-OD1...HN1	2.19	-8.1 -5.8 -5.9	<i>in</i> <i>out</i> <i>out</i>
					GLU8E-OE1...HN2	2.15		
					LEU65C-OXT...HN1	2.01		
					GLU8A-OE1...HN1	2.03		

Примечание. ePL— полная энергия связывания блокатора с белком E, ккал/моль;  
 eVDW— энергия Ван-дер-Ваальсовых взаимодействий;  
 eCoul— энергия электростатических взаимодействий;  
 eHb— энергия водородных связей;  
 Y— акцептор водородной связи молекулы блокатора с атомами E-канала;  
 HX— донор водородной связи;  
 eSolv— взаимодействие с растворителем;  
*in(out)*— позиция связывания лиганда, внутри/вне E-канала.

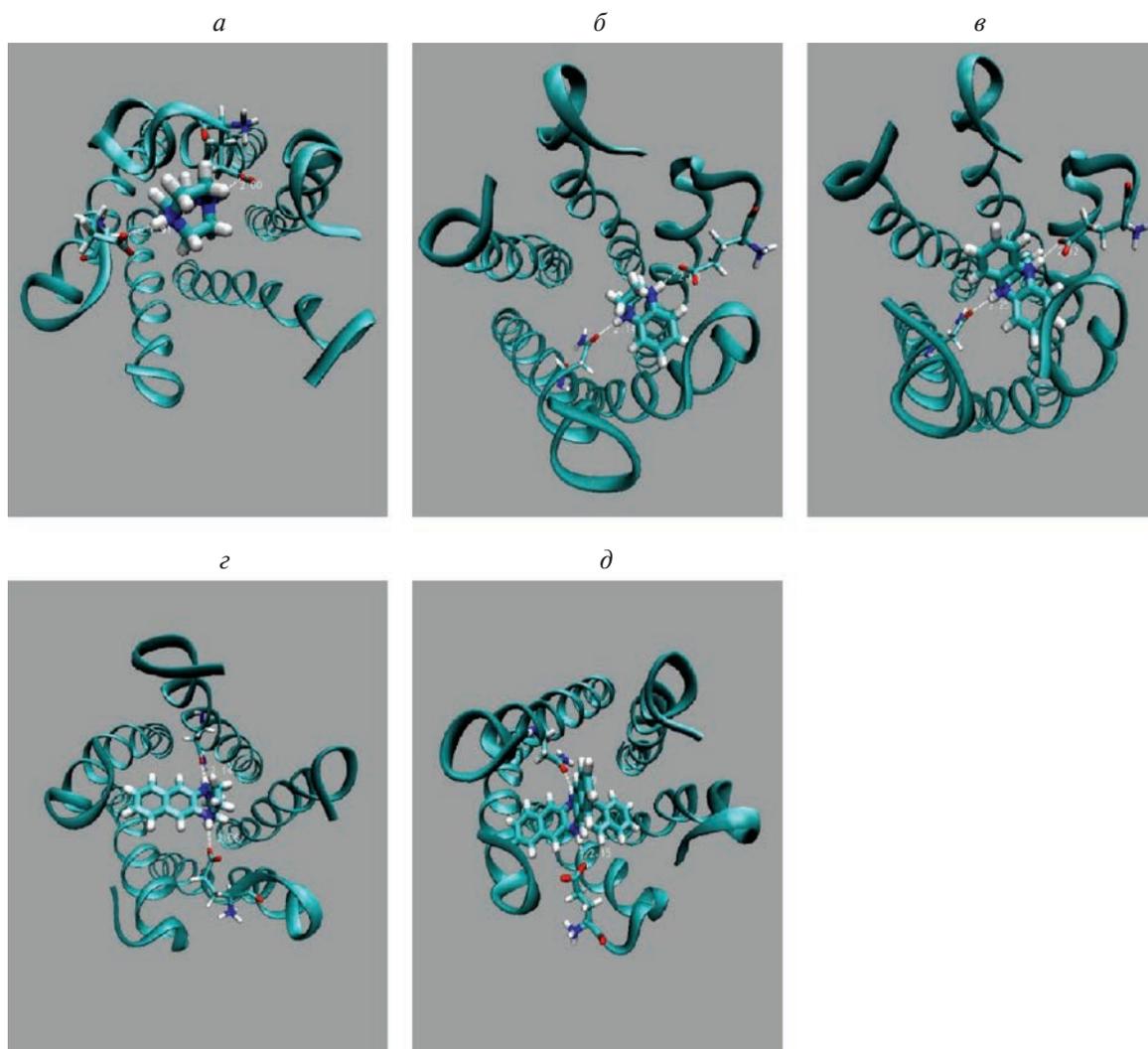


Рис. 5. Структура наивероятных, с наибольшей энергией связи, комплексов нативного E-канала с оптимально связанными молекулами блокаторов: DABCO (а); DABCOB (б); DABCO3B (в); DABCON (г); DABCO3N (д).

демонстрируют наибольшую степень взаимного перекрытия (рис. 4). Значительное перекрытие потенциальных энергий вдоль МД траекторий показывает, что нейтральные остатки ASN15A(B, C, D, E) могут с большой вероятностью спонтанно превращаться в отрицательно заряженный аминокислотный остаток ASP15A(B, C, D, E).

