ФЕДЕРАЛЬНОЕ БЮДЖЕТНОЕ УЧРЕЖДЕНИЕ НАУКИ
«ГОСУДАРСТВЕННЫЙ НАУЧНЫЙ ЦЕНТР ВИРУСОЛОГИИ И БИОТЕХНОЛОГИИ «ВЕКТОР» ФЕДЕРАЛЬНОЙ СЛУЖБЫ ПО НАДЗОРУ В СФЕРЕ ЗАЩИТЫ ПРАВ ПОТРЕБИТЕЛЕЙ И БЛАГОПОЛУЧИЯ ЧЕЛОВЕКА

На правах рукописи

Алиев Тимур Илхамович

Создание двойных редакторов оснований MmCas12m-TadDE и dSpCas9-TadDE и системы их доставки для мутагенеза гена *gag* провируса ВИЧ-1

1.5.3 – Молекулярная биология

Диссертация на соискание учёной степени
кандидата биологических наук

Научный руководитель:
кандидат биологических наук
Иматдинов Ильназ Рамисович

Кольцово – 2026

ОГЛАВЛЕНИЕ

СПИСОК СОКРАЩЕНИЙ 7
ВВЕДЕНИЕ 9
ГЛАВА 1. ОБЗОР ЛИТЕРАТУРЫ 21
1.1.ВИЧ-инфекция 21
1.2.Вирус иммунодефицита человека 1-го типа 23
1.3.Генная терапия ВИЧ-инфекции 30
1.4.CRISPR-Cas системы в геномном редактировании 36
1.5.Разнообразие редакторов оснований ДНК 41
1.5.1.Редакторы оснований цитозина (Cytosine base editors – CBE) .. 41
1.5.2.Редакторы оснований аденина (Adenine base editors – ABE) 44
1.5.3.Двойные редакторы оснований (Dual base editors) 45
1.5.4.Метод праймированного редактирования (Prime editing) 47
1.6.Заключение 51
ГЛАВА 2. МАТЕРИАЛЫ И МЕТОДЫ 52
2.1. Материалы 52
2.1.1.Коммерческие наборы 52
2.1.2.Питательные среды, сыворотки 52
2.1.3.Растворы, буферы и реагенты 53
2.1.4.Ферменты 54
2.1.5.Праймеры 54
2.1.6.Плазмиды 57
2.1.7.Клеточные культуры эукариот 59
2.1.8.Клеточные культуры прокариот 59
2.1.9.Программное обеспечение 59
2.2. Методы 59
2.2.1.ПЦР 59
2.2.2.Химический синтез олигонуклеотидов ДНК 60
2.2.3.Ферментативный гидролиз ДНК 61

2.2.4.Электрофоретическое фракционирование ДНК 61
2.2.5.Выделение фрагментов ДНК из агарозного геля 61
2.2.6.Очистка нуклеиновых кислот с помощью магнитных частиц .. 62
2.2.7.Клонирование фрагментов ДНК 62
2.2.8.Трансформация клеток *E. coli* методом электропорации 63
2.2.9.ПЦР-скрининг целевых клонов-трансформантов *E. coli* 63
2.2.10.Выделение плазмидной ДНК 64
2.2.11.Секвенирование ДНК 64
2.2.12.Хранение и культивирование линий клеток эукариот 65
2.2.13.Трансфекция в монослой эукариотических клеток 66
2.2.14.Световая микроскопия 66
2.2.15.Проточная цитофлуорометрия 66
2.2.16.Выделение геномной ДНК из эукариотических клеток 67
2.2.17.Выявление событий редактирования с использованием эндонуклеазы I бактериофага T7 67
2.2.18.Продукция VLP 68
2.2.19.Продукция лентивирусных частиц 69
2.2.20.Электронная микроскопия 69
2.2.21.Вестерн-блот 71
2.2.22.Определение трансдукционной способности клеток TZM-bl. 72
2.2.23.Обработка клеток TZM-bl вирусоподобными частицами 72
2.2.24.Секвенирование нового поколения (NGS) 73
2.2.25.Количественная ПЦР в реальном времени 75
2.2.26.Статистическая обработка 76
2.2.27.Моделирование пространственной структуры рибонуклеопротеиновых комплексов 76
ГЛАВА 3. РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ 78
3.1.Создание системы скрининга эффективности связывания химерных Cas-белков и действия редакторов оснований 78

3.2. Получение кодон-оптимизированных генов NLS-MmCas12m-NLS-TadDE и NLS-dSpCas9-NLS-TadDE 85
3.3. Подтверждение функциональной активности MmCas12m и dSpCas9 в комплексе с отдельными sgRNA, обеспечивающими таргетирование на область старт-кодона гена *gag* ВИЧ-1 88
3.4. Редактирование генома TZM-bl с помощью двойных редакторов оснований MmCas12m-TadDE и dSpCas9-TadDE 95
3.5. Создание плазмид для продукции вирусоподобных частиц VLP-MmCas12m-TadDE и VLP-dSpCas9-TadDE 105
3.6. Продукция, морфологическое описание и подтверждение функциональности вирусоподобных частиц VLP-MmCas12m-TadDE и VLP-dSpCas9-TadDE 109
3.7. Редактирование генома TZM-bl с использованием двойных редакторов оснований MmCas12m-TadDE и dSpCas9-TadDE, доставляемых вирусоподобными частицами 115
3.8. Повышение эффективности мутагенеза старт-кодона гена *gag* путём дизайна оптимальных sgRNA 118
3.9. Оценка безопасности применения вирусоподобных частиц VLP-MmCas12m-TadDE и VLP-dSpCas9-TadDE при мутагенезе старт-кодона *gag* в клетках линии TZM-bl 121
3.10. Моделирование пространственной структуры редакторов оснований MmCas12m-TadDE и dSpCas9-TadDE 122
ЗАКЛЮЧЕНИЕ 128
ВЫВОДЫ 130
ДАЛЬНЕЙШИЕ НАПРАВЛЕНИЯ ИССЛЕДОВАНИЙ 132
СПИСОК ЛИТЕРАТУРЫ 133
Приложение 1. Диаграммы рассеяния (дот-плоты), полученные при анализе методом проточной цитометрии клеток HEK293T, трансфицированных скрининговыми плазмидами phMGFP, pKW, pKW_R1, pKW_R3, pKW_R1m, pKW_R3m. 157

Приложение 2. Список олигонуклеотидов, использованных для сборки гена MmCas12m..... 158
Приложение 3. Список олигонуклеотидов, использованных для сборки рекомбинантного гена TadA-8e – TadDE 161
Приложение 4. Ампликоны, использованные для создания pMmCas12m-TadDE-sgRNA, psgRNA-MmCas12m и pdSpCas9-TadDE-sgRNA, psgRNA-dSpCas9 162
Приложение 5. Выравнивание sgRNA на область старт-кодона гена *gag*..... 163
Приложение 6. Применение системы скрининга для тестирования sgRNA1-9 при таргетировании MmCas12m и dSpCas9 на область старт-кодона гена *gag* провируса ВИЧ-1 164
Приложение 7. Электрофореграмма продуктов амплификации, кодирующих MmCas12m-TadDE и dSpCas9-TadDE 165
Приложение 8. Нуклеотидная последовательность, кодирующая двойной редактор оснований MmCas12m-TadDE..... 166
Приложение 9. Нуклеотидная последовательность, кодирующая двойной редактор оснований dSpCas9-TadDE 167
Приложение 10. Нуклеотидная последовательность, кодирующая полипептид-прекурсор Gag-MmCas12m-TadDE для продукции VLP-MmCas12m-TadDE 169
Приложение 11. Нуклеотидная последовательность, кодирующая полипептид-прекурсор Gag-dSpCas9-TadDE для продукции VLP-dSpCas9-TadDE..... 171
Приложение 12. Электрофореграмма продуктов амплификации, кодирующих полипептид-прекурсор Gag-MmCas12m-TadDE и Gag-dSpCas9-TadDE..... 174
Приложение 13. Необработанные результаты вестерн-блота с наличием маркера молекулярного веса белков 175

Приложение 14. Схема, отображающая принцип эксперимента для подтверждения функциональности VLP с использованием плазмид системы скрининга..... 176
Приложение 15. Диаграммы рассеяния (дот-плоты), полученные при анализе методом проточной цитометрии клеток HEK293T, трансфицированных скрининговой плазмидой pKW_R1 и обработанных VLP. 177
Приложение 16. Диаграммы рассеяния (дот-плоты), полученные при анализе методом проточной цитометрии клеток HEK293T, трансфицированных скрининговой плазмидой pKW_R3 и обработанных VLP. 178
Приложение 17. Двумерные тепловые карты взаимодействия доменов редакторов оснований MmCas12m-TadDE и dSpCas9-TadDE при редактировании старт-кодона *gag* провируса ВИЧ-1. 179

СПИСОК СОКРАЩЕНИЙ

AAV – (англ. adeno-associated virus) аденоассоциированный вирус;

ABE – (англ. adenine base editor) редактор основания аденина;

CA – (англ. capsid) капсид;

CBE – (англ. cytosine base editor) редактор оснований цитозина;

CCR5 – (англ. C-C chemokine receptor type 5) C-C рецептор хемокина 5-го типа;

CRF – (англ. circulating recombinant form) циркулирующая рекомбинантная форма ВИЧ;

CRISPR – (англ. clustered regularly interspaced short palindromic repeats) короткие палиндромные повторы, регулярно расположенные группами;

CXCR4 – (англ. C-X-C chemokine receptor type 4) C-X-C рецептор хемокина 4-го типа;

DPBS – фосфатно-солевой буфер Дульбекко;

env – (англ. envelope gene) ген белков оболочки;

FBS – (англ. fetal bovine serum) эмбриональная телячья сыворотка;

g – относительное центробежное ускорение;

gag – (англ. group antigen gene) ген внутренних белков вириона;

GFP – (англ. green fluorescent protein) зеленый флуоресцентный белок;

gp – (англ. glycoprotein) гликопротеин;

LTRs – (англ. long terminal repeats) длинные концевые повторы;

LV – (англ. lentivirus) лентивирус;

MA – (англ. matrix) матриксный белок;

NC – (англ. nucleocapsid) нуклеокапсид;

nef – (англ. negative regulator factor gene) ген негативного фактора экспрессии;

NGS – (англ. next-generation sequencing) секвенирование нового поколения;

NLS – (англ. nuclear localization signal) сигнал ядерной локализации;

PAM – (англ. protospacer adjacent motif) прилегающий мотив протоспейсера;

PBS – (англ. primer-binding site) сайт связывания праймера;

PBS-буфер – (англ. phosphate buffered saline) фосфатно-солевой буфер;

PE – (англ. prime editor) праймированный редактор;

pegRNA – (англ. prime editing guide RNA) направляющая РНК для праймированного редактирования;

pol – (англ. polymerase gene) ген полимеразы;

PSI – (англ. packaging signal HIV-1) сигнал упаковки ВИЧ-1;

rev – (англ. regulator of expression of virus protein genes) ген белка, регулирующего экспрессию генов вирусных белков;

sgRNA – (англ. single guide RNA) направляющая РНК;

TadDE – рекомбинантная дезаминаза TadA-8e;

TALEN – (англ. transcription activator-like effector nucleases) эффекторные нуклеазы, подобные активаторам транскрипции;

tat – (англ. transactivator gene) ген трансактивирующего фактора;

UGI – (англ. uracil DNA glycosylase inhibitor) ингибитор урацил-ДНК-гликозилазы;

vif – (англ. virus infectivity factor gene) ген вирусного инфекционного фактора;

VLP – (англ. virus-like particles) вирусоподобные частицы;

vpr – (англ. viral protein R gene) ген вирусного белка R;

vpu – (англ. viral protein U gene) ген вирусного белка U;

ВИЧ-1 – вирус иммунодефицита человека 1-го типа;

ОРС – открытая рамка считывания;

п.н. – пар нуклеотидов;

СПИД – синдром приобретённого иммунодефицита;

т.п.н. – тысяч пар нуклеотидов;

ТАЕ – (англ. tris-acetate-EDTA) трис-ацетатный ЭДТА буфер;

ВВЕДЕНИЕ

Актуальность темы исследования

В настоящее время геномное редактирование является стремительно развивающимся направлением молекулярной биологии. Инструменты геномного редактирования должны обладать приемлемой эффективностью, контролируемым уровнем нецелевых мутаций и отсутствием геномных перестроек. С открытием нуклеаз TALEN (Transcription Activator-Like Effector Nucleases) и нуклеаз на основе белков, содержащих домены «цинковые пальцы» (ZFN), начались активные работы по точечному редактированию генов и геномов [1]. Данные нуклеазы позволили приступить к избирательному внесению двухцепочечных разрывов в ДНК. Подобные разрывы позволяют инициировать мутации, сдвигать рамки считывания, удалять или вставлять фрагмент ДНК в геном живой клетки [2]. Однако для перепрограммирования нацеливания на конкретный участок возникает необходимость замены обширных аминокислотных последовательностей белков обеих групп TALEN и ZFN, а сама оптимизация при нацеливании на конкретный участок оказывается весьма трудоёмким процессом [3, 4]. Многообещающей альтернативой выступила открытая позже система геномного редактирования CRISPR-Cas. В классическом варианте комплекс CRISPR-Cas состоит из нуклеазы Cas и направляющей РНК (sgRNA). sgRNA комплементарна участку нацеливания на ДНК или РНК в зависимости от белка Cas [5]. При перепрограммировании на другие участки задача заключается в замене только небольшой по длине sgRNA. Это сделало систему CRISPR-Cas мобильнее по сравнению с ZFN и TALEN, где заменять надо целые ДНК-связывающие белковые домены. К тому же CRISPR-Cas система оказалась проще в отношении оптимизации нацеливания на участок редактирования [2]. В то же время принцип действия CRISPR-Cas по-прежнему основан на внесении двухцепочечных разрывов в ДНК, что не исключает нежелательные геномные перестройки, возникающие в ходе репарации генома клетки. Поиск безопасного инструмента геномного редактирования привёл к

разработке редакторов оснований ДНК. Принцип действия подобных редакторов заключается в точечном редактировании азотистого основания по механизму дезаминирования. При этом в ДНК клетки не вносятся двухцепочечные разрывы, что существенно снижает мутагенный эффект при редактировании [6]. Однако сами по себе редакторы оснований не могут нацеливаться на искомый участок редактирования. Для эффективного таргетирования редакторов оснований к месту редактирования применяют упомянутую CRISPR-Cas систему. Для предотвращения двухцепочечных разрывов чаще всего используют никазные формы Cas-белков (nCas) и реже – каталитически неактивные (dCas) [7]. В многочисленных исследованиях, описанных в обзоре литературы и в обсуждении, представлены результаты создания редакторов оснований с использованием различных Cas-белков, таких как SpCas9, SaCas9, SauriCas9, Cas12a и др. В составе редакторов оснований применяют различные дезаминазы и линкеры. Активно ведутся работы по модификации и направленной эволюции Cas-белков и дезаминаз для повышения точности, эффективности и придания новых свойств редакторам оснований. Особое место редакторы оснований получили в области генной терапии вирусных и наследственных заболеваний человека, как менее генотоксичный аналог классической CRISPR-Cas системы.

Таким образом, формируется актуальность поиска и модификации Cas-белков и дезаминаз для создания новых редакторов оснований с особыми свойствами. Среди недавно открытых Cas-белков особое внимание привлекает вариант MmCas12m, выделенный из микроорганизма *Mycolicibacterium mucogenicum* [8]. Данный вариант имеет ряд особенностей. MmCas12m состоит из 596 аминокислотных остатков (амк), что принципиально отличает его от широко применяемого SpCas9 (1368 амк) и от компактного аналога SaCas9 (1053 амк). Подобный сверхкомпактный размер позволяет закодировать MmCas12m в составе редактора оснований даже в одну экспрессионную кассету аденоассоциированного вируса (AAV), которому отдаётся предпочтение в качестве платформы доставки терапевтических инструментов при генной

терапии [9]. Даже компактный и широко применяемый аналог SaCas9 в составе редактора оснований не может быть закодирован в одну экспрессионную кассету AAV. К тому же компактные размеры облегчают доставку белковых комплексов в составе вирусоподобных частиц (VLP). В результате изучения биологических свойств MmCas12m установлено, что даже дикий тип не обладает каталитической активностью по отношению к целевой ДНК, но сохраняет способность к связыванию с ДНК, что объясняется его нативной функцией: подавлять транскрипцию экзогенных нуклеиновых кислот путём связывания с ДНК без внесения двухцепочечных разрывов. MmCas12m способен катализировать процессинг собственной sgRNA и узнавать 5’-TTN-3’ PAM (protospacer adjacent motif). Отсутствие каталитической активности по отношению к ДНК также позволяет клетке ингибировать экспрессию генов, экзогенных и эндогенных мобильных генетических элементов [8]. Как и другие Cas12-белки, MmCas12m имеет двухдоменную структуру, состоящую из REC (recognition — распознавание) и NUC (nuclease — нуклеаза) доменов. REC домен состоит из трёх частей: WED, REC1 и REC2, а NUC домен – из RuvC и TNB частей. Область распознавания PAM находится между REC1 и WED. Характерной особенностью каталитически активного нуклеазного домена RuvC Cas12-белков является наличие в активном центре трёх (реже четырех) отрицательно заряженных аминокислот, которые координируют два двухвалентных катиона магния или марганца (Mg^{2+}/Mn^{2+}). У всех групп Cas12-белков (за исключением двух вариантов), три остатка активного центра являются высококонсервативными: аспарагиновая кислота, глутаминовая кислота, аспарагиновая кислота (D-E-D). Два исключения – группа каталитически неактивных Cas12k, полностью утративших остатки активного центра, и группа Cas12m, в которой остатки, координирующие двухвалентные катионы, сильно варьируются. MmCas12m в данных позициях активного центра содержит гистидин и два остатка аспарагиновой кислоты (H-D-D), что исключает нуклеазную активность центра RuvC [8].

Среди текущего многообразия рекомбинантных дезаминаз стоит отметить вариант TadDE, полученный в ходе направленной эволюции исходной дезаминазы TadA-8e. TadDE проявляет способность к дезаминированию аденина и цитозина с относительно одинаковой эффективностью, что было установлено в ходе оригинального исследования [10]. Размеры TadDE составляют всего 167 амк, что делает её чрезвычайно компактной. Стандартно при создании двойных редакторов оснований объединяют домены аденозиндезаминаз и цитозиндезаминаз в единый эффекторный белок. Новая дезаминаза TadDE позволит создать двойные редакторы оснований с одним доменом дезаминазы, что уменьшит размеры и упростит доставку комплекса редактирования в клетки.

Принимая во внимание открытые MmCas12m, TadDE и их свойства, мы предположили, что новый двойной редактор оснований из MmCas12m и TadDE может быть крайне полезен для таргетного множественного мутагенеза. Компактные размеры позволят упростить доставку в клетки-мишени. К тому же, MmCas12m не обладает нуклеазной активностью, что минимизирует риск внесения разрывов в ДНК. Сама природа белков Cas12m исключает риск реверсии каталитической активности, который проявляется у точечных мутантов SpCas9, имеющих D10A или H840A. Таким образом, новый двойной редактор оснований MmCas12m-TadDE должен иметь сниженный риск нецелевой нуклеазной активности, а подход – направлен на минимизацию риска генотоксичности. Поскольку любое спланированное исследование должно включать контрольную группу для корректного сопоставления результатов, то мы параллельно создали наиболее близкий аналог dSpCas9-TadDE на основе каталитически неактивной формы широко применяемого SpCas9 (D10A/H840A).

Разработка любого инструмента геномного редактирования ориентирована на решение конкретных прикладных задач. Двойные редакторы оснований применяются в случаях, где необходим множественный мутагенез. В частности, они эффективны при необходимости нокаута генов для получения трансгенных клеточных линий или моделей лабораторных животных с

заданными свойствами для вирусологических исследований. На текущий день актуальным остается вопрос терапии ВИЧ-инфекции. Разработанные нами двойные редакторы оснований представляют значительный интерес для адаптации и применения в данной области. В рамках валидации эффективности систем в условиях, приближенных к *in cellulo*, целесообразно применение созданных нами двойных редакторов оснований для направленного мутагенеза старт-кодона гена *gag/gag-pol* провируса ВИЧ-1 в клетках линии TZM-bl. Ген *gag* кодирует основные структурные компоненты вириона и играет ключевую роль в его морфогенезе. Таким образом, инактивация старт-кодона *gag* (и других генов ВИЧ-1) интегрированного провируса ВИЧ-1 рассматривается как перспективная стратегия для разработки подходов к генной терапии ВИЧ-инфекции. Выбор двойных редакторов оснований обусловлен их способностью к множественному мутагенезу, который и требуется при необходимости инактивации гена.

В качестве платформы доставки редакторов оснований в клетки разработана система VLP на основе модифицированного лентивирусного вектора, исключающего упаковку в вирионы геномной РНК, а также обратной транскриптазы (RT) и интегразы (IN). Применение VLP подразумевает доставку в клетки функционально активных, предварительно собранных комплексов редактирования. Отсутствие в составе VLP вирусного генетического материала полностью нивелирует риск инсерционного мутагенеза, характерного для лентивирусных векторов. Разработка данной платформы согласуется с общей концепцией нашего исследования, направленной на минимизацию генотоксических эффектов.

Таким образом, обобщение теоретических предпосылок и накопленных к настоящему времени данных релевантных мировых исследований достаточно для обоснования актуальности разработки новых вариантов редакторов оснований ДНК. Особого внимания заслуживает тенденция к использованию вирусоподобных частиц в качестве универсальной и безопасной платформы для

внутриклеточной доставки различных систем прецизионного редактирования генома.

Цели и задачи работы

Целью диссертационной работы является создание и экспериментальная валидация новых двойных редакторов оснований на базе каталитически неактивных Cas-белков (MmCas12m, dSpCas9) и рекомбинантной двойной дезаминазы TadDE, а также разработка системы доставки данных редакторов в клетки на основе вирусоподобных частиц для направленного мутагенеза старт-кодон гена *gag* провируса ВИЧ-1.

Для достижения цели поставлены следующие задачи:

1. Выполнить биоинформатический дизайн, синтез *de novo* и направленное клонирование в экспрессирующих векторах кодон-оптимизированных химерных генов, кодирующих сконструированные двойные редакторы оснований MmCas12m-TadDE или dSpCas9-TadDE;

2. Разработать репортерную систему скрининга эффективности связывания химерных Cas-белков и действия редакторов оснований в целевой области редактирования петли SL4 сигнала упаковки PSI, в непосредственной близости от старт-кодона гена *gag* провируса ВИЧ-1;

3. Выявить особенности редактирования двойных редакторов оснований MmCas12m-TadDE и dSpCas9-TadDE в области старт-кодона гена *gag* в клетках перевиваемой линии TZM-bl, содержащих рекомбинантный фрагмент провируса ВИЧ-1;

4. Разработать и применить систему на основе вирусоподобных частиц для доставки двойных редакторов оснований MmCas12m-TadDE и dSpCas9-TadDE в эукариотические клетки *in vitro*;

5. С использованием вирусоподобных частиц, доставляющих двойные редакторы оснований MmCas12m-TadDE и dSpCas9-TadDE, осуществить мутагенез старт-кодона гена *gag* в клетках линии TZM-bl и оценить нецелевой эффект редактирования и персистенцию чужеродной ДНК.

Научная новизна

В рамках настоящей работы получены оригинальные конструкции репортерной системы скрининга для оценки эффективности связывания химерных Cas-белков и действия редакторов оснований в целевой области, включающей старт-кодон гена *gag* провируса ВИЧ-1. Данная система использована для анализа функциональности оригинальных направляющих РНК при таргетировании MmCas12m и dSpCas9 на область SL4 сигнала упаковки PSI провируса ВИЧ-1. Впервые сконструирован и синтезирован химерный ген NLS-MmCas12m-NLS-TadDE – двойной редактор оснований, состоящий из MmCas12m, рекомбинантной дезаминазы TadDE и сигналов ядерной локализации. Получен аналог NLS-dSpCas9-NLS-TadDE, в котором последовательность MmCas12m заменена на последовательность, кодирующую каталитически неактивную форму dSpCas9. Проведены дизайн и сборка генетических конструкций для продукции VLP на основе репликативно-дефектного лентивируса, предназначенных для *in vitro* доставки редактора оснований MmCas12m-TadDE или dSpCas9-TadDE. Проведено экспериментальное подтверждение способности оригинальных VLP проникать в клетки *in vitro* и доставлять функционально активный рибонуклеопротеиновый комплекс, состоящий из редактора оснований MmCas12m-TadDE или dSpCas9-TadDE и sgRNA. Функционально отобраны sgRNA, которые обеспечивают нацеливание редакторов оснований на старт-кодон гена *gag* провируса ВИЧ-1. При использовании девяти основных и двух дополнительных sgRNA установлено, что MmCas12m-TadDE обеспечивает эффективность редактирования до 10–11 % в позиции некоторых оснований, dSpCas9-TadDE – до 40 %. При этом MmCas12m-TadDE проявил себя как более стабильный

вариант в условиях неоптимального PAM, что позволяет использовать его в тех случаях, где применение dSpCas9-TadDE невозможно или сопряжено с низкой эффективностью. Нами впервые описана и опубликована стратегия генной терапии ВИЧ-инфекции с использованием двойных редакторов оснований, доставляемых VLP, для мутагенеза старт-кодона гена *gag*. Представленные результаты научно доказуемы и впервые получены в ходе проведения диссертационной работы.

Теоретическая и практическая значимость работы

Результаты, полученные в данной работе, дополняют имеющиеся знания и экспериментальный опыт в области разработки инструментов геномного редактирования и генной терапии ВИЧ-инфекции.

В рамках создания системы скрининга получены генетические конструкции серии pKW_R, позволяющие проводить анализ связывания специфичных к ДНК Cas-белков с областью старт-кодона гена *gag* провируса ВИЧ-1. Данные конструкции также могут быть использованы для выявления событий редактирования, результат которого будет отражаться на изменении уровня продукции белка-репортера. При необходимости исследователь может заменить имеющуюся нуклеотидную последовательность фрагмента провируса ВИЧ-1 на другую, желаемую для таргетирования. Для этой цели в генетические конструкции заложены сайты эндонуклеаз рестрикции *BmtI*, *EcoRV* для клонирования последовательности для таргетирования. Это делает генетические конструкции системы скрининга универсальными.

Экспериментально выявлены эффективные sgRNA, обеспечивающие таргетирование редакторов оснований MmCas12m-TadDE и dSpCas9-TadDE на область старт-кодона гена *gag* провируса ВИЧ-1. Представленные в работе sgRNA могут использоваться для таргетирования инструментов геномного

редактирования или иных белковых компонентов к данной области с использованием MmCas12m и dSpCas9.

Два набора полученных генетических конструкций (pGag-MmCas12m-TadDE, psgRNA-MmCas12m и pGag-dSpCas9-TadDE, psgRNA-dSpCas9) могут быть использованы для продукции вирусоподобных частиц, обеспечивающих доставку двойного редактора оснований MmCas12m-TadDE или dSpCas9-TadDE в эукариотические клетки *in vitro*. При перепрограммировании исследователю необходимо произвести лишь одноэтапное клонирование двухцепочечного фрагмента ДНК, кодирующего искомую sgRNA, в плазмиду psgRNA-MmCas12m или psgRNA-dSpCas9. Стоит отметить, что редакторы оснований, доставляемые в синтезированном виде с помощью VLP, сохраняют эффективность и ширину окна редактирования, как и в случае прямой трансфекции плазмидами.

Разработанные двойные редакторы оснований MmCas12m-TadDE и dSpCas9-TadDE могут использоваться для различных целей в области генной терапии наследственных и инфекционных заболеваний человека, для трансгенеза лабораторных животных и перевиваемых эукариотических клеточных линий. Компактные размеры химерного гена NLS-MmCas12m-NLS-TadDE также позволяют закодировать полный редактор оснований в единую экспрессионную кассету AAV, размер которой не может превышать ~4700 нт. Это принципиально отличает данный редактор от существующих двойных редакторов на основе SpCas9 или SaCas9, для которых требуется двухвекторная стратегия. Разработанная платформа доставки на основе VLP обеспечивает доставку уже синтезированных комплексов редактирования и исключает попадание в клетку чужеродных нуклеиновых кислот, кодирующих чужеродные белковые компоненты. Таким образом, подход с применением каталитически неактивных Cas-белков и VLP предполагает минимизацию инсерционного мутагенеза и исключает длительную эписомальную транскрипцию, что обеспечивает снижение генотоксичности.

Методология и методы исследования

Для достижения поставленной цели и выполнения задач применяли общенаучные и специальные методы исследования. В теоретической главе диссертационной работы при написании обзора литературы, построения гипотезы, дизайна модельной системы использовали общелогические методы, гипотетический метод, формальные методы логики и системный анализ. Для получения практических результатов использовали методы молекулярной биологии, генетической инженерии, вирусологии, биоинформатики и математической статистики.

Положения, выносимые на защиту

1. Последовательности кодон-оптимизированных генов сконструированных двойных редакторов оснований обеспечивают синтез функционально активных продуктов MmCas12m-TadDE и dSpCas9-TadDE *in vitro*;

2. Репортерная система скрининга пригодна для количественной оценки взаимодействия ДНК-специфичных Cas-белков (MmCas12m и SpCas9) с областью старт-кодона гена *gag* провируса ВИЧ-1;

3. Сконструированные плазмиды pGag-MmCas12m-TadDE, psgRNA-MmCas12m и pGag-dSpCas9-TadDE, psgRNA-dSpCas9 могут быть использованы для продукции вирусоподобных частиц VLP-MmCas12m-TadDE и VLP-dSpCas9-TadDE соответственно, которые способны к самосборке с дальнейшим созреванием и доставке в клетки *in vitro* функционально активных рибонуклеопротеиновых комплексов.

Степень достоверности и апробации результатов

Положения и выводы являются научно обоснованными. Достоверность результатов подтверждена статистической обработкой и воспроизводимостью в повторах. Результаты всех экспериментов получены с применением современных молекулярно-биологических, вирусологических и генно-инженерных методов исследования.

Результаты были представлены на III Гомельском международном конгрессе «Инфекционные болезни, микробиология и иммунология» (г. Гомель, Республика Беларусь), на Международной научной конференции молодых учёных в области биотехнологии, молекулярной биологии, вирусологии, биофизики, биоинформатики и фундаментальной медицины «OpenBio 2025» (р.п. Кольцово, Российская Федерация), на Главной международной научно-практической конференции по молекулярной диагностике в России «Молекулярная Диагностика – 2025» (г. Москва, Российская Федерация), на Международной научной студенческой конференции МНСК-2026 (г. Новосибирск, Российская Федерация), на 15-ой Международной мультиконференции «Биоинформатика регуляции и структуры геномов/системная биология» – BGRS/SB-2026 (г. Новосибирск, Российская Федерация).

Список публикаций по теме диссертации

По материалам диссертационной работы опубликовано 10 научных работ, в том числе 4 научные статьи в рецензируемых журналах, рекомендованных ВАК Министерства образования и науки РФ для защиты диссертаций, и 5 тезисов в рецензируемых сборниках трудов международных научных конференций. Получен 1 патент РФ.