Переход между нейтральным и заряженным состояниями остатков может происходить в различных комбинациях, ионизированный/нейтральный в каждой из пяти субъединиц. Следовательно, появляется новая степень свободы, т.е. состояние ионизации ASN/ASP, у пяти остатков ASP15A(B, C, D, E), что приводит к уменьшению энтропии ионизационного состояния на  $\sim -kT \times \ln N_{iz}$ , где  $N_{iz} = 2^5$  — это количество ионизационных состояний пяти остатков ASN15/ASP15(A, B, C, D, E). Большое перекрытие флуктуаций энергий нативных, *mut0*, и мутантных, *mut1*, структур, повышает вероятность спонтанных переходов между

нативной структурой и *mut1*. Перекрытие энергий между нативным и мутантными, *mut2*, состояниями мало, поэтому спонтанные переходы в нейтральное состояние остатков GLU8A(B, C, D, E) при нейтральных остатках ASN15A(B, C, D, E) маловероятно. Результаты докинга бициклических и трициклических производных DABCO на нативный белок E показаны на рис. 5 и в табл. 7.

#### Связывание полициклических производных DABCO с мутантными структурами E-каналов

Мутации аминокислотных остатков белка E, образующих внутреннюю поверхность канала, ASN15/ASP15 и GLU8/GLN8 случаются с вероятностями  $42 \times 10^{-4}$  и  $35 \times 10^{-4}$  соответственно [55]. Рассмотрены два наивероятных мутанта нативного E-канала, *mut1*, мутация ASN15/ASP15 во всех цепях (A, D, C, D, E) белка E и мутант *mut2*, GLU8/GLN8. Мутантные структуры стабильны при длительном моделировании методом МД (рис. 4).

Средние энергии по траекториям МД обеих мутантных структур выше, чем у нативного белка *E*, *mut1* менее выгодна на 30.1, а мутант *mut2* на 94.4 ккал/моль (табл. 6). Амплитуды тепловых флуктуаций потенциальных энергий, по траекториям МД, составляют 53.2 (для *mut0*), 50.1 (*mut1*), 47.3 (*mut2*), ккал/моль. Достаточно большие амплитуды тепловых флуктуаций потенциальных энергий структур *mut0* и *mut1* и их перекрытие (рис. 4) свидетельствуют, что мутанты *mut0* и *mut1* могут спонтанно переходить один в другой, ASN15/ASP15, с высокой вероятностью при комнатной температуре.

Основные моды связывания молекул блокаторов DABCO, DABCOB, DABCO3B, DABCON и DABCO3N с белком *E* и энергии их взаимодействия с мутантом *mut1* приведены на рис. 6 и в табл. 8. Отметим, что наиболее сильной модой связывания для всех пяти лигандов (рис. 1) являются структуры с лигандом внутри Е-канала. Энергии связывания молекул блокаторов с мутантом *mut1* (ASN15/ASP15) больше на ~5–11 ккал/моль, по

сравнению с энергиями связывания с нативным белком *E*. Это связано с положительным зарядом лигандов (+2 е.у.), производных DABCO и возрастания электростатических взаимодействий с мутантными отрицательно заряженными остатками ASP15(A, B, C, D, E). Основные сайты связывания характеризуются образованием двух водородных связей между ASP15 из двух разных белковых цепей Е-канала (табл. 8) с двумя атомами азота молекул лигандов (рис. 6). Табл. 8 показывает, что молекулы DABCO3B и DABCO3N связываются наиболее сильно со внутренней поверхностью канала *E*, тогда как более простые DABCOB и DABCON связаны заметно слабее, на 13–15 ккал/моль. Увеличение энергии связывания внутри Е-канала обусловлено более сильными электростатическими и Ван-дер-Ваальсовыми взаимодействиями положительно заряженных молекул блокатора с мутантным, *mut1* ASN15/ASP15, Е-каналом (табл. 8).

Второй вероятный мутант Е-канала, *mut2*, GLU8/GLN8, во всех пяти цепях А, В, С, D, Е