Структура и объём диссертации

Диссертационная работа состоит из таких разделов, как введение, обзор литературы, материалы и методы, результаты и их обсуждение, заключение, выводы, список цитируемой литературы и приложение. Работа изложена на 179 страницах, включает 27 рисунков, 7 таблиц. Список литературы включает 185 источников. Приложение включает 17 дополнительных материалов.

Личный вклад автора

Автор принимал непосредственное участие в постановке задач, решаемых в рамках диссертационной работы, все представленные экспериментальные данные получены лично или при участии автора. Автор принимал непосредственное участие в статистическом анализе, интерпретации данных, представлении результатов на конференциях, а также написании и публикации научных статей в рецензируемых журналах. Перевод результатов на английский язык при написании научных статей выполнен автором лично.

Благодарности

Автор выражает благодарность Ильназу Рамисовичу Иматдинову, своему научному руководителю, за фундаментальный вклад в создание данной работы. Автор благодарит коллектив лаборатории векторных систем на основе вирусных геномов ГНЦ ВБ «Вектор» Роспотребнадзора.

Автор искренне признателен и благодарен каждому, кто помог делом и добрым словом.

ГЛАВА 1. ОБЗОР ЛИТЕРАТУРЫ

1.1 ВИЧ-инфекция

Синдром приобретённого иммунодефицита (СПИД), вызываемый вирусом иммунодефицита человека 1-го типа (ВИЧ-1), является серьёзной угрозой для общественного здравоохранения. По последним оценкам Всемирной организации здравоохранения во всем мире проживает не менее 44,5 млн инфицированных ВИЧ-1. Каждый год в мире умирает около 800 тысяч человек из-за последствий ВИЧ-инфекции [11]. Согласно Государственному докладу «О состоянии санитарно-эпидемиологического благополучия населения в Российской Федерации в 2025 году», представленному в начале июня 2026 года, число лабораторно подтверждённых носителей ВИЧ-инфекции в Российской Федерации достигло значения не менее 1,25 млн человек [12]. Несмотря на то, что заболеваемость ВИЧ-инфекцией в 2025 году оказалась рекордно низкой за последние 12 лет (45638 новых случаев), а охват населения по проведённым тестам – 53,8 млн, ситуация по ВИЧ-инфекции требует пристального внимания (рис. 1).

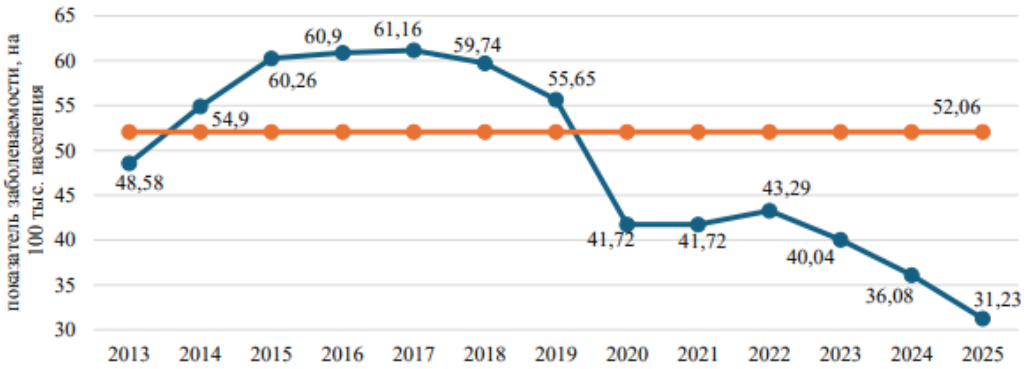


Рисунок 1. Динамика заболеваемости ВИЧ-инфекцией в Российской Федерации в 2013–2025 гг. (на 100 тыс. населения). Оранжевым выделена линия среднемноголетнего показателя, синим – ежегодного. Взято из Государственного доклада «О состоянии санитарно-эпидемиологического благополучия населения в Российской Федерации в 2025 году» [12].

ВИЧ-инфекция вызывает прямые и косвенные экономические затраты. Общий объём прямых медицинских расходов федерального бюджета и бюджетов субъектов Российской Федерации на реализацию мер по противодействию распространения ВИЧ-инфекции в 2025 г. составил 83,1 млрд руб. Косвенные экономические затраты, обозначенные как потери ВВП от смертности и инвалидизации населения в связи с ВИЧ-инфекцией, составили 181 млрд руб., выплаты пособий по инвалидности в связи с ВИЧ-инфекцией – 15,2 млрд руб. Суммарно экономическое бремя ВИЧ-инфекции только за период 2025 г. в Российской Федерации можно оценить в 279,2 млрд руб. За 2024 г. данный показатель был немного ниже – 276,1 млрд руб.

Всё чаще новые случаи заражения ВИЧ-1 определяются среди наиболее трудоспособного населения. Так, в 2025 г. 34,8 % всех случаев впервые выявленной ВИЧ-инфекции приходились на долю граждан возраста 40–49 лет и 32,2 % – на долю возраста 30–39 лет. Больше половины зарегистрированных случаев составили мужчины – 60,3 %. В результате опроса среди впервые инфицированных в 2025 г. граждан установлено, что 82,3 % сообщили только о половом традиционном контакте, 5,4 % – о нетрадиционном, и 11,1 % – о внутривенном употреблении наркотиков [12]. Таким образом, основным путём передачи на сегодняшний день остается половой.

Наиболее пораженными регионами с тяжёлой эпидемиологической обстановкой остаются: Кемеровская (2039,25 поражённых ВИЧ-инфекцией на 100 тыс. населения), Свердловская (2063,25), Иркутская (1941,05), Оренбургская (1743,36) и Самарская (1474,54) области. На территории Новосибирской области данный показатель составляет 1429,98. Согласно результатам молекулярно-генетических исследований, проводимых на базе научных организаций Роспотребнадзора, на территории Российской Федерации преобладает субтип А6 ВИЧ-1 (83,8 %), существенно реже выявляется субтип В (6,1 %) и некоторые циркулирующие рекомбинантные формы CRF63_02A6 (3,1 %), CRF02_AG (2,3 %), рекомбинантные формы между всеми указанными вариантами и субтип G (0,9 %). В редких случаях удаётся обнаружить нетипичные для Российской

Федерации генетические варианты ВИЧ-1: A7 и CRF06_cpx, характерные для Западной Африки [13, 14], CRF18_cpx, характерный для Кубы и Центральной Африки [15].

1.2 Вирус иммунодефицита человека 1-го типа

ВИЧ-1 является представителем семейства *Retroviridae* рода *Lentivirus*. Зрелый вирион ВИЧ-1 имеет сферическую форму диаметром около 100-120 нанометров [16]. В состав зрелого вириона входит несколько тысяч белковых молекул различных типов, каждая из которых выполняет собственную функцию для обеспечения полного жизненного цикла (рис. 2А). ВИЧ-1 является оболочечным вирусом, поэтому каждый вирион покрыт липидным бислоем, который образуется из плазматической мембраны клетки-продуцента, когда вирион отпочковывается от её поверхности [17]. Вирусную оболочку пронизывают до 100 поверхностных гликопротеиновых комплексов, каждый из которых образован тремя молекулами поверхностного гликопротеина gp120 (SU) и тремя молекулами трансмембранного гликопротеина gp41 (TM) [18]. Впрочем, поздние исследования указывают на то, что количество данных тримерных комплексов значительно меньше и варьируется в пределах 7–14 [19]. С помощью гликопротеина gp120 ВИЧ-1 взаимодействует с рецептором CD4 и корецепторами CCR5 или CXCR4, представленными на поверхности Т-хелперов, моноцитов, макрофагов, дендритных клеток и NK-клеток человека [20, 21]. Трансмембранный гликопротеин gp41 играет важную роль в закреплении всего гликопротеинового комплекса на поверхности вирусной оболочки [22]. Под вирусной оболочкой располагается матрикс, структурной субъединицей которого является белок p17 (MA). На N-конце MA располагается сигнал модификации миристиновой кислотой и высокоосновная область (позиции 17–31 амк от N-конца), что обеспечивает локализацию данного домена и всех незрелых молекул Gag/p55 и Gag-Pol/p160 на поверхности мембраны [23, 24]. Три молекулы MA образуют тример, который организуется с другими тримерами в более сложную гексагональную архитектуру, покрывающую

пространство под вирусной оболочкой зрелого вириона [25]. Под матриксом располагается капсид, образованный из молекул белка p24 (CA). Именно белок p24 является диагностически значимым и служит для выявления ВИЧ-инфекции на ранних стадиях заражения [26]. Зрелый капсид имеет вытянутую форму, напоминающую усеченный конус, и состоит из около 1500 молекул белка p24, собранных в гексамерную структуру.

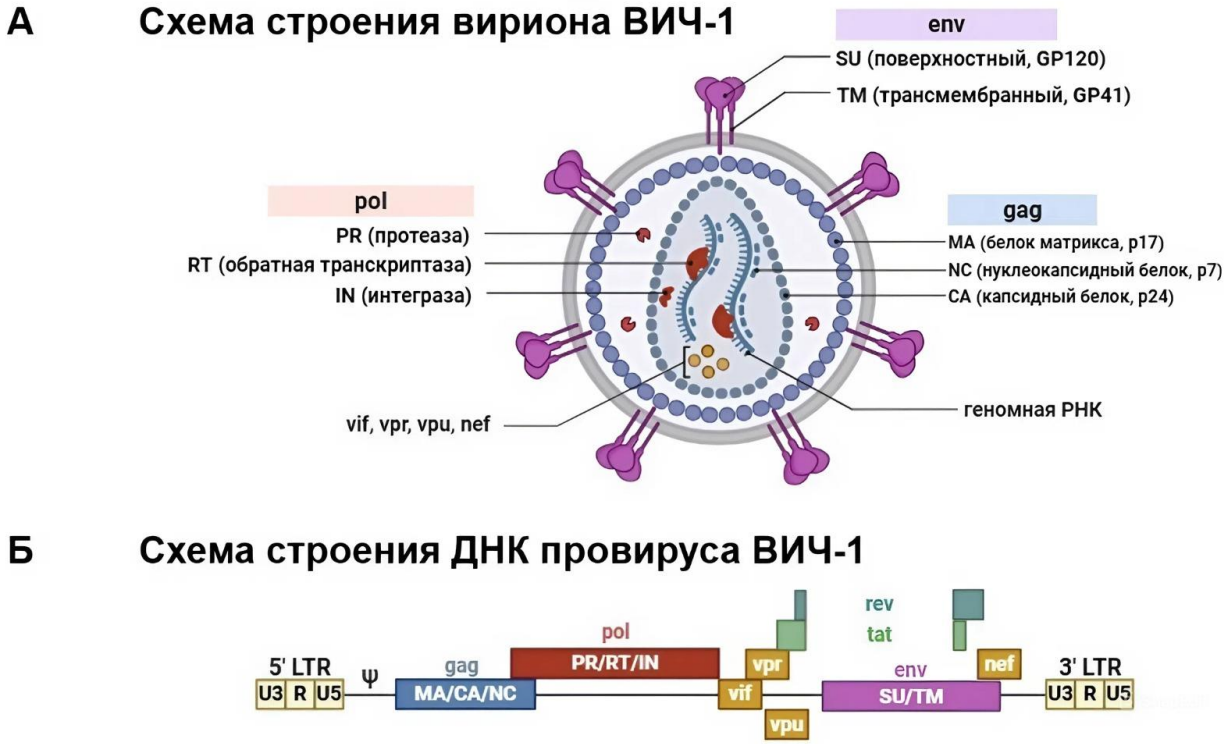


Рисунок 2. Схематичное изображение строения вириона (А) и ДНК провируса ВИЧ-1 (Б) с указанием основных структурных элементов. Подробное описание в тексте. Сделано с помощью BioRender.

Внутри капсида находятся две копии полноразмерной молекулы геномной РНК длиной около 9750 нуклеотидов [27]. Каждая молекула геномной РНК состоит из регуляторных элементов R и U5 с 5’-конца, U3 и R с 3’-конца, между которыми закодированы девять генов *gag*, *pol*, *vif*, *vpr*, *vpu*, *env*, *tat*, *rev* и *nef* [28]. В форме провирусной ДНК регуляторные элементы образуют длинные концевые повторы (long terminal repeats – LTRs) (рис. 2Б). Молекулы геномной РНК связаны с белком нуклеокапсида p7 (NC). NC имеет в своем строении домены

«цинковых пальцев», чем и обеспечивает захват геномной РНК внутрь будущего капсида во время сборки незрелого вириона на плазматической мембране клетки [29]. NC окружен двумя спейсерными пептидами SP1 и SP2, которые играют роль в созревании вириона, поскольку содержат сайты гидролиза вирусной протеазы [30]. Домен p6, локализованный на С-конце молекул Gag/p55, принимает участие в захвате белка Vpr внутрь вириона [31]. Внутри капсида также располагаются вирусные ферменты: интеграза, обратная транскриптаза, в то время как протеаза локализуется и вне капсида [32-35]. В капсиде также выявлен вспомогательный белок Nef, который играет роль в подавлении синтеза белков главного комплекса гистосовместимости MHC-I, и белок Vif, задача которого заключается в предотвращении упаковки эндогенной дезаминазы APOBEC3G в вирион [36, 37]. Вирусный вспомогательный белок Vpu не упаковывается в вирион. Он обнаруживается на плазматической мембране и эндоплазматическом ретикулуме инфицированной клетки [38]. Его основная задача заключается в блокировании клеточного белка BST-2, который препятствует отщеплению зарождающегося вириона от плазматической мембраны клетки [39, 40]. Несмотря на то, что целенаправленная упаковка Vpu в вирион не происходит, его локализация на мембране не исключает данное событие.

Проникновение ВИЧ-1 в клетку-мишень осуществляется путём взаимодействия поверхностного вирусного белка gp120 с CD4-рецептором клетки. После чего вирус взаимодействует с хемокиновым корецептором CCR5 или CXCR4 (рис. 3, 1). Происходит конформационное изменение поверхностного гликопротеина вируса, в результате которого вирион сближается с мембраной клетки, вызывая дальнейшее слияние своей оболочки и мембраны клетки [41].

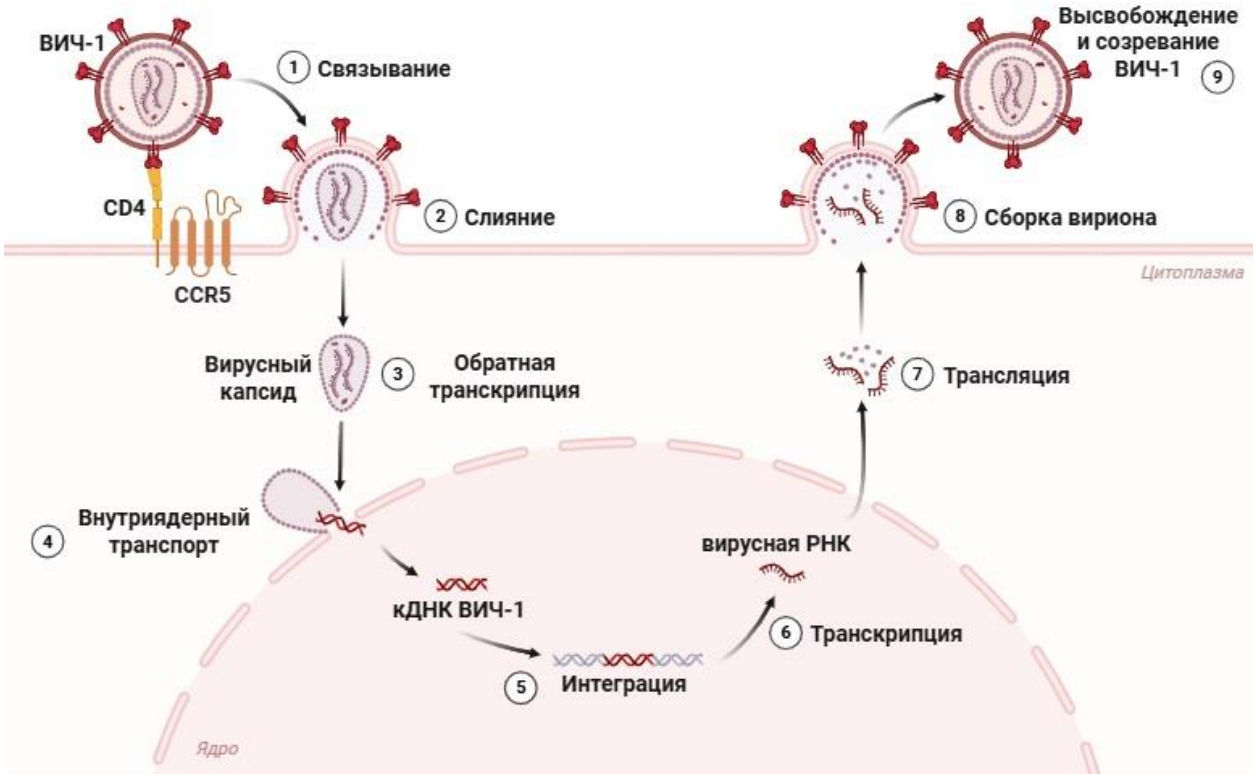


Рисунок 3. Схематичное изображение жизненного цикла ВИЧ-1 с указанием основных стадий. Подробное описание в тексте. Сделано с помощью BioRender.

После проникновения вируса в клетку начинается постепенное разрушение капсида, сопряженное с обратной транскрипцией (рис. 3, 2). Синтез кДНК обеспечивает вирусная обратная транскриптаза, которая обладает активностью РНК-зависимой ДНК-полимеразы, РНКазы H и ДНК-зависимой ДНК-полимеразы. Обратная транскрипция происходит во время транспорта капсида от клеточной мембраны к ядерной (рис. 3, 3). Поскольку обратная транскриптаза не способна начать синтез *de novo*, то в качестве затравки она использует тРНК-Lys3 клетки-хозяина, 3’-конец которой комплементарен сайту связывания праймера (primer-binding site – PBS) ВИЧ-1 [42]. По мере синтеза обратная транскриптаза вируса гидролизует цепь РНК в химерной молекуле ДНК-РНК, проявляя свойства РНКазы H [43, 44]. В результате образуется двухцепочечная молекула вирусной ДНК. Около ядерных пор происходит высвобождение вирусного генома с последующим проникновением в ядро.

Таким образом, процесс обратной транскрипции происходит под защитой вирусного капсида, что предотвращает деградацию нуклеиновых кислот ВИЧ-1 клеточными нуклеазами.

Синтезированная вирусная ДНК продолжает взаимодействие с вирусными и клеточными белками, образуя нуклеопротеиновый комплекс предварительной интеграции (PIC), в состав которого входят Vpr, MA, вирусная интеграза, некоторые белки клетки-хозяина, в частности фактор BANF1, фактор роста LEDGF/p75, белки группы высокой подвижности (HMGs), интегразный взаимодействующий фактор 1 (INI1) [45–48]. Вирусная интеграза образует гомодимеры, каждый из которых взаимодействует с одним из LTRs двухцепочечной вирусной ДНК в комплексе PIC. Интеграза производит гидролиз двухцепочечной молекулы в сайте узнавания интегразы, удаляя динуклеотид GT с каждого 3’-конца вирусной ДНК, образуя 5’-выступ (5′-overhang). На укороченном 3’-конце остается консервативный CA динуклеотид [49, 50]. В таком виде PIC проникает в ядро через ядерные поры, используя сеть микротрубочек и активных филаментов для направленного движения к ядерной мембране (рис. 3, 4). Процесс непосредственного проникновения через ядерные поры опосредуется белком импортином клетки-хозяина [51]. Особенностью ВИЧ-1 относительно многих других ретровирусов является его способность к инфекции как делящихся, так и неделящихся клеток. Главным образом это связано с наличием белка Vpr, который опосредует внутриклеточный транспорт в ядро и предотвращает пролиферацию инфицированных клеток [52, 53]. 3’-конец каждого LTR вирусной ДНК образует новые фосфодиэфирные связи с ДНК клетки-хозяина, что обусловлено активностью вирусной интегразы (рис. 3, 5). Стоит отметить, что выбор позиции для интеграции вирусной ДНК в геном клетки-хозяина имеет неслучайный характер. На места рекомбинации ДНК влияет структура хроматина. Замечена тенденция к интеграции в промоторные и энхансерные области, GC-богатые участки, участки с эпигенетическими метками триметилированного гистона 3-лизина 36 [54]. Также ключевую роль играет и пространственная конформация

хроматина. События интеграции чаще регистрировались на периферии хроматина ближе к ядерной мембране. Даже активный хроматин в центре ядра имел частоту интеграции ниже в сравнении с таковым на периферии [55]. После успешной интеграции копия провирусной ДНК ВИЧ-1 может реплицироваться вместе с ДНК клетки-хозяина, транскрибироваться или сохраняться в латентном состоянии [56].

В процессе транскрипции задействованы ферменты матричного синтеза клетки-хозяина (рис. 3, 6). Непосредственный синтез полноразмерной вирусной РНК обеспечивает основной фермент транскрипции эукариот – РНК-полимераза II [57]. Важную роль в инициации транскрипции играет вирусный белок Tat. Белок Tat связывается с TAR-элементом (Trans-activation response element), который находится на 5'-концах зарождающихся транскриптов ВИЧ-1. После связывания с TAR-элементом Tat привлекает положительный комплекс элонгации транскрипции – P-TEFb, который фосфорилирует С-конец РНК-полимеразы II, снимая её с паузы транскрипции. Таким образом Tat предотвращает транскрипцию абортивных коротких молекул вирусной РНК и увеличивает синтез полноразмерной вирусной РНК [58, 59]. В отсутствие Rev белка происходит сплайсинг полноразмерного транскрипта и его ядерный экспорт белками клетки-хозяина. Один из сплайсированных вариантов РНК вируса обеспечивает трансляцию белка Rev. Синтезированные молекулы Rev белка проникают в ядро благодаря сигналу ядерной локализации в своём строении. Внутриклеточный транспорт белка Rev в ядро обеспечивает белок импортин-β [60]. Связываясь в ядре с RRE-элементом (Rev response element) полноразмерной молекулы вирусной РНК, Rev меняет конформацию таким образом, что его сигнал ядерной локализации становится недоступным, а сигнал ядерного экспорта выносится на поверхность и становится доступным для транспортных белков клетки-хозяина, в числе которых экспортин XPO1 [61]. После экспорта из ядра несплайсированная полноразмерная вирусная РНК не только упаковывается в вирион, но и служит матрицей для синтеза структурных белков Gag/p55, Gag-Pol/p160. По матрице сплайсированных РНК происходит

трансляция остальных вирусных белков: Tat, Vpr, Vif, Vpu, Nef и белка-предшественника Env (рис. 3, 7). Сборка вириона происходит на внутренней поверхности плазматической мембраны клетки-хозяина. Основу будущего вириона образуют белки-предшественники Gag/p55, Gag-Pol/p160 [62]. Синтез белка Gag-Pol/p160 происходит в результате сдвига рамки считывания при движении рибосомы. Соотношение Gag/p55:Gag-Pol/p160 варьируется в пределах от 10:1 до 20:1. Данные белки локализуются преимущественно в местах сосредоточения фосфатидилинозитол-4,5-бисфосфата, электростатическое взаимодействие с которым обусловлено присутствием высокоосновной области и миристиновой кислоты на N-конце MA домена белков Gag/p55 и Gag-Pol/p160 [63]. Межмолекулярные взаимодействия доменов «цинковых пальцев» Gag/p55 и Gag-Pol/p160 с вирусной РНК происходят только с полноразмерным транскриптом, поскольку именно он содержит последовательность сигнала упаковки, который играет ключевую роль для захвата РНК в созревающий вирион [64]. Сосредоточение молекул белков Gag/p55 и Gag-Pol/p160 на плазматической мембране также способствует локализации тримеров Env в местах сборки. Обеспечивается взаимодействием между MA и C-хвостом гликопротеина Env [65]. Скопление и мультимеризация Gag/p55 и Gag-Pol/p160 приводит к физической деформации мембраны (рис. 3, 8). В результате изгиба начинает образовываться сферическая форма зарождающегося вириона. Полное высвобождение вириона от поверхности плазматической мембраны опосредуется активностью клеточных эндосомальных сортировочных комплексов, необходимых для транспортного механизма – ESCRT [66]. Окончательное созревание вирусной частицы происходит уже после отщепления (рис. 3, 9). Только зрелый вирион после структурной реорганизации является инфекционным. Созревание начинается с автокаталитической активации вирусной протеазы. После отщепления протеазы от предшественника Gag-Pol/p160 происходит последовательное расщепление белков-предшественников Gag/p55 и Gag-Pol/p160 по строго расположенным сайтам гидролиза [67]. В результате ферментативной обработки высвобождаются вирусные ферменты и

структурные белки (MA/p17, CA/p24 и NC/p7). Белки начинают собираться в характерные структуры, описанные выше. Зрелый вирион ВИЧ-1 готов к новому циклу репликации.

1.3 Генная терапия ВИЧ-инфекции

В настоящее время не разработано эффективной вакцины для профилактики или препарата для полного излечения от ВИЧ-инфекции. Существующая комбинированная антиретровирусная терапия (АРТ) является важным противоэпидемическим мероприятием, способным подавлять репликацию вируса, но только до тех пор, пока пациент принимает соответствующие противовирусные препараты [68]. Подобные мероприятия снижают вероятность передачи вируса от инфицированных людей. В 2025 году снижение контагиозности (подавлена вирусная нагрузка) было установлено у 75,5 % больных, проходящих курс АРТ [12]. Тем не менее, данный показатель не является достаточным для радикального снижения частоты передачи ВИЧ-1 и смертности от ВИЧ-инфекции и сопутствующих заболеваний. Невозможность полного излечения от ВИЧ-инфекции с помощью АРТ обусловлена сохраняющейся в геноме инфицированного провирусной ДНК на протяжении всей жизни и её способностью долгое время пребывать в латентном состоянии.

Осложняет положение и возникновение устойчивости ВИЧ-1 к антиретровирусным препаратам (АРВП), что обусловлено высоким уровнем генетической изменчивости самого вируса [69]. Так, например, частота ошибок обратной транскриптазы при синтезе двухцепочечной молекулы ДНК вируса составляет 10^{-5}-10^{-3} ошибок/п.н., это значит, что каждая провирусная ДНК ВИЧ-1 длиной около 9,7 тысяч п.н. с высокой вероятностью будет нести минимум одну мутацию *de novo* [70]. И это не единственный источник генетической изменчивости. Нужно учитывать, что мутации могут возникнуть на этапе обратной транскрипции полноразмерной вирусной РНК в результате рекомбинации между двумя различными молекулами геномной РНК, а также в

результате воздействия белков эндогенных дезаминаз семейства APOBEC [71–73].

В связи с этим актуальным является развитие подходов генной терапии ВИЧ-инфекции. Одни из первых подходов предполагали прямое воздействие на ДНК провируса ВИЧ-1 с помощью CRISPR-Cas систем. Данная стратегия может заключаться как в полном удалении провируса, так и во внесении мутаций в открытые рамки считывания генов вирусных белков. Целью редактирования становятся нуклеотидные последовательности вирусных генов, играющих ключевую роль для полного жизненного цикла вируса. Так, предпринимались попытки внесения мутаций в гены *rev, tat, env, gag, pol* ВИЧ-1 с помощью опосредованных CRISPR-Cas9 двухцепочечных разрывов в соответствующую часть провирусной ДНК. Подобное редактирование приводило к снижению вирусной нагрузки вследствие синтеза с отредактированного провируса ВИЧ-1 мутантных транскриптов, по матрице которых проходил синтез нефункциональных вирусных белков [74–76]. Первым препаратом для генной терапии ВИЧ-инфекции, проходящий клинические испытания, стал EBT-101. Это экспериментальный препарат, разработанный американской компанией Excision BioTherapeutics. Препарат предназначен для полного удаления провирусной ДНК ВИЧ-1 из генома инфицированной клетки по принципу «функционального излечения». Основой препарата послужила CRISPR-Cas9 система, которая нацеливается на консервативные участки LTRs провируса, вносит двухцепочечные разрывы и тем самым удаляет провирус из генома. Доставка осуществляется с использованием вектора на основе AAV 9-го серотипа. Несмотря на предварительную безопасность, препарат не смог полностью вылечить участников испытаний. После прекращения антиретровирусной терапии спустя 1-4 месяца у пациентов снова наблюдали повышение вируса в крови [77].

Несмотря на кажущуюся простоту и очевидные преимущества, подобный подход с инактивацией одного или нескольких генов всё же имеет некоторые ограничения. Мутации, возникающие в отдельных генах после репарации по

механизму негомологичного соединения концов, могут генерировать жизнеспособные формы провируса ВИЧ-1, которые к тому же становятся устойчивы к повторному воздействию CRISPR-Cas с конкретной направляющей РНК [78]. Помимо этого, мутации в одном гене не исключают синтез других вирусных белков, которые могут продолжать оказывать на клетку токсичный эффект [79, 80]. Чтобы избежать подобного эффекта, применялся подход с полным удалением провируса ВИЧ-1 из генома клетки-хозяина с использованием CRISPR-Cas9 системы, нацеленной на LTRs провируса [81]. Однако данный подход требует учитывать, что частота делеций ниже частоты мутаций. При этом вероятность делеции уменьшается с увеличением размера удаляемого фрагмента [82]. Таким образом вместо искомой делеции провируса может возникнуть две мутации в каждом LTR, которые не решают проблему окончательно. Дополнительно выявлено, что после подобного удаления провирусы ВИЧ-1 долгое время способны сохраняться в клетке в эписомальной форме, обеспечивая экспрессию вирусных белков [83]. Учитывая, что в каждом геном инфицированной клетки может быть несколько копий провирусной ДНК, то самым опасным фактором может стать неправильная репарация множественных двухцепочечных разрывов, способная привести к хромосомным перестройкам с летальным исходом.

Ещё одним подходом генной терапии ВИЧ-инфекции стало создание устойчивых к вирусу клеток путём редактирования генов корецепторов CCR5 и CXCR4. Сразу несколько научных групп сообщили об эффективной инактивации гена *CCR5* с использованием CRISPR-Cas9, что придало 25–30 % отредактированным CD4+ лимфоцитам и CD34+ гемопоэтическим стволовым клеткам устойчивость к R5-тропному ВИЧ-1 [84, 85]. В 2019 году было сообщено об успешной аллогенной трансплантации модифицированных *ex vivo* с помощью CRISPR-Cas9 гемопоэтических стволовых клеток с нокаутом гена *CCR5* пациенту с ВИЧ-инфекцией [86]. Однако, данная терапия оказалась неэффективна, поскольку из-за низкой доли модифицированных CD4+ лимфоцитов (около 5–8 %) не удалось остановить распространение ВИЧ-1.