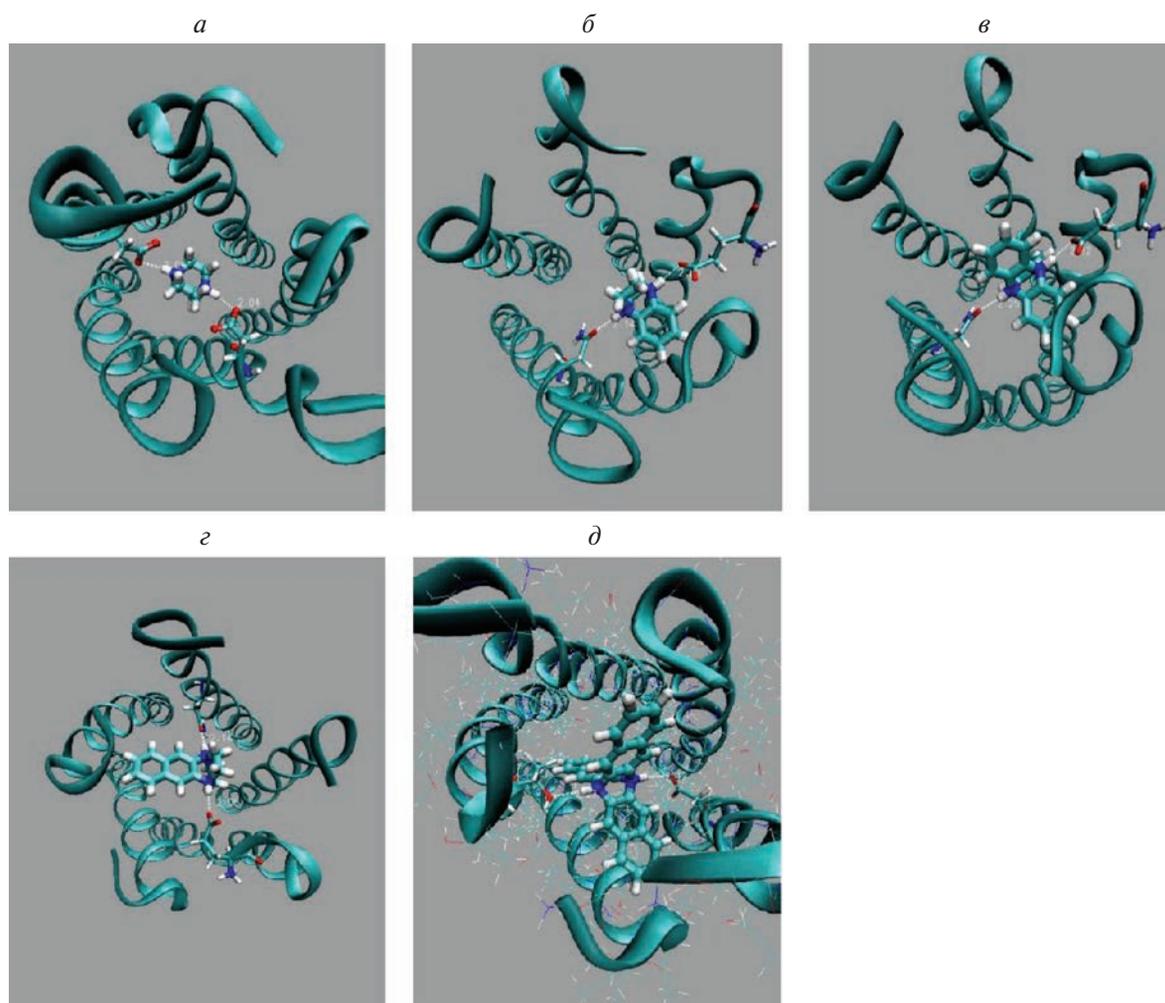


Рис. 6. Структура наиболее вероятных комплексов мутантного, *mut1*, Е-канала с оптимально связанными молекулами блокаторов: DABCO (а), DABCOB (б), DABCO3B (в), DABCON (г), DABCO3N (д).

**Таблица 8.** Сайты связывания производных DABCO с мутантным, *mut1*, E-белком вируса SARS-CoV-2 [Glu8-Asp<sup>15</sup>(A, B, C, D, E)]

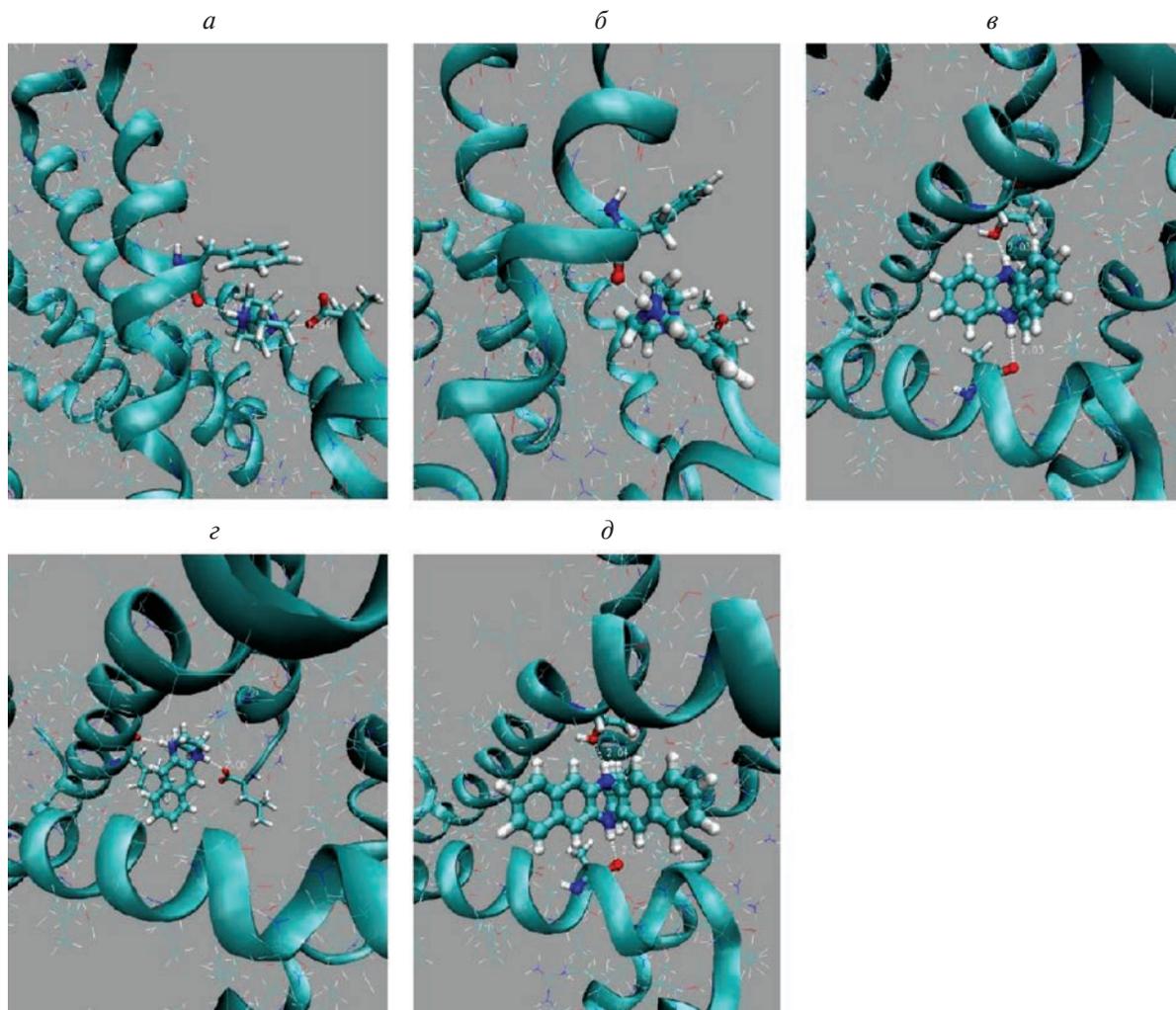
Молекула-блокатор	ePL	eVDW	eCou	eHb	Y...HX	Расстояние, Å	eSolv	<i>in(out)</i>
DABCO <sup>a</sup>	-36.5 -35.3 -30.3	-10.2 -10.5 -15.1	-15.1 -12.7 -6.0	-9.8 -10.0 -9.8	ASP15A-OD1...HN1	2.04	-1.4 -1.1 0.3	<i>in</i> <i>in</i> <i>in</i>
					ASP15C-OD2...HN2	2.06		
					ASP15A-OD2...HN1	1.88		
					ASP15C-OD1...HN2	2.0		
					ASP15A-OD1...HN1	2.05		
					THR11B-OG1 HN2...	2.06		
DABCOB	-43.0 -38.4 -36.6	-15.8 -18.9 -18.5	-16.8 -8.0 -10.2	-9.9 -10.0 -8.6	ASP15A-OD1...HN1	1.90	-0.8 -1.4 0.7	<i>in</i> <i>in</i> <i>in</i>
					ASP15C-OD1...HN2	2.04		
					ASP15A-OD1...HN1	1.93		
					THR11D-OG1 HN2	2.04		
					ASP15B-OG1...HN2	1.99		
					LEU12A-O...HN1	2.26		
DABCO3B	-56.5 -47.2 -46.2	-32.3 -29.8 -32.9	-17.4 -11.0 -4.9	-9.7 -9.4 -9.7	ASP15A-OD2...HN1	2.09	2.8 3.0 1.3	<i>in</i> <i>in</i> <i>in</i>
					ASP15D-OD1...HN2	2.05		
					ASP15A-OD1...HN1	2.13		
					THR11D-OG1...HN2	2.11		
					ASP15A-OD1...HN1	2.02		
					THR11B-O...HN2	2.10		
DABCON	-45.2 -42.4 -38.6	-19.7 -19.3 -23.5	-16.1 -11.9 -5.7	-9.9 -10.1 -9.1	ASP15A-OD1...HD1	1.95	0.6 -1.1 -0.3	<i>in</i> <i>in</i> <i>in</i>
					ASP15C-OD1...HN2	2.04		
					ASP15A-OD1...HN1	1.95		
					ASP15C-OD1...HN2	2.04		
					ASP8-OD1...HN1	1.95		
					LEU63-O...HN2	2.12		
DABCO3N	-60.5 -57.3 -57.7	-47.0 -45.0 -54.1	-5.3 -9.8 3.4	-8.9 -4.8 -9.1	ASP15A-OD2...HN1	2.12	0.7 2.3 2.0	<i>in</i> <i>in</i> <i>in</i>
					THR11D-OG1...HN2	2.23		
					ASP15A-OG1...HN1	2.04		
					THR11B-O...HN2	2.40		
					SER60D-OG...HN1	2.21		
					ALA43A-O...HN2	2.15		