Впрочем, инактивация одного лишь *CCR5* не решает задачу блокирования распространения вируса в организме уже инфицированного пациента, поскольку по мере прогрессирования заболевания ВИЧ-1 использует для проникновения ещё один хемокиновый рецептор – CXCR4. Следовательно, для создания резистентности к ВИЧ-1 независимо от его тропности необходимо инактивировать гены обоих корецепторов: CXCR4 и CCR5. Однако CXCR4 играют ключевую роль в развитии гемопоэтических стволовых клеток и их дальнейшего дифференцирования в Т-лимфоциты [87]. Это ограничивает терапевтический потенциал нокаута гена *CXCR4*. Дополнительно стоит упомянуть, что описанный ранее риск нежелательных эффектов от двухцепочечных разрывов в ДНК по-прежнему остается. Подобные геномные перестройки могут нарушить экспрессию генов-супрессоров опухолевого роста. Следствием подобной терапии могут стать злокачественные образования [88].

Привлекательной альтернативой является более деликатное воздействие на геном клетки с помощью редакторов оснований ДНК. Подобные ферменты точечно без внесения двухцепочечных разрывов в ДНК способны модифицировать нуклеотиды по механизму дезаминирования азотистых оснований. Однако редактор оснований ДНК не может самостоятельно связываться с искомым окном редактирования, по этой причине для таргетирования применяют CRISPR-Cas системы. Cas-белок обеспечивает таргетирование на целевой участок с помощью sgRNA, при этом для предотвращения нуклеазной активности в его каталитический центр вносятся мутации [7]. В работе Knipping и др. для инактивации генов *CCR5* и *CXCR4* применен подход на основе внесения преждевременных стоп-кодонов или элиминации старт-кодона с использованием редакторов оснований [89]. Данный пример близок к предложенной нами стратегии, поэтому рассмотрим его подробнее.

Для инактивации генов корецепторов посредством внесения преждевременных стоп-кодонов применён редактор оснований цитозина BE4max (рис. 4А). Дополнительно для инактивации гена *CCR5* путём

элиминации старт-кодона применён редактор оснований аденина ABE8e (рис. 4Б).

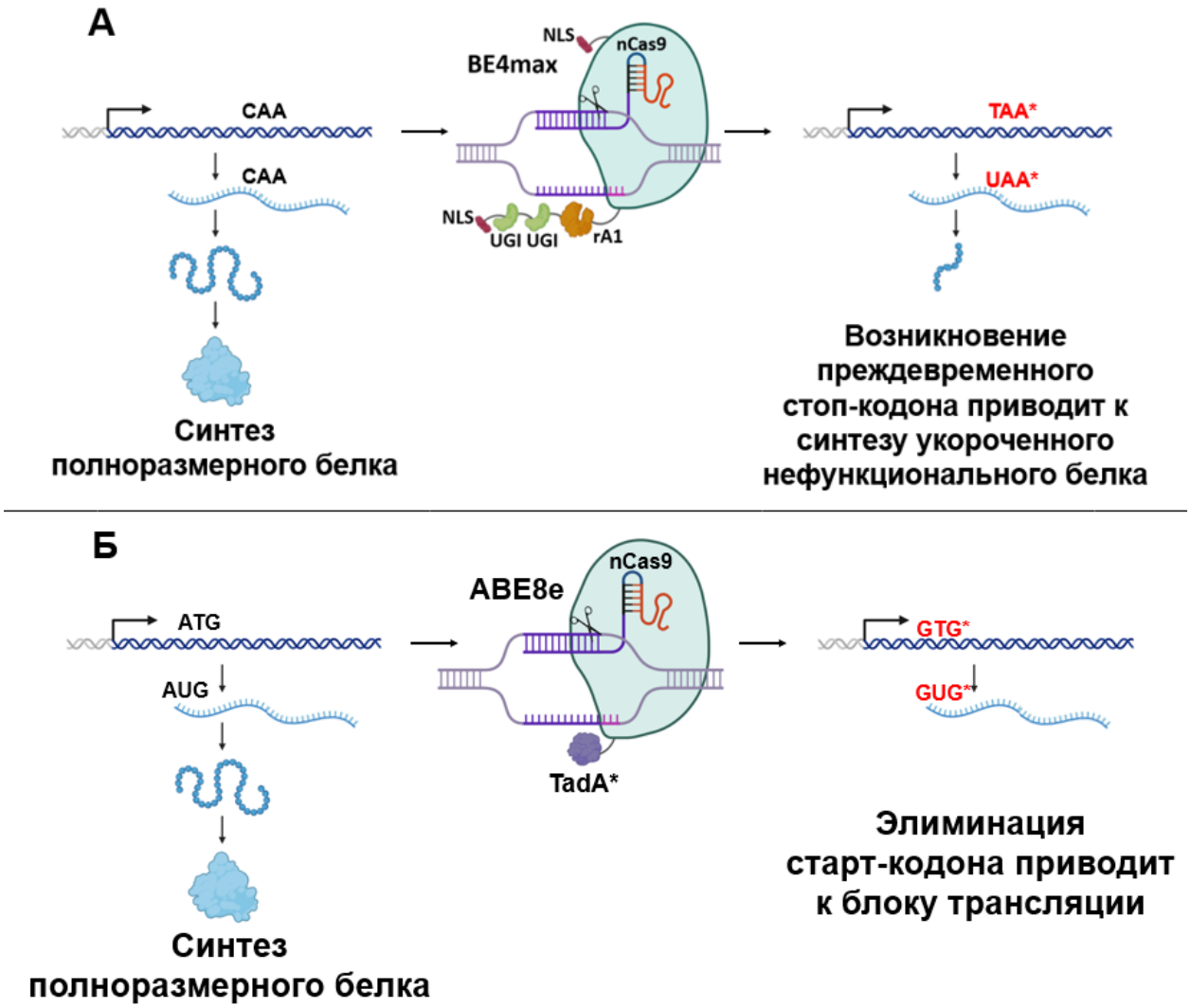


Рисунок 4. Принцип применения редактора оснований цитозина для внесения преждевременного стоп-кодона с целью инактивации гена (А). Принцип применения редактора оснований аденина для элиминации старт-кодона с целью инактивации гена (Б). Сделано с помощью BioRender.

мРНК, кодирующие редакторы оснований и различные sgRNA, вводили в активированные Т-лимфоциты методом электропорации. С помощью высокопроизводительного секвенирования нового поколения (NGS) проводили оценку эффективности редактирования для каждой sgRNA. При инактивации *CXCR4* sgRNA, именуемая как X4-2, привела к самой высокой эффективности

редактирования – 89,1 (±2,6) % последовательностей содержали желаемые мутации. Эффективность инактивации гена *CCR5* с помощью BE4max достигла 68,4 (±8) % в случае применения R5-1 sgRNA. Элиминация старт-кодона *CCR5* с помощью ABE8e показала самую высокую эффективность – 87 (±7,6) % целевых последовательностей содержали мутированный старт-кодон. Дополнительный анализ нецелевой активности выявил неспецифическое редактирование преимущественно в межгенных и интронных областях генома. После выбора sgRNA, обеспечивающих самую высокую эффективность редактирования, проводили оценку уровня экспрессии генов корецепторов в Т-клетках. Для этого на 11 день после редактирования клетки анализировали методом проточной цитометрии. Каждая из примененных sgRNA обеспечивала стабильное снижение экспрессии *CXCR4* и *CCR5* на 75–86 % и 80–94 % соответственно. Чтобы оценить влияние инактивации генов корецепторов на приобретение устойчивости к ВИЧ-1, отредактированные Т-клетки подвергали заражению репликативно-дефектными лентивирусами (LV) различной тропности: к CCR5 или CXCR4 корецептору. Для простоты детекции событий инфицирования LV содержали гены флуоресцентных белков: GFP для CCR5-тропного LV и RFP для CXCR4-тропного LV. Эффективность защиты от инфекции достигла около 88 % и 86 % соответственно. Результаты одновременной инактивации обоих генов показали сопоставимую эффективность. Анализ жизнеспособности отредактированных Т-клеток подтвердил сохранение функциональности.

Принимая во внимание описанные в данном обзоре подходы генной терапии, можно заключить, что редакторы оснований на основе CRISPR-Cas систем являются перспективным инструментом не только для генной терапии ВИЧ-инфекции, но и других вирусных и наследственных заболеваний. Данный подход заимствован из природы: в основе лежат механизмы противодействия мобильным генетическим элементам (МГЭ), повсеместно распространённым в геномах прокариот и эукариот [90, 91]. Стоит отметить, что данные механизмы подразумевают не только удаление или модификацию нуклеотидной последовательности с помощью редакторов оснований, но и интерференцию

транскрипции с МГЭ путём прочного связывания с ними белковых комплексов защиты. Подход, предложенный нами, заключается в мутагенезе старт-кодона гена *gag* с использованием двойных редакторов оснований. Отличием нашего подхода является применение в составе редактора новой рекомбинантной аденозиндезаминазы TadDE, полученной методом направленной эволюции, что концептуально позволит внести множественные мутации в локус петли SL4 PSI провируса ВИЧ-1 [10]. Для нацеливания предполагается использование нативной формы каталитически неактивного MmCas12m и широко применяемого dSpCas9. Полное отсутствие каталитической активности, даже исключение никазной формы, является взвешенным шагом, направленным на снижение генотоксичности, возникающей по причине внесения одно- и двухцепочечных разрывов.

1.4 CRISPR-Cas системы в геномном редактировании

Прокариотические организмы подвержены постоянному воздействию различных факторов окружающей среды. Среди таких факторов стоит отметить проникновение чужеродной – экзогенной – ДНК в результате вирусной трансдукции, трансформации или конъюгации. Проникновение экзогенной ДНК способно стать причиной гибели клетки, например, в результате заражения бактериофагом [92]. В ответ на подобные угрозы у прокариотических организмов появились защитные механизмы, позволяющие клетке отличать собственную ДНК от экзогенной и принимать меры по элиминации чужеродных последовательностей. Стратегия основана на сохранении экзогенной ДНК в виде коротких палиндромных повторов – Clustered Regularly Interspaced Short Palindromic Repeats, сокращённо – CRISPR. Данные последовательности служат шаблоном для интерференции чужеродной ДНК.

Защитный механизм CRISPR-Cas реализуется в три этапа: адаптации, экспрессии и интерференции. На первом этапе последовательность вторгшегося генетического элемента, называемая протоспейсером, встраивается в массив CRISPR, образуя новый спейсер. Это событие позволяет организму-хозяину

запомнить генетический материал чужеродного агента и демонстрирует адаптивную природу этой «иммунной системы». Процесс гидролиза и интеграции экзогенной ДНК обеспечивает белковый комплекс Cas1-Cas2. Важно отметить, что выбор целевой последовательности для интеграции в локус CRISPR не является случайным. Так, например, в системах CRISPR-Cas I, II и V типов короткая последовательность, называемая прилегающим мотивом протоспейсера – PAM, находится непосредственно рядом с протоспейсером и играет ключевую роль в адаптации [93–98]. Процесс интеграции протоспейсеров осуществляется с помощью фактора интеграции хозяина (IHF). IHF связывается с ДНК хозяина в области лидерной последовательности CRISPR, сворачивает ДНК в U-образную конфигурацию, после чего происходит интеграция протоспейсера в геном хозяина между лидерной последовательностью и 5’-фосфатом проксимального ряда повторов-спейсеров [98–101]. Интегрированный в геном ДНК-протоспейсер теперь именуется спейсером. Следующая стадия – экспрессия – заключается в синтезе длинного транскрипта pre-crRNA (незрелого предшественника crRNA) и его процессировании с образованием коротких зрелых молекул crRNA, содержащих последовательность спейсера [102]. У CRISPR-Cas систем I и III типа процессинг обеспечивается рибонуклеазами семейства белков Cas6 [103]. CRISPR-Cas системы типа II задействуют для процессинга фермент РНКазу III [104]. На этапе интерференции зрелая crRNA служит направляющей для Cas-белков и обеспечивает распознавание последовательности-мишени в чужеродной ДНК или РНК. Чужеродная нуклеиновая кислота гидролизуется, либо происходит интерференция транскрипции путём прочного связывания рибонуклеопротеинового комплекса с чужеродной ДНК [5]. Данный механизм адаптирован для широкого применения в лабораторных исследованиях.

Одной из самых широко применяемых в исследованиях нуклеаз является Cas9, выделенная из бактерии *Streptococcus pyogenes* (SpCas9). Однако находят применение и другие варианты, например, SaCas9, выделенный из *Staphylococcus aureus*. Вариант SaCas9 отличается более компактными

размерами (1053 амк) по сравнению с SpCas9 (1368 амк), что даже позволяет поместить его в экспрессионную кассету AAV [105]. Взаимодействие Cas-белка с нуклеотидной последовательностью ДНК опосредуется sgRNA, которая может быть представлена в виде комплекса из tracrRNA (транс-активирующая РНК) и crRNA, либо объединённой единой молекулы – sgRNA [106]. Локус, на который предполагается таргетирование, должен содержать короткую последовательность нуклеотидов, примыкающую к его 3'-концу, именуемую PAM (protospacer adjacent motif, прилегающий мотив протоспейсера) [107]. Например, для SpCas9 PAM является последовательность 5′-NGG-3′, где N – любой из четырёх нуклеотидов. Однако сохраняется способность связывания с целевой последовательностью, примыкающей к неоптимальному PAM. Так, например, для SpCas9 выявлено сохранение нуклеазной активности в случае применения sgRNA, обуславливающих таргетирование белкового комплекса к неоптимальной последовательности 5′-NAG-3′ и 5′-NGA-3′ [108]. Ортологи Cas12a (AsCpf1, LbCpf1), имея в качестве PAM последовательность 5'-TTTV-3', способны к узнаванию и таких неоптимальных PAM, как 5′-CTTV-3′, 5′-TCTV-3′ и 5′-TTCV-3′ [109].

Система CRISPR-Cas внесла значительный вклад в развитие генного редактирования. CRISPR-Cas активно применяется в области сельского хозяйства, получения лабораторных моделей животных и в генной терапии, что описано в многочисленных оригинальных исследованиях и обзорах [110–112]. Желаемый результат может достигаться путём делеции целевого локуса. Для этого подбирают несколько sgRNA, обеспечивающих таргетирование комплекса CRISPR-Cas по обе стороны от удаляемого локуса (рис. 5А). Одновременное внесение двухцепочечных разрывов в местах, фланкирующих локус, приводит к его выпадению из генома [113, 114]. При этом размер удаляемого фрагмента может составлять порядка 100 т.п.н. [115]. В отдельных случаях удавалось достичь и большей длины удаляемого фрагмента, однако стоит учитывать, что эффективность удаления снижается по мере увеличения размера фрагмента [116]. Если задача заключается в инактивации целевого гена путём сдвига рамки

считывания, то sgRNA для CRISPR-Cas подбирают на белок-кодирующую последовательность гена. В данном случае важно, чтобы репарация двухцепочечного разрыва осуществлялась по пути негомологичного соединения концов (рис. 5Б). Данный путь репарации отличается невысокой точностью, что приводит к замене, выпадению или вставке лишних нуклеотидов [117]. Преждевременный стоп-кодон или сдвиг рамки считывания приводят к синтезу нефункционального продукта. Так, в частности, были получены иммунодефицитные линии мышей с нокаутом гена *Foxn1* путём воздействия на его первый экзон системой CRISPR-Cas9 [118]. Возможно и применение CRISPR-Cas систем для инсерции трансгенов или других интересующих последовательностей ДНК в целевой локус. CRISPR-Cas, как и в предыдущих применениях, вносит двухцепочечный разрыв в место предполагаемой инсерции. Но в данном случае в системе дополнительно присутствует донорная матрица с гомологичными участками по обе стороны от места разрыва, которая направляет репарацию по пути гомологичного соединения концов. Это приводит к встраиванию экзогенной ДНК в геном клетки-хозяина (рис. 5В). Размер инсерции может достигать несколько т.п.н. [119, 120]. Часто при выборе места инсерции применяют так называемые «безопасные гавани» – «safe harbor». Инсерция в безопасные гавани не приводит к нарушению функций или уровня экспрессии эндогенных генов [121, 122]. У мышей, в частности, такими «гаванями» являются Rosa26, у приматов – AAVS1, у кроликов – rbRosa26 [123, 124].