Примечание. ePL – полная энергия связывания блокатора с белком E, ккал/моль;  
eVDW – энергия Ван-дер-Ваальсовых взаимодействий;  
eCou – энергия электростатических взаимодействий;  
eHb<sup>a</sup> – энергия водородных связей;  
Y – акцептор водородной связи молекулы блокатора с атомами E-канала;  
HX – донор водородной связи;  
eSolv – взаимодействие с растворителем;  
*in(out)* – позиция связывания лиганда, *in/out* – внутри/вне E-канала.

характеризуется заменой заряженных кислотных остатков GLU8 (A, B, C, D, E) на нейтральные GLN8. Моделирование мутанта GLU8/GLN8 показывает заметное увеличение энергии структуры *mut2* на ~ 60 ккал/моль относительно энергии мутанта *mut1* и на ~90 ккал/моль относительно нативного белка E. Структуры сайтов, предпочтительных по энергии связывания, для пяти молекул лигандов DABCO, DABCOB, DABCO3B, DABCON, DABCO3N, с мутантом *mut2* E-канала приведены в табл. 9. Наиболее предпочтительные структуры связывания пяти молекул блокаторов с мутантным, *mut2*, E-каналом показаны на рис. 7. Видно, что только молекула блокатора, DABCO3N, имеет сайт

связывания внутри канала E (табл. 9) с энергией, близкой к энергии оптимального сайта связывания вне канала. То есть молекула блокатора DABCO3N, связываясь в E-канале, может блокировать канал с вероятностью ~1/2.

Анализ результатов для мутанта E-канала, *mut2*, табл. 9 и сравнение с расчетами для нативной, *mut0*, структуры с зарядом 1 е.у., мутантом *mut1* с зарядом 2 е.у., показывает, что нейтральный мутант, *mut2*, E-канала обладает наименьшей вероятностью связывания молекул блокаторов внутри E-канала. Можно отметить, что энергия связывания молекулы лиганда DABCO3N линейно коррелирует с зарядами остатков, формирующих вход в E-канал,



**Рис. 7.** Структура наиболее вероятных комплексов мутанта E канала, *mut2*, с молекулами блокаторов: (а) DABCO; (б) DABCOB; (в) DABCO3B; (г) DABCON; (д) DABCO3N.

энергии равны  $-61.5$ ,  $-58.1$  и  $-54.6$  ккал/моль, соответственно, для структур *mut1*, *mut0*, *mut2*, имеющих заряды  $-2$ ,  $-1$  и  $0$  е.у. соответственно.