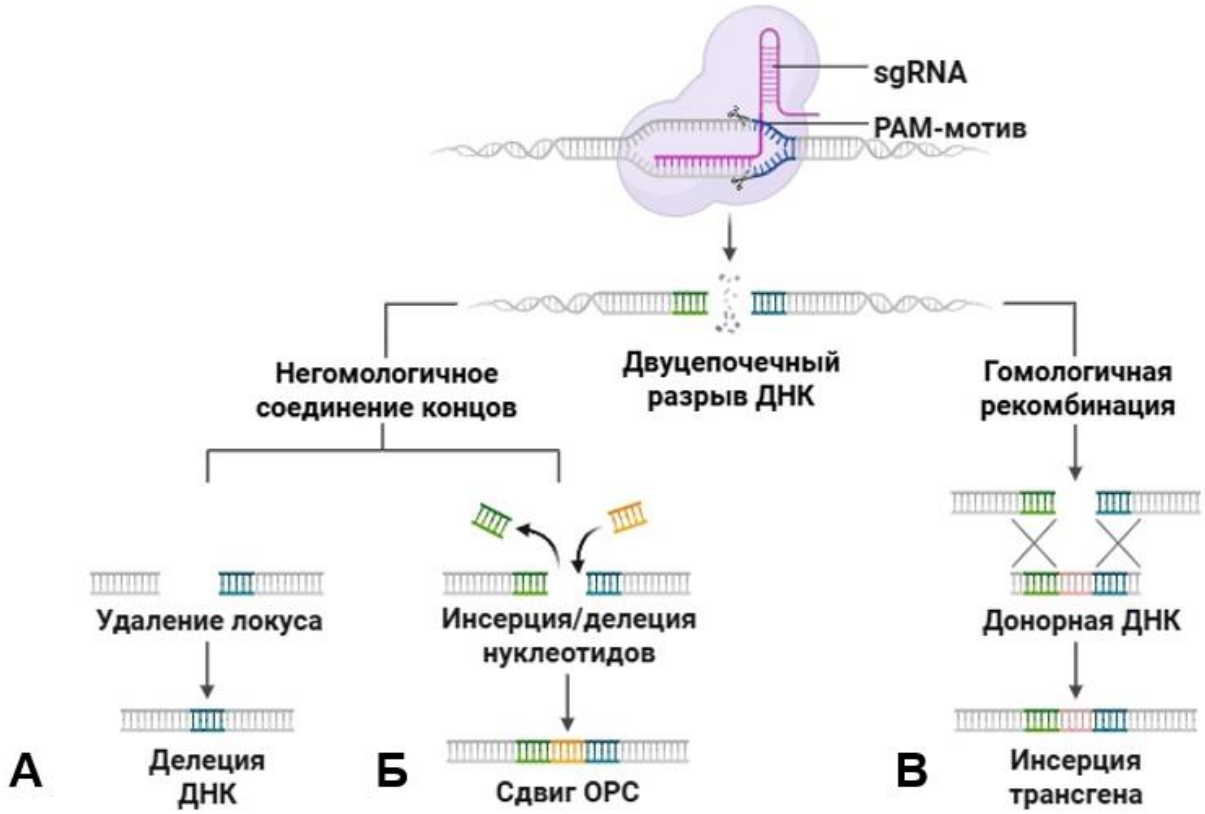


Рисунок 5. Варианты применения CRISPR-Cas систем: делеция локуса (А), инактивация гена в результате сдвига открытой рамки считывания – ОРС (Б), инсерция целевого трансгена в место двухцепочечного разрыва по матрице экзогенной ДНК (В). Сделано с помощью BioRender.

CRISPR-Cas системы могут использоваться не только для внесения двухцепочечных разрывов в ДНК, но и как платформы для таргетирования других белковых компонентов к искомой последовательности ДНК. Так, CRISPR-Cas системы успешно применялись для доставки ДНК-метилтрансфераз, белковых факторов активации и интерференции транскрипции, интеграз [125–128]. Безусловно, CRISPR-Cas системы послужили и эффективной платформой для таргетирования рекомбинантных вариантов дезаминаз, став основой для создания инструмента под названием редактор оснований ДНК [129–131]. Опубликованы многочисленные научные работы, описывающие применение редакторов оснований для разработки подходов генной терапии не только наследственных, но и вирусных заболеваний человека [132]. В феврале 2025 года первым в мире пациентом, получившим

персонализированный препарат редактора оснований на основе CRISPR-Cas, стал младенец из Филадельфии, страдающий от недостаточности карбамоилфосфатсинтетазы 1 в результате мутации гена *CPS1* [133]. Продукт данного гена отвечает за преобразование аммиака в карбамоилфосфат на первом этапе орнитинового цикла («цикл мочевины»). Мутация в гене *CPS1* приводит к невозможности данного процесса и, как следствие, к накоплению токсичного аммиака в крови [134]. Спустя 7 недель после первой инфузии пациент мог получать повышенное количество белка с пищей и сниженную вдвое по сравнению с начальной дозу препарата, связывающего азот, без возникновения нежелательных побочных эффектов. Серьёзных побочных эффектов не выявлено, наблюдения продолжаются. Данное достижение является ярким примером актуальности разработки генотерапевтических препаратов, которые позволят повысить качество жизни людей с наследственными заболеваниями.

1.5 Разнообразие редакторов оснований ДНК

1.5.1 *Редакторы оснований цитозина (Cytosine base editors – CBE)*

Основой для CBE послужила группа природных ферментов цитозиндезаминаз. Способность представителей данной группы ферментов заключается в дезаминировании цитозина (C), что приводит к его превращению в урацил (U), который образует водородную связь с аденином (A). В таком случае воздействие CBE приводит к замене пары C-G на пару T-A после репликации. У человека разнообразие цитозиндезаминаз представлено семейством AID/APOBEC, представители которого задействованы для внутриклеточных процессов: защита от вирусов, редактирование матричных РНК, внесение однонуклеотидных мутаций в ДНК в ходе соматического гипермутагенеза и переключение изотипов антител [135]. Первый рекомбинантный ДНК-специфичный CBE был создан путём слияния крысиного APOBEC1 (rAPOBEC1) с dSpCas9 посредством линкера XTEN (рис. 6). Так был получен рекомбинантный белок, показавший 50 % эффективность дезаминирования цитозина ДНК в условиях *in vitro*. Впрочем, испытания *in vivo*

продемонстрировали эффективность на уровне 0,8–7,7 % [136]. Низкая эффективность объясняется ответной реакцией ферментов репарации на образующийся в результате дезаминирования гетеродуплекс U-G в ДНК. Неканоническая пара оснований, вызванная возникновением U в положении C, приводит к активации урацил-ДНК-гликозилазы (UDG). Задача UDG заключается в удалении U из ДНК в геноме клеток [137]. Дальнейшие механизмы эксцизионной репарации оснований приводят к восстановлению C в участке редактирования [138].

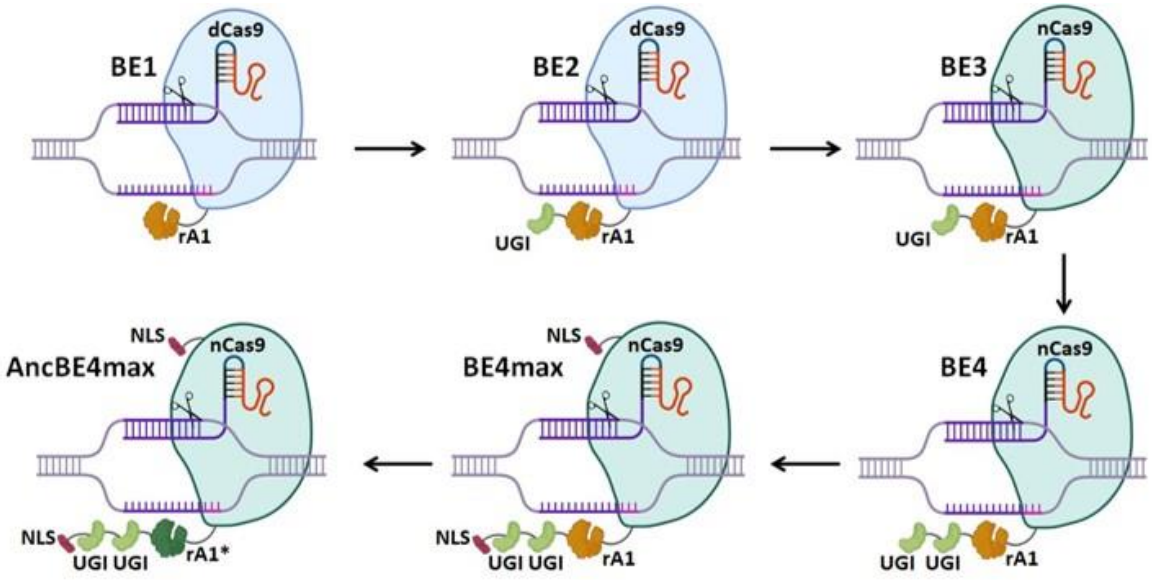


Рисунок 6. Схематичные изображения редакторов оснований цитозина различных поколений. rA1 – rAPOBEC1, rA1* – кодон-оптимизированный rAPOBEC1, dCas9 – каталитически неактивная форма Cas9, nCas9 – никазная форма Cas9, UGI – ингибитор урацил-ДНК-гликозилазы, NLS – сигнал ядерной локализации. Подробное описание в тексте. Сделано с помощью BioRender.

Чтобы повысить эффективность защиты урацила от ферментов репарации к C-концу BE1 добавили ингибитор урацил-ДНК-гликозилазы бактериофага PBS1 – UGI [136, 139]. Так возник редактор оснований второго поколения – BE2 (рис. 6). Эффективность редактирования в условиях *in vivo* при использовании BE2 была в среднем в 3 раза выше, чем при использовании BE1, что привело к повышению доли редактированных последовательностей до 20 % от общего

объёма секвенированной ДНК. Чтобы повысить эффективность редактирования оснований сверх этого предела, предпринимались попытки инициировать репарацию неотредактированной цепи ДНК, содержащей G [136]. Для этого в dSpCas9 восстановили остаток гистидина в положении 840, что привело к возможности внесения одноцепочечного разрыва в цепь ДНК, спаренной с sgRNA. Внесение подобного разрыва приводит к репарации по матрице уже отредактированной цепи, в которой после редактирования располагается U. Полученный редактор третьего поколения (BE3) сохраняет мутацию D10A в SpCas9, что предотвращает внесение двухцепочечных разрывов в ДНК. UGI также сохранен на С-конце белкового комплекса (рис. 6). Эффективность редактирования целевых участков генома клеток линии HEK293 достигла около 37 %, что в 1,5–2 раза выше в сравнении со вторым поколением. Дальнейшие работы по улучшению редакторов оснований цитозина привели к созданию новых вариантов. Одна из основных модификаций заключалась в добавлении второй копии UGI к С-концу BE3 (рис. 6). Так удалось получить редактор четвертого поколения – BE4 [140]. Дальнейшее добавление сигналов ядерной локализации (NLS) как на N-, так и на С-конец белкового комплекса (BE4max) и замена rAPOBEC1 на оптимизированный APOBEC (AncBE4max) привели к повышению эффективности редактирования до 67 % *in vivo* [140] (рис. 6). Для всех редакторов оснований цитозина характерно широкое окно редактирования в позициях 1–14 протоспейсера. Пик эффективности, как правило, приходится на позиции 4–8 протоспейсера. Реже для создания редакторов оснований цитозина применяют цитозиндезаминазу CDA1, полученную из морской миноги (*Petromyzon marinus*). В целом, подобные редакторы оснований с использованием pmCDA1 обеспечивают эффективность редактирования на уровне BE4max. Однако в отличие от APOBEC1 pmCDA1 обеспечивает одинаковую эффективность редактирования при расположении как на N-конце от Cas-белка, так и на С-конце. APOBEC1 в свою очередь почти нефункциональна при расположении на С-конце от Cas-белка [141]. Это ещё раз подчеркивает специфику редактирования в зависимости от применяемой

дезаминазы. Таким образом возникает необходимость тщательного анализа особенностей белковых комплексов при создании новых редакторов оснований.

1.5.2 *Редакторы оснований аденина (Adenine base editors – ABE)*

При дезаминировании аденозина образуется инозин (I), который воспринимается ферментами матричного синтеза как G, в результате чего при синтезе дочерней цепи ДНК происходит замена T на C. Природные аденозиндезаминазы существуют, но субстратом для них преимущественно являются молекулы РНК. Так, например, у человека целое семейство аденозиндезаминаз, специфичных к двухцепочечной РНК, кодируются генами группы ADAR [142]. Разнообразие расширяют многочисленные аденозиндезаминазы тРНК – TadA, специфичные к одноцепочечной РНК [143]. Тестирование различных природных аденозиндезаминаз на способность редактирования молекул ДНК не показало обнадеживающих результатов [144]. С целью создания аденозиндезаминазы, специфичной к ДНК, применен подход направленной эволюции. За исходную форму для создания ДНК-специфичной аденозиндезаминазы взяли тРНК-аденозиндезаминазу *Escherichia coli* – ecTadA, которая преобразует A в I в одноцепочечной части антикодоновой петли тРНК-Arg [145]. Выбор обусловлен тем, что по структуре ecTadA близка к APOBEC/AID (цитидиндезаминаза), которая обладает нативной способностью взаимодействовать с одноцепочечной ДНК. В дополнение ecTadA проявляет природное взаимодействие с одноцепочечной частью тРНК, что соответствует заданной цели – редактирования одноцепочечного участка ДНК [146]. Внесение мутаций A106V и D108N в ecTadA и её слияние с nSpCas9 (D10A) посредством XTEN-линкера привело к созданию первого рекомбинантного редактора оснований аденина, названного ABE1.2. Однако эффективность редактирования в условиях *in vitro* целевых последовательностей ДНК посредством полученного фермента оказалось низкой – около 3,2 %. Но приобретение специфичности к ДНК послужило отправной точкой к созданию последующих поколений ABE с уже большей эффективностью и сниженным нецелевым эффектом. Полученный спустя ряд модификаций ABE7.10 показал 58 % эффективность редактирования

в целевых локусах клеток линии HEK293 [145]. В исходном варианте ABE7.10 представлял гетеродимерную конструкцию, в которой нуклеазный домен nSpCas9 был слит одновременно с нативной и с рекомбинантной формами ecTadA. Тем не менее последующие экспериментальные данные продемонстрировали, что присутствие нативной копии ecTadA в такой системе не является обязательным. [146, 147]. Дополнительно стоит отметить, что применение ингибиторов ДНК-гликозилаз, как в случае CBE, не требуется, поскольку удаление инозина при репарации является маловероятным событием [148, 149]. Дальнейшие модификации ABE7.10 привели к созданию двух новых вариантов ABE8e и ABE8.20, которые показали повышенную эффективность и скорость редактирования. В составе данных редакторов оснований используется модифицированная дезаминаза TadA-8e, несущая 8 аминокислотных замен в сравнении с TadA7.10. Внесенные модификации позволили повысить связывание с ДНК и увеличить эффективность редактирования [150]. Если считать от PAM-дистального конца протоспейсера, то ABE7.10 редактирует преимущественно позиции 4–8 протоспейсера, ABE8.20 — позиции 3–9, а ABE8e — позиции 3–11. Такое расширение области редактирования объясняется повышением каталитической эффективности и процессивности фермента TadA в ходе направленной эволюции [151]. Стоит отметить, что повышенная эффективность связывания TadA-8e с ДНК привела к схожей эффективности редактирования при расположении домена дезаминазы на N- или C-конце от Cas-белка [152].

1.5.3 *Двойные редакторы оснований (Dual base editors)*

Создание единого редактора оснований, способного к дезаминированию аденозина и цитозина, делает систему редактирования более универсальной, что может быть удобным и оправданным в некоторых случаях, в частности, при необходимости множественного мутагенеза. Подобные редакторы оснований получены путём слияния дезаминаз из BE3 и ABE7.10 с nSpCas9 (D10A) [153]. Анализ эффективности выявил закономерность: активность домена

цитозиндезаминазы значительно повышена, когда она присоединена к N-концу аденозиндезаминазы (рис. 7А). Добавление NLS и двух копий UGI привело к повышению эффективности редактирования. При тестировании эффективности редактирования оснований на клетках линии HeLa доля одновременных мутаций обоих типов (A>G, C>T) в одном окне редактирования сильно варьирует в диапазоне 2–30 % в зависимости от локуса [153].

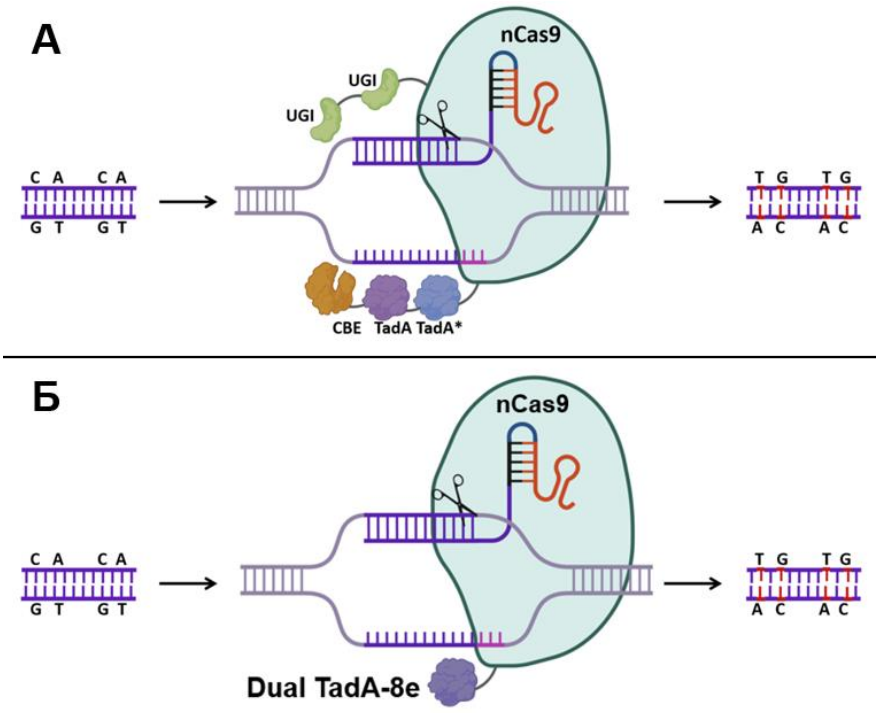


Рисунок 7. Схематичное изображение строения и принципа действия двойного редактора оснований. Двойной редактор оснований ДНК старого поколения (А): nCas9 – никазная форма Cas9, UGI – ингибитор урацил-ДНК-гликозилазы, CBE – рекомбинантная цитозиндезаминаза, TadA – рекомбинантная аденозиндезаминаза, TadA* – нативная аденозиндезаминаза. Двойной редактор оснований ДНК нового поколения, полученный путём направленной эволюции (Б): Dual TadA-8e – рекомбинантная аденозиндезаминаза, обладающая способность к дезаминированию аденина и цитозина. Сделано с помощью BioRender.

Безусловно, подобная система, сочетающая в себе несколько функционально значимых доменов, имеет специфику, которая может являться недостатком в зависимости от цели исследователя. Во-первых, оба домена

стерически конкурируют в целевом локусе. Во-вторых, большие размеры подобной системы затрудняют её доставку в клетки тканей и органов, если речь идёт о генной терапии. Широко распространенные для генной терапии векторы на основе AAV имеют ограничения по размеру переносимой экспрессионной кассеты – около 4700 нуклеотидов [154, 155]. Это накладывает определённые ограничения на применение AAV для доставки двойных редакторов оснований. С целью решения данной задачи предприняты попытки создания сверхкомпактного двойного редактора оснований. С помощью направленной эволюции рекомбинантной аденозиндезаминазы TadA-8e удалось получить новый редактор оснований, обладающий способностью к одновременному дезаминированию A и C с эффективностью около 80 % и 73 % соответственно [10] (рис. 7Б). Новый фермент, получивший название TadDE (TadA-8e-based Dual Editor), обладает компактными размерами – 167 амк, что может значительно упростить его доставку в клетки *in vitro* или *in vivo* с помощью AAV.

1.5.4 *Метод праймированного редактирования* (*Prime editing)*

Несмотря на широкие возможности CBE, ABE и двойных редакторов оснований, они способны вносить лишь точечные мутации, ограниченные четырьмя типами транзиций в целевом участке ДНК. Ни один из данных редакторов оснований не способен заменить одну последовательность ДНК на другую, строго выбранную исследователем. В работе Anzalone и др. описан метод, позволяющий преодолеть эти недостатки, – праймированное редактирование [156]. Как и в случае с описанными выше редакторами оснований, праймированное редактирование не требует внесения двухцепочечных разрывов в ДНК. Редактирование осуществляется с помощью инструмента, названного prime editor (PE), который состоит из nSpCas9 (H840A) и модифицированной обратной транскриптазы в комплексе с направляющей для праймированного редактирования РНК (prime editing guide RNA – pegRNA) (рис. 8, 1). Важно отметить, что pegRNA существенно отличается от sgRNA для CRISPR-Cas систем. pegRNA содержит не только последовательность, комплементарную целевому участку, но и дополнительную последовательность,

определяющую требуемые изменения [157]. 5'-конец pegRNA взаимодействует с участком связывания праймера (PBS) на ДНК, освобождая вторую некомплементарную цепь. Таким образом, одноцепочечная ДНК, содержащая PAM, гидролизуется с помощью nSpCas9. В результате возникает освободившийся 3’-конец, который служит затравкой для обратной транскриптазы. Гидролизованная PAM-содержащая цепь удлиняется с помощью обратной транскриптазы, которая использует внутреннюю часть pegRNA в качестве матрицы, последовательность матрицы определяется исследователем в соответствии с задачей (рис. 8, 2). В результате этого этапа образуется повтор PAM-содержащей цепи ДНК: отредактированный 3'-участок, полученный по матрице pegRNA, и исходный, неотредактированный 5'-участок. Выбор участка, который будет гибридизоваться с цепью ДНК, не содержащей PAM, — это равновесный процесс, в котором термодинамически предпочтительным, скорее всего, будет идеально комплементарный 5'-участок (рис. 8, 3). На рис. 8 изображен «выбор» редактированной цепи для гибридизации. В случае такого «выбора» нативный фрагмент разрушится клеточными эндонуклеазами, а возникшие несоответствия между нуклеотидными последовательностями редактированной и нередактированной цепи будут устранены в ходе репарации неспаренных оснований [158–160] (рис. 8, 5).

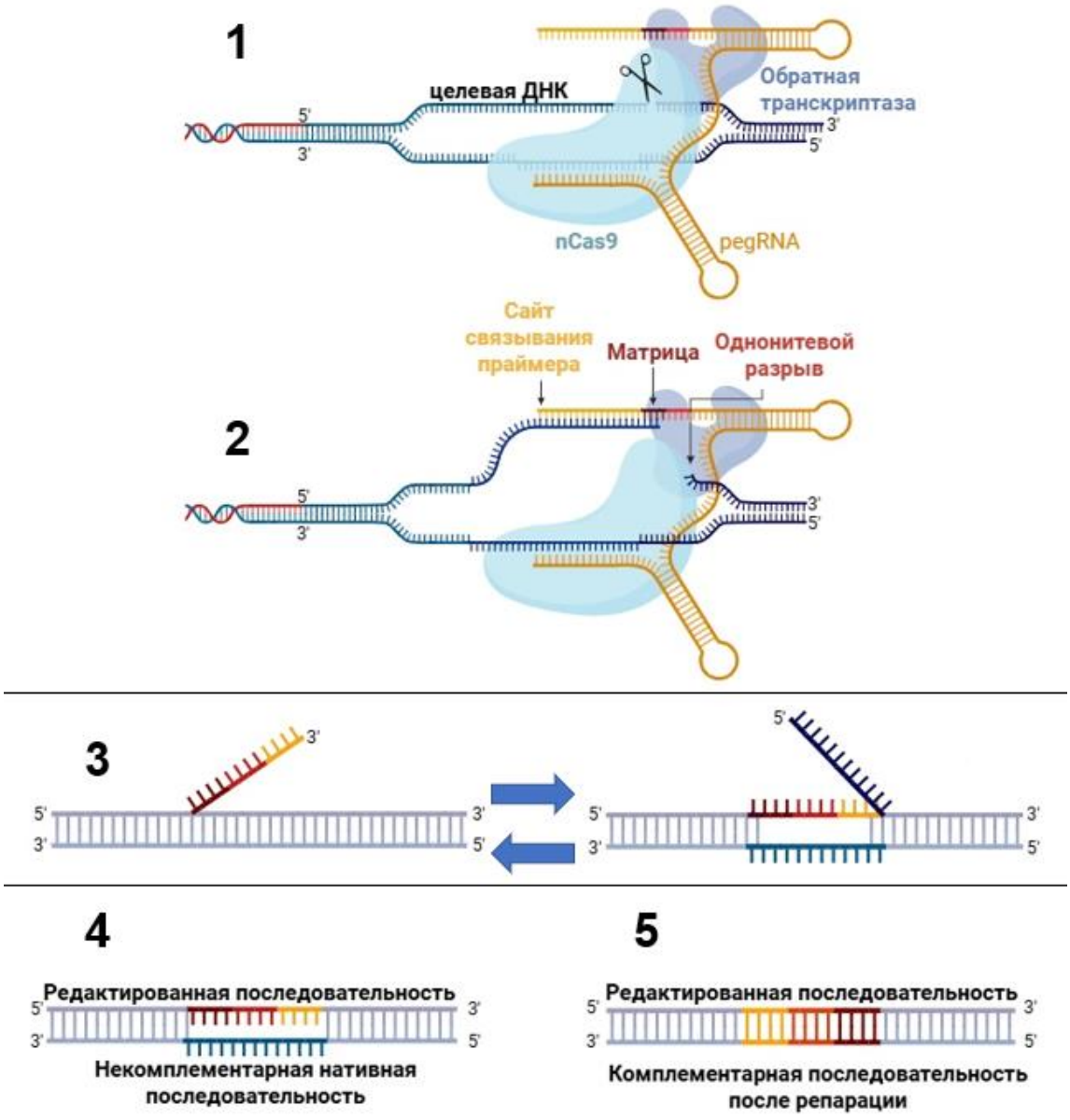


Рисунок 8. Схематичное изображение принципа метода праймированного редактирования. Подробное описание в тексте. Сделано с помощью BioRender.

Первое поколение PE (PE1) состояло из обратной транскриптазы вируса лейкемии мышей Молони (MMLV RT), присоединённой к С-концу nSpCas9 (H840A), и pegRNA, синтез которой обеспечивала вторая плазмида. Эффективность редактирования *in vitro* с помощью PE1 достигала максимального уровня в 0,7–5,5 % [156]. При создании второго поколения PE2 в обратную транскриптазу были внесены мутации D200N, L603W, T330P, T306K и W313F, приводящие к повышению эффективности редактирования в несколько

раз по сравнению с PE1 [156]. В рамках реализации третьего поколения PE3 была добавлена sgRNA для внесения разрыва в неотредактированную цепь. Различные sgRNA, обуславливающие дополнительный разрыв на расстоянии 14–116 п.н. от действия pegRNA, показали повышение эффективности редактирования до 12 %. Оптимальным, однако, было выбрано расстояние 40–90 п.н. от pegRNA-индуцированного разрыва [156]. Впрочем, разрыв в неотредактированной цепи должен вноситься только после репарации первого разрыва, опосредованного pegRNA. В противном случае мутагенный эффект вследствие двухцепочечного разрыва сохраняется. Для достижения этой цели были разработаны sgRNA с протоспейсерами, которые соответствуют редактированной цепи и могут образовывать устойчивый дуплекс только с ней. При использовании этой стратегии, названной PE3b, несовпадения между протоспейсером и неотредактированной цепью должны препятствовать внесению разрыва, опосредованного sgRNA, до тех пор, пока не произойдёт редактирование цепи, содержащей PAM для pegRNA. В результате использования PE3b средняя частота инсерций/делеций (0,74 %) была в 13 раз ниже, чем при использовании PE3, без какого-либо заметного снижения эффективности редактирования [156]. В рамках создания последующих поколений к PE2 и PE3 добавлен доминантно-негативный мутант белка MLH1dn, нативная форма которого является компонентом комплекса репарации ошибочно спаренных оснований MutSα–MutLα. MLH1dn временно подавляет репарацию, конкурируя с нативным белком, что обеспечивает копирование последовательности отредактированной цепи на неотредактированную. Новые редакторы назвали PE4 и PE5 соответственно. Эти системы повысили эффективность редактирования в 7,7 раза (PE4 по сравнению с PE2) и в 2 раза (PE5 по сравнению с PE3) [158]. При создании поколения PE6 применяли компактные варианты обратных транскриптаз Ec48 *E. coli* и Tf1 *S. pombe*, что позволило сократить ОРС редактора в среднем до 5,5 т.п.н. [159]. Кодон-оптимизация последовательности обратной транскриптазы MMLV RT для клеток человека, а также добавление двух сигналов ядерной локализации в поколениях PE2–5 позволили создать PEmax

[158]. Добавление к С-концу PEmax белка La привело к созданию последнего поколения PE7, который обеспечивает эффективность редактирования до 25 %. La усиливает прайм-редактирование в клетках млекопитающих путём связывания с pegRNA для защиты от экзонуклеаз [160]. Праймированное редактирование будет оправдано в тех случаях, где необходима замена одной последовательности на другую, строго определённую. Подход позволяет проводить вставку желаемых фрагментов ДНК длиной до 150–200 п.н. [161]. Впрочем, фактическая эффективность может быть существенно ниже – около 40 п.н. [162].

1.6 Заключение

В обзоре представлены данные об эпидемиологической обстановке ВИЧ-инфекции в Российской Федерации, о морфологии и жизненном цикле ВИЧ-1. Описаны перспективные подходы генной терапии ВИЧ-инфекции с использованием инструментов геномного редактирования на основе CRISPR-Cas систем и редакторов оснований. Рассмотрены структурные и функциональные особенности различных инструментов редактирования оснований. Можно заключить, что создание новых вариантов и усовершенствование уже имеющихся инструментов геномного редактирования – повышение их эффективности, уменьшение размеров комплекса, снижение нецелевого эффекта – несомненно расширяет перспективы и возможности их применения. Редакторы оснований являются крайне удобным инструментами для внесения преждевременных стоп-кодонов или элиминации старт-кодонов. Это позволяет точечно инактивировать ген, представляющий интерес, без внесения разрывов в ДНК, что снижает риск нежелательных геномных перестроек. Разработка препаратов генной терапии на основе редакторов оснований является перспективным направлением прецизионной медицины.