Анализ результатов, представленных в табл. 6–9, показывает, что для положительно заряженных  $+2$  е.у., молекул блокаторов DABCO, DABCOB, DABCO3B, DABCON, DABCO3N критически важны электростатические взаимодействия с нативной и мутантными структурами ионного канала E. Вероятность захвата ионов  $H^+/K^+$  в канал E с последующей диффузией ионов вдоль канала будет значительно снижена для нейтрального мутанта *mut2*, по сравнению со связыванием для заряженных нативной,  $-1$  е.у., и *mut1*,  $-2$  е.у., структур E-канала.

### ЗАКЛЮЧЕНИЕ

В работе представлены результаты конструирования молекул лекарственных препаратов, положительно,  $+2$  е.у., заряженных производных

диазабициклооктана, которые могут эффективно связываться внутри M2- и E-каналов вирусов гриппа A и SARS-CoV-2 и блокировать транспорт ионов  $H^+/K^+$  через клеточную мембрану, нарушая функционирование вирусов. В физиологических условиях проведено моделирование тепловой динамики структур нативных белков, формирующих M2- и E-каналы, а также наиболее вероятных мутантных структур каналов. Методом исчерпывающего докинга определены наиболее вероятные позиции и энергии связывания серии положительно заряженных молекул блокаторов с нативными и мутантными структурами M2- и E-каналов в физиологических условиях. Определены оптимальные структуры молекул блокаторов, связывающихся в M2- и E-каналах с высокой энергией,  $30$ – $50$  ккал/моль.

Исследование поддержано грантом РФФИ № 18-04-00005-а и бюджетным проектом, финансируемым Правительством Российской Федерации для Института химической

**Таблица 9.** Сайты связывания производных DABCO с мутантным, *mut2*, E-каналом вируса SARS-CoV-2 [Gln8Asn15]

Молекула блокатор	ePL	eVDW	eCoul	eHb	Y ... HX Dist, Å	eSolv	in(out)
DABCO	-25.6	-17.7	8.1	-9.8	THR35C-OG1...HN1 2.09	-6.8 -3.8 -3.7	out out in
	-25.3	-11.9	0.1	-9.7	ASN45B-OD1...HN2 2.06		
	-18.5	-13.6	8.3	-9.5	LEU65C-OXT...HN1 1.94		
					PHE20D-O...HN2 2.14		
					ASN15B-OD1...HN1 2.11		
DABCOB	-32.9	-14.7	-4.3	-9.7	PHE20D-O...HN1 2.15	-4.3 -6.4 -2.7	out out in
	-32.6	-18.9	2.5	-9.8	LEU65C-O...HN2 1.87		
	-28.9	-19.6	3.0	-9.6	ASN45B-OD1...HN1 2.03		
					THR35C-O...HN2 2.15		
					ASN15C-OD1...HN1 2.04		
DABCO3B	-43.6	-35.3	0.8	-9.4	ALA24E-O...HN1 2.03	0.2 -0.6 -2.9	out out in
	-41.0	-33.5	1.3	-8.2	THR30D-OE1...HN2 2.02		
	-31.1	-24.1	5.7	-9.8	LEU37D-O...HN1 2.30		
					ASN45C-OD1...HN2 2.03		
					GLN10A-O...HN1 2.07		
DABCON	-37.4	-21.1	-1.8	-9.8	THR9B-OG1...HN2 2.15	-4.7 -5.4 -0.9	out out in
	-33.8	-18.8	0.2	-9.8	PHE20D-O...HN1 2.08		
	-33.9	-25.7	2.4	-9.7	LEU65C-OXT...HN2 2.03		
					LEU65B-O...HN1 1.90		
					THR55C-OG1...HN2 2.13		
DABCO3N	-54.6	-45.3	1.5	-9.3	ASN15C-OG1...HN1 2.06	-1.5 -6.2 -1.4	in out out
	-53.7	-40.4	1.9	-9.0	ASN15C-OG1...HN2 2.12		
	-45.3	-50.9	2.4	-5.0	ALA32E-O...HN1 2.04		
					THR30D-OG1H...N2 2.04		

Примечание. ePL – полная энергия связывания лиганда с белком E, ккал/моль;

eVDW – энергия электростатических взаимодействий;

eCoul – энергия электростатических взаимодействий;

eHb – энергия водородных связей;

Y – акцептор водородной связи молекулы блокатора с атомами E-канала;

HX – донор водородной связи;

eSolv – взаимодействие с растворителем;

in/out – позиция связывания лиганда, внутри/вне E-канала.

биологии и фундаментальной медицины Сибирского отделения Российской академии наук: № АААА-А17-117020210022-4.

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

Автор заявляет об отсутствии конфликта интересов.

## СПИСОК ЛИТЕРАТУРЫ

- Rudneva I., Ignatieva A., Timofeeva T., Shilov A., Kushch A., Masalova O., Klimova R., Bovin N., Mochalova L., Kaverin N. (2012) Escape mutants of pandemic influenza A/H1N12009 virus: variations in antigenic specificity and receptor affinity of the hemagglutinin. *Virus Res.* **166**, 61–67. doi: 10.1016/j.virusres.2012.03.003
- Wever P.C., Bergen L. (2014) Death from 1918 pandemic influenza during the First World War: a perspective from personal and anecdotal evidence. *Influenza Other Respir. Viruses.* **8**, 538–546. doi:10.1111/irv.12267
- Ullrich S., Ekanayake K.B., Otting G., Nitsche C. (2022) Main protease mutants of SARS-CoV-2 variants remain susceptible. *Bioorg. Med. Chem. Lett.* **62**, 128629. doi:10.1016/j.bmcl.2022.128629
- Wang Q., Ye S.B., Zhou Z.J., Song A.L., Zhu X., Peng J.M., Liang R.M., Yang C.H., Yu X.W., Huang X., Yu J., Qiu Y., Ge X.Y. (2022) Key mutations in the spike protein of SARS-CoV-2 affecting neutralization resistance and viral internalization. *J. Med Virol.* **95**, e28407. https://doi.org/10.1002/jmv.28407
- Stouffer A.L., Acharya R., Salom D., Levine A.S., Costanzo L.D., Soto C.S., Tereshko V., Nanda V., Stayrook S., DeGrado W.E. (2008) Structural basis for the function

- and inhibition of an influenza virus proton channel. *Nature*. **451**(7178), 596–599. doi:10.1038/nature06528
6. Vorobjev Y.N. (2021) An effective molecular blocker of ion channel of M2 protein as anti-influenza drug. *J. Biomol. Struct. Dynamics*. **39**(7), 2352–2363. doi.org/10.1080/07391102.2020.1747550
  7. Mandala V.S., McKay M.J., Shcherbakov A.A., Dregni A.J., Kolocouris A., Hang M. (2022) Structure and drug binding of the SARS-CoV-2 envelope protein transmembrane domain in lipid bilayers. *Nat. Struct. Mol. Biol.* **27**, 1202–1208. doi:10.1038/s41594-020-00536-8
  8. Surya W., Li Y., Torres J. (2018) Structural model of the SARS coronavirus E channel in LMPG micelles. *Biochim. Biophys. Acta Biomembr.* **1860**, 1309–1317. doi.org/10.1016/j.bbmem.2018.02.017
  9. Pinto L.H., Holsinger L.J., Lamb R.A. (1992) Influenza virus M2 protein has ion channel activity. *Cell*. **69**(3), 517–528. doi.org/10.1016/0092-8674(92)90452-I
  10. Ruch T.R., Machamer C.E. (2011) The hydrophobic domain of infectious bronchitis virus E protein alters the host secretory pathway and is important for release of infectious virus. *J. Virol.* **85**, 675–685. doi: 10.1128/JVI.01570-10
  11. Nieto-Torres J.L., DeDiego D.I., Verdià-Báguena C., Jimenez-Guardeño J.M., Regla-Nava J.A., Fernandez-Delgado R., Castaño-Rodríguez C., Alcaraz A., Torres J., Aguilera V.M., Enjuanes L. (2014) Severe acute respiratory syndrome coronavirus envelope protein ion channel activity promotes virus fitness and pathogenesis. *PLoS Pathog.* **10**(5), e1004077. doi:10.1371/journal.ppat.1004077
  12. Singh Tomar P.P., Arkin I.T. (2020) SARS-CoV-2 E protein is a potential ion channel that can be inhibited by Gliclazide and Memantine. *Biochem. Biophys. Res. Commun.* **530**, 10–14. doi: 10.1016/j.bbrc.2020.05.206
  13. Gurumallappa, Arun Renganathan R.R., Hema M.K., Karthik C.S., Rani S., Nethaji M., Jayanth H.S., Mallu P., Lokanath N.K., Ravishankar Rai V. (2021) 4-acetamido-3-nitrobenzoic acid — structural quantum chemical studies, ADMET and molecular docking studies of SARS-CoV2. *J. Biomol. Struct. Dynamics*. **40**(14), 6656–6670. doi:10.1080/07391102.2021.1889664
  14. Schoeman D., Fielding B.C. (2019) Coronavirus envelope protein: current knowledge. *J. Virol.* **16**, 69–91. doi:10.1186/s12985-019-1182-0
  15. DeDiego M.L., Alvarez E., Almazán F., Rejas M.T., Lami-rande E., Roberts A., Shieh W.-J., Zaki S.R., Subbarao K., Enjuanes L. (2007) A severe acute respiratory syndrome coronavirus that lacks the E gene is attenuated in vitro and in vivo. *J. Virol.* **81**, 1701–1713. doi:10.1128/JVI.01467-06
  16. Torres J., Maheswari U., Parthasarathy K., Lifang Ng, Liu D.X., Gong X. (2007) Conductance and amantadine binding of a pore formed by a lysine-flanked transmembrane domain of SARS coronavirus envelope protein. *Protein Sci.* **16**, 2065–2071. doi:10.1110/ps.062730007
  17. Verdià-Báguena C., Nieto-Torres J.L., Alcaraz A., DeDiego M.L., Torres J., Aguilera V.M., Enjuanes L. (2012) Coronavirus E protein forms ion channels with functionally and structurally-involved membrane lipids. *Virology*. **432**, 485–494. doi:10.1016/j.virol.2012.07.005
  18. Surya W., Li Y., Verdià-Báguena C., Aguilera V.M., Torres J. (2015) MERS coronavirus envelope protein has a single transmembrane domain that forms pentameric ion channels. *Virus Res.* **201**, 61–66. doi:10.1016/j.virusres.2015.02.023
  19. Wilson L., Gage P., Ewart G. (2006) Hexamethylene amiloride blocks E protein ion channels and inhibits coronavirus replication. *Virology*. **353**, 294–306. doi:10.1016/j.virol.2006.05.028
  20. To J., Surya W., Fung T.S., Li W., Verdià-Báguena C., Queralt-Martin M., Aguilera V.M., Liu D.X., Torres J. (2017) Channel-inactivating mutations and their revertant mutants in the envelope protein of infectious bronchitis virus. *J. Virol.* **91**, e02158–16. doi:10.1128/JVI.02158-16
  21. Pervushin K., Tan E., Parthasarathy K., Lin X., Jiang F.L., Yu D., Vararattanavech A., Soong T.W., Liu D.X., Torres J. (2009) Structure and inhibition of the SARS coronavirus envelope protein ion channel. *PLoS Pathog.* **5**, e1000511. doi:10.1371/journal.ppat.1000511
  22. Torres J., Parthasarathy K., Lin X., Saravanan R., Kukol A., Liu D.X. (2009) Model of a putative pore: the pentameric alpha-helical bundle of SARS coronavirus E protein in lipid bilayers. *Biophys. J.* **91**, 938–947. doi:10.1529/biophysj.105.080119
  23. Mandala V.S., McKay M.J., Shcherbakov A.A., Dregni A.J., Kolocouris A., Hong M. (2020) Structure and drug binding of the SARS-CoV-2 envelope protein transmembrane domain in lipid bilayers. *Nat. Struct. Mol. Biol.* **27**, 1202–1208. https://doi.org/10.1038/s41594-020-00536-8
  24. Parthasarathy K., Ng L., Lin X., Liu D.X., Pervushin K., Gong X., Torres J. (2008) Structural flexibility of the pentameric SARS coronavirus envelope protein ion channel. *Biophys. J.* **95**, L39–L41. doi:10.1529/biophysj.108.133041
  25. Hanning Chen, Yujie Wu, Voth G.A. (2007) Proton transport behavior through the influenza A M2 channel: Insights from molecular simulation. *Biophys. J.* **93**, 3470–3479. doi: 10.1529/biophysj.107.105742
  26. Tang Y., Zaitseva F., Lamb R.A., Pinto L.H. (2002) The gate of the influenza virus M2 proton channel is formed by a single tryptophan residue. *J. Biol. Chem.* **277**, 39880–39886. doi:10.1074/jbc.M206582200
  27. Воробьев Ю.Н. (2020) Конструирование эффективного ингибитора ионного канала белка M2 вируса гриппа А. *Молекуляр. биология*. **54**, 321–332.
  28. Vorobjev Y.N. (2021) Design of effective molecular blocker of e protein channel as anti SARS-CoV-2 virus. *Drug Des.* **10**, 183.
  29. Vorobjev Y.N. (2022) Molecular blocker of native and mutant E protein ion channel of SARS-CoV-2 virus. *Drug Des.* **11**, 205.
  30. Попов А.В., Воробьев Ю.Н. (2010) Программа GUI-BioPASED для моделирования молекулярной биополимеров с графическим пользовательским интерфейсом. *Молекуляр. биология*. **44**(4), 735–742.
  31. Vorobjev Y.N., Almagro J.C., Hermans J. (1998) Discrimination between native and intentionally misfolded conformations of proteins: ES/IS, a new method for calculating conformational free energy that uses both dynamics

- simulations with an explicit solvent and an implicit solvent continuum model. *Proteins*. **32**, 399–413. [https://doi.org/10.1002/\(SICI\)1097-0134\(19980901\)32:4<399::AID-PROT1>3.0.CO;2-C](https://doi.org/10.1002/(SICI)1097-0134(19980901)32:4<399::AID-PROT1>3.0.CO;2-C)
32. Vorobjev Y.N. (2011) Advances in implicit models of water solvent to compute conformational free energy and molecular dynamics of proteins at constant pH. *Adv. Protein Chem. Struct. Biol.* **85**, 281–322. <https://doi.org/10.1016/B978-0-12-386485-7.00008-9>
  33. Vorobjev Y.N. (2014) Modeling of Electrostatic Effects in Macromolecules. Ed. Liwo A. In: *Computational Methods to Study the Structure and Dynamics of Biomolecules and Biomolecular Processes: From Bioinformatics to Molecular Quantum Mechanics*. Springer-Verlag Berlin Heidelberg, 1, pp. 151–194. doi: 10.1007/978-3-642-28554-7\_6
  34. Vorobjev Y.N., Scheraga H.A., Vila J.A. (2018) Coupled molecular dynamics and continuum electrostatic method to compute the ionization pKa's of proteins as a function of pH. Test on a large set of proteins. *J. Biomol. Struct. Dynamics*. 563–574. doi:10.1080/07391102.2017.1288169
  35. Cornell W.D., Cieplak P., Bayly C.I., Gould I.R., Merz K.M., Ferguson D.M., Spellmeyer D.C., Fox T., Caldwell J.W., Kollman P.A. (1995) A second generation for the simulation of proteins, nucleic acids and organic molecules. *J. Am. Chem. Soc.* **117**, 5179–5197. doi: 10.1021/ja00124a002
  36. Wang J., Cieplak P., Kollman P.A. (2000) How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J. Comput. Chem.* **21**, 1049–1074. doi.org/10.1002/1096-987X(200009)21:12%3C1049::AID-JCC3%3E3.0.CO;2-F
  37. Wang J., Wolf R.M., Caldwell J.W., Kollman P.A., Case D.A. (2004) Development and testing of a general amber force fields. *J. Comput. Chem.* **27**, 1157–1174. doi:10.1002/jcc.20035
  38. Vorobjev Y.N. (2010) Blind docking method combining search of low-resolution binding sites with ligand pose refinement by molecular dynamics-based global optimization. *J. Comp. Chem.* **31**, 1080–1092. doi: 10.1002/jcc.21394
  39. Mohga F.M., Youssef A.A.A. (2001) Dielectric permittivity and AC conductivity investigation for the new model lipid bilayer material: (CH<sub>2</sub>)<sub>10</sub> (NH<sub>3</sub>)<sub>2</sub>CdCl<sub>4</sub> Z. *Naturforsch.* **56A**, 568–578.
  40. Vorobjev Y.N., Hermans J. (1997) SIMS, computation of a smooth invariant molecular surface. *Biophysical J.* **73**, 722–732. doi: 10.1016/S0006-3495(97)78105-0
  41. Löffler G., Schreiber H. (1997) The frequency-dependent conductivity of a saturated solution of ZnBr<sub>2</sub> in water: A molecular dynamics simulation. *J. Chem. Phys.* **107**, 3135–3143. doi.org/10.1063/1.474703
  42. Li C., Li L., Zhang J., Alexov E. (2012) Highly efficient and exact method for parallelization of grid-based algorithms and its implementation in DelPhi. *J. Comput. Chem.* **33**, 1960–1966. doi:10.1002/jcc.23033
  43. Li L., Li C., Zhang Z., Alexov E. (2013) On the dielectric “Constant” of proteins: smooth dielectric function for macromolecular modeling and its implementation in DelPhi. *J. Chem. Theory Comput.* **9**, 2126–2136. <https://doi.org/10.1021/ct400065j>
  44. Moffat J.S., Vijayvergiya V., Gao P.F., Cross T.A., Woodbury D.J., Busath D.D. (2008) Proton transport through influenza A virus M2 protein reconstituted in vesicles. *Biophys. J.* **94**, 434–445. doi: 10.1529/biophysj.107.109082
  45. Hari Z.S., Moorthy N., Poongavanam V., Pratheepa V. (2014) Viral M2 Ion channel protein: a promising target for anti-influenza drug discovery. *Med. Chem.* **14**, 819–830. doi: 10.2174/138955751410141020150822
  46. Olsen R.W. (2006) Picrotoxin-like channel blockers of GABA<sub>A</sub> receptors. *Proc. Nat. Acad. Sci. USA.* **103**, 6081–6082. [www.pnas.org/cgi/doi/10.1073/pnas.0601121103](http://www.pnas.org/cgi/doi/10.1073/pnas.0601121103)
  47. Thomaston J.L., Polizzi N.F., Konstantinidi A., Wang J., Kolocouris A., DeGrado W.F. (2018) Inhibitors of the M2 proton channel engage and disrupt transmembrane networks of hydrogen-bonded waters. *J. Am. Chem. Soc.* **140**, 15219–15226. doi: 10.1021/jacs.8b06741
  48. Mustafa M., Henderson D.J., Busath D.D. (2009) Free-energy profiles for ions in the influenza M2-TMD channel. *Proteins*. **76**, 794–807. doi:10.1002/prot.22376
  49. Homeyer N., Ioannidis H., Kolarov F., Gauglitz G., Zikos C., Kolocouris A., Gohlke H. (2016) Interpreting thermodynamic profiles of aminoadamantane compounds inhibiting the m2 proton channel of influenza A by free energy calculations. *J. Chem. Inf. Model.* **56**, 110–126. doi: 10.1021/acs.jcim.5b00467
  50. Holsinger L.J., Nichani D., Pinto L.H., Lamb R.A. (1994) Influenza A virus M2 ion channel protein: a structure-function analysis. *J. Virol.* **68**, 1551–1563. doi: 10.1128/jvi.68.3.1551-1563.1994
  51. Wang J., Kim S., Kovacs F., Cross T.A. (2001) Structure of the trans membrane region of the M2 protein H<sup>+</sup> channel. *Protein Sci.* **10**, 2241–2250. doi:11.1110/ps.17901
  52. Cady S.D., Mishanina T.V., Hong M. (2009) Structure of amantadine-bound M2 transmembrane peptide of influenza A in lipid bilayers from magic-angle-spinning solid-state NMR: the role of Ser31 in amantadine binding. *J. Mol. Biol.* **385**, 1127–1141. doi.org/10.1016/j.jmb.2008.11.022
  53. Cady S.D., Schmidt-Rohr K., Wang J., Soto C.S., DeGrado W.F., Hong M. (2010) Structure of the amantadine binding site of influenza M2 proton channels in lipid bilayers. *Nature*. **463**, 689–692. <https://doi.org/10.1038/nature08722>
  54. Hong M., DeGrado W.F. (2012) Structural basis for proton conduction and inhibition by the influenza M2 protein. *Protein Sci.* **21**, 1620–1633. doi: 10.1002/pro.2158
  55. PAM1 Mutation Matrix (1978) Atlas of protein sequence and structure. **5**(Suppl. 3). <http://profs.scienze.univr.it/~liptak/ACB/files/PAMMutationMatrices.pdf>

## Molecular Ion Channel Blockers of Influenza A and SARS-CoV-2 Viruses

Yu. N. Vorobjev<sup>1, \*</sup>

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

*\*e-mail: ynvorob@niboch.nsc.ru*

Drug molecules that block the functional cycle of influenza A and SARS-CoV-2 viruses are proposed. The blocker molecules effectively binding inside the M2 and E-channels of influenza A and SARS-CoV-2 viruses and blocking the diffusion of H<sup>+</sup>/K<sup>+</sup> ions destroy the functional cycle of viruses. A family of positively charged, +2 e.u., molecular blockers of H<sup>+</sup> /K<sup>+</sup> ion diffusion through M2 and E-channels is proposed. The blocker molecules, derivatives of diazabicyclooctane, was investigated for its binding affinity to the channels M2 and E. Thermal dynamics and binding affinity were modeled by the exhaustive docking method sites. Blocker molecules with higher affinity for the blocking sites were proposed. The most probable mutations of amino acids of protein M2 and E channels are considered, the effectiveness of channel blocking are analyzed and optimal structures of blocker molecules are proposed.

**Keywords:** influenza A and SARS-CoV-2 viruses, M2 and E protein ion channel, molecular dynamics, M2 and E ion channel blockers, diazabicyclooctane derivatives, binding and blocking of M2 and E channels