ГЛАВА 2. МАТЕРИАЛЫ И МЕТОДЫ

2.1. Материалы

2.1.1 Коммерческие наборы

- Набор для выделения ДНК из агарозных гелей QIAquick Gel Extraction Kit («QIAGEN», ФРГ);

- Набор для выделения плазмидной ДНК Monarch Plasmid Miniprep Kit («New England Biolabs», США);

- Набор для выделения плазмидной ДНК без эндотоксинов EndoFree Plasmid Maxi Kit («QIAGEN», ФРГ);

- Набор для выявления контаминации микоплазмой культур клеток MycoReport («Евроген», РФ);

- Набор для очистки ДНК из агарозного геля и реакционных смесей Cleanup Mini («Евроген», РФ);

- Набор для подготовки библиотек к нанопоровому секвенированию CycloneSEQ 24 Barcode Library Prep Set («MGI Tech», КНР);

- Набор для постановки количественной ПЦР в реальном времени БиоМастер HS-qPCR (2×) («Биолабмикс», РФ);

- Набор для постановки ПЦР БиоМастер HS-Taq ПЦР (2×) («Биолабмикс», РФ);

- Набор для секвенирования по методу Сэнгера Brilliant Dye v3.1 mix («NimaGen», США);

- Набор РИБО-преп для выделения ДНК/РНК из биологического материала (ФБУН ЦНИИ Эпидемиологии Роспотребнадзора, РФ).

2.1.2 Питательные среды, сыворотки

- Питательная среда DMEM/F12 («БиолоТ», РФ) с добавлением 100 ед/мл пенициллина и 100 мкг/мл стрептомицина («БиолоТ», РФ);

- Эмбриональная телячья сыворотка – FBS («Thermo Fisher Scientific», США);

- Бактериальная среда LB-агар (100 мл: 1 г бактотриптон, 0,5 г дрожжевой экстракт, 1 г NaCl, 1,5 г бактоагар);

- Бактериальная среда LB (100 мл: 1 г бактотриптон, 0,5 г дрожжевой экстракт, 1 г NaCl).

2.1.3 Растворы, буферы и реагенты

- Раствор DPBS 1-кратный полный («БиолоТ», РФ);

- Раствор Трипсина-Версена (1:1) («БиолоТ», РФ);

- Диметилсульфоксид, для культур клеток («БиолоТ», РФ);

- Раствор Tris-HCl, pH 8,0, 1 М;

- Стерильный раствор трипанового синего 0,4 % («БиолоТ», РФ);

- 1×TAE буфер (40 мМ Tris-CH_3COOH, 2 мМ ЭДТА, pH 7,6);

- Буфер для ДНК-лигазы бактериофага T4 («Thermo Fisher Scientific», США);

- Буфер rCutSmart для эндонуклеаз рестрикции («New England Biolabs», США);

- Буфер для ДНК-полимеразы Q5 High-Fidelity («New England Biolabs», США);

- Буфер для нанесения образцов в гель-электрофорезе («SibEnzyme», РФ);

- Реагент для трансфекции Lipofectamine 3000 («Thermo Fisher Scientific», США);

- Буфер переноса (0,192 М глицин, 0,025 М Tris-HCl, 20 % этиловый спирт, pH 8,3);

- PBST (0,1 М NaCl, 0,02 М NaH_2PO_4, 0,08 М Na_2HPO_4, 0,05 % Tween-20, pH 7,4);

- Проявляющий раствор (фосфатно-цитратный буфер, 0,006 % перекись водорода, 0,04 % 4-хлор-1-нафтол (50 мг/мл в этаноле));

- Блокирующий раствор желатина для иммунохимии 10× Fish Gelatin Blocking Agent («Biotium Inc», США);

- Высоко-деионизированный формамид Hi-Di («Thermo Fisher Scientific», США);

- Маркер молекулярного веса ДНК 1 т.п.н. – 1kb ДНК-маркер («SibEnzyme», РФ);

- Маркер молекулярного веса ДНК 100 п.н. – 100+ bp ДНК-маркер («SibEnzyme», РФ);

- Маркер молекулярного веса белков, 20-200 кДа, предокрашенный («Servicebio», КНР);

- Изопропанол («База №1 Химреактивов», РФ);

- Хлороформ («База №1 Химреактивов», РФ);

- Агароза («Servicebio», КНР);

- Магнитные частицы для очистки ДНК из реакционных смесей AMPure XP («Beckman Coulter», США);

- ЭДТА («Sigma-Aldrich», США);

- Triton X-100 («MP Biomedicals», ФРГ);

- Фенилметилсульфонил фторид («Диаэм», РФ);

- Бромистый этидий («Sigma-Aldrich», США).

2.1.4 Ферменты

- ДНК-полимераза Q5 High-Fidelity («New England Biolabs», США);

- ДНК-лигаза бактериофага T4 («Thermo Fisher Scientific», США);

- Эндонуклеазы рестрикции *BmtI*, *BsaI*, *EcoRI*, *EcoRV*, *HindIII*, *KpnI*, *NheI*, *PciI*, *PvuI*, *SphI*, *XbaI*, *XcmI* («New England Biolabs», США);

- Эндонуклеаза I бактериофага T7 («New England Biolabs», США).

2.1.5 Праймеры

Олигонуклеотидные праймеры, использованные в работе для постановок ПЦР и секвенирования, представлены в таблице 1.

Таблица 1. Олигонуклеотидные праймеры, использованные в работе:

Название праймера	**Нуклеотидная последовательность 5′ – 3′**

F-mGFP_EcoRV	TAATAAGATATCATGGGTGTGATCAAGCCCGA
F-X_PA_7	TCACTCCAGCTAAGTTAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAG
F-partGAG_XbaI	TTATTATCTAGACCAAAAATTTTGACTAGCGGA
R-partGAG_EcoRV_1	ATCTAATTCTCCCCCGCTTAA
R-partGAG_EcoRV_3	ATCCAATTTTCCCCCGCTCAATACTGACGCTCTCGCACCCATGC
F-dATG_GAG	AGGAGAGAGATAGGTGCGAGAGC
R-dATG_GAG	TCTCGCACCTATCTCTCTCCT
F-dATG_GFP	ATTGGATATCATAGGTGTGATCA
F-SV40NLS_Cas12m_XmaI_1	CCCGGGCCCATGGGGCCCAAGAAAAAGCGGAAAGTG
F-SV40NLS_Cas12m_XmaI_2	AGAAAAAGCGGAAAGTGGGGCCCAGCGGC
F-Linker_1	AGGCCAAAAAGAAAAAGGGTCTCAGCGGTGGTAGCAGCGGCGG
F-Linker_2	GCGGTGGTAGCAGCGGCGGGCCTACCATGAGCGAAG
52bC12m	CCTTCTTTGTGGCGGCAGGTCTTTTGCCAGGATTGCTA
53bC12m	CTGCCGCCACAAAGAAGGCCGGACAGGCCAAAAAGAA
F-SV40NLS_Cas9	AGCGGAAAGTGGGGCCCAGCGGCGACAAGAAGTACAGC
R-nNLS_Cas9	CCTTCTTTGTGGCGGCAGGTCTTTTGCCGTCGCCTCCCAGCTG

F-PGKp_SV40NLS_PciI	TATATTATAAGGGTAGGGGAGGCGCTTTTCC
R-PGKp_SV40NLS	TTTCTTGGGCCCCATCGAAAGGCCCGGAGA
F-U6p_crR_C12m_1	GAAAGGACGAAACACCGTCATAGCCCAGCTTGGCGGGCGA
R-U6p_crR_C12m_2	GCTTGTGTCGGTCTCCGTCTTGGCCTTCGCCCGCCAAGCT
R-U6p_crR_C12m_3	AAAAAAAAAGAGACCACTGAAGCTTGTGTCGGTCTCCG
F-U6p_Cas9	GAGGGCCTATTTCCCATGATT
R-U6p_crR_C12m_fused	TCGCCCGCCAAGCTGGGCTATGACGGTGTTTCGTCCTTTC
R-U6p_Cas9	ACAAAAAAGCACCGACTCGGTGCCA
F-P2A_GFP	ATCAAGCTAGCCGCCACCATGGT
R-P2A_GFP	TTTCTTGGGCCCCATGGGCCCGGGGTTCTCCT
F-SV40NLS_TadDE	CGGCGGGCCTACCATGAGCGAAGTGGAATT
R-TadDE_2×STOP	GGTACCTTATTAATTAATGCTGCTCTGTGC
F-TZM-bl_T7EI	CTCTCTCGACGCAGGACTC
R-TZM-bl_T7EI	AGGCCAGGATTAACTGCGAAT
F-NGS_TZM-bl	AAGAGGCGAGGGGCGGC
R-NGS_TZM-bl	TAACTGCGAATCGTTCTAGC
F-GAGmod_1	TAGGCTAGCCGCCACCATGGGTGCCCGCGCCAGCGTGCTCA
F-GAGmod_2	GCGCCAGCGTGCTCAGCGGGGGAGAATTAGATCGATG

R-Gag_Cas_uni	TTTCTTGGGCCCCCAGGACTAATG GGAAAATTTAAAG
F-Gag_Cas_fused	ATTAGTCCTGGGGGGCCCAAGAAAA AGCGGA
R-TadDE_XcmI	TATACCATATATATATTGGTCATTAA TTAATGCTGCTCTG
F-U6p_PciI	TATATTATAAGAGGGCCTATTTCCC ATG
R-U6p_KpnI	TATAGGTACCTTCTCCAATTGTCCCT CTAG
OT1-12-F	GTGCCACAAACCATGCTTATATT
OT1-12-R	AACACTTGAGGCCATTATAGGG
OT2-12-F	TCTTAGAGGCCATTGTGGGG
OT2-12-R	TGGGGCACCATGAACTGTAT
OT3-12-F	GGGCAGTGACAATCAACAGG
OT3-12-R	TTAGGCACCCTGGGAGAATG
OT4-12-F	TACTCAGTTCCAAAGATTCT
OT4-12-R	AAATAAATTCATGTTCTTTA
OT1-9-F	GCGTACTCATAGGTTTTGGGG
OT1-9-R	CCACACTGTGCAAACATGGAG
OT2-9-F	GTCCGGCTGTAGCTTCCTTC
OT2-9-R	TGGCAGGGCAAGTTCATGTT
OT3-9-F	CAGCTGTTTCACCTCCGACT
OT3-9-R	ATCACTGGCAGCAGACACTC
OT4-9-F	CTCTGGCACACTCGCTAAGT
OT4-9-R	TGAGTCTATCGGGGGTGGAC

2.1.6 Плазмиды

Коммерческие и оригинальные плазмиды, использованные в работе, представлены в таблице 2.

Таблица 2. Плазмиды, использованные и полученные в ходе работы:

Название	**Источник**
pLenti_Pack	Коллекция плазмид лаборатории векторных систем на основе вирусных геномов ФБУН ГНЦ ВБ «Вектор» Роспотребнадзора
pLenti_Env (#8454, Addgene)	
pLenti CMV GFP Puro (#17448, Addgene)	
PX458-3×HA-dSpCas9 (#201953, Addgene)	
pSpCas9_sgRNA (#122089, Addgene)	
phMGFP	
pGFP_P2A	
pKW	Получены в данной работе
pKW _R1	
pKW _R1m	
pKW _R3	
pKW _R3m	
pMmCas12m	
pTadDE	
pMmCas12m-sgRNA1-9	
pdSpCas9-sgRNA1-9	
pMmCas12m-TadDE-sgRNA1-9	
pdSpCas9-TadDE-sgRNA1-9	
psgRNA-MmCas12m	
psgRNA-dSpCas9	
pGag-MmCas12m-TadDE	
pGag-dSpCas9-TadDE	

2.1.7 Клеточные культуры эукариот

1. HEK293T – перевиваемая линия клеток почки эмбриона человека; получена из коллекции культур клеток ФБУН ГНЦ ВБ «Вектор» Роспотребнадзора (кат. № 273, Кольцово, РФ);

2. TZM-bl (JC53-bl) – линия HeLa/CD4/CCR5; получена из коллекции культур клеток ФБУН ГНЦ ВБ «Вектор» Роспотребнадзора (кат. № 307, Кольцово, РФ).

2.1.8 Клеточные культуры прокариот

Компетентные клетки NEB Stable *E. coli* («New England Biolabs», США).

2.1.9 Программное обеспечение

Дизайн генетических конструкций проводили с использованием SnapGene v 7.2.1 («Dotmatics», США). Статистические анализы проводили с помощью программного обеспечения GraphPad Prism v 9.3.1 («GraphPad Software», США). Анализ олигонуклеотидов проводили с использованием OligoAnalyzer Tool v 3.1 («Integrated DNA Technologies», США). Оптимизацию кодонной последовательности генов проводили с использованием программного обеспечения Codon Optimization Tool («Integrated DNA Technologies», США). Для предсказания вероятных нецелевых сайтов редактирования использовали программное обеспечение Cas-OFFinder v 2.4.1 (Сеульский национальный университет, Республика Корея), Unipro UGENE v 53.0 («УНИПРО», РФ). Выравнивание прочтений, полученных в ходе NGS секвенирования, осуществляли с помощью minimap2 2.31 (режим -ax map-ont). Последующую коррекцию ошибок нанопоровых прочтений проводили инструментом medaka 2.2.2 с использованием модели «r941_min_sup_g507».

2.2 Методы

2.2.1 ПЦР

Постановку ПЦР проводили на приборе C1000 Touch Thermal Cycler («Bio-Rad», США). Общий объём реакционной смеси составлял 50 мкл. Реакционная

смесь содержала: 1 мМ раствор dNTP; по 100 пМ прямого и обратного праймера; буфер согласно рекомендациям производителя; 0,25 ед.ак. Q5 ДНК полимеразы или 0,25 ед.ак. HS-Taq ДНК-полимеразы; 20 нг матричной ДНК. При использовании Q5 ДНК-полимеразы амплификацию фрагментов ДНК проводили по температурно-временному профилю:

Параметры	**Стадия**				
	Инкубация	**25-30 циклов**			**Хранение**
		Денатурация	**Отжиг**	**Элонгация**	
Температура	98 °C	96 °C	52-60 °C	72 °C	4 °C
Время	0:45	0:15	0:20	0:15-3:00	∞

В случае использования HS-Taq ДНК-полимеразы амплификацию фрагментов ДНК проводили по следующему температурно-временному профилю:

Параметры	**Стадия**				
	Инкубация	**25-30 циклов**			**Хранение**
		Денатурация	**Отжиг**	**Элонгация**	
Температура	95 °C	95 °C	52-60 °C	72 °C	4 °C
Время	5:00	0:30	0:30	0:15-3:00	∞

При проведении всех ПЦР после крайнего цикла следовал дополнительный этап финальной элонгации длительностью 2:00 при 72°C. Некоторые параметры ПЦР (температура отжига и время элонгации) подбирали индивидуально в зависимости от характеристик праймеров и особенностей необходимого ампликона.

2.2.2 Химический синтез олигонуклеотидов ДНК

Синтез праймеров и олигонуклеотидов для сборки генов проводили на универсальном синтезаторе ДНК/РНК ASM-32 («БИОССЕТ», РФ).

Синтез осуществляли согласно протоколу 2-цианэтильного фосфорамидитного метода. Для синтеза использовали коммерческие дезоксирибонуклеозидные мономеры («LeapChem», КНР).

2.2.3 Ферментативный гидролиз ДНК

Общий объём реакционной смеси составлял 25 мкл. Общее количество ДНК в каждой пробе не превышало 5 мкг. В зависимости от фермента использовали специфический буфер, рекомендованный производителем. Количество эндонуклеаз рестрикции брали из расчета 1 ед.ак. на 1 мкг ДНК. Гидролиз проводили при 37 °С, если производителем не предусмотрено иное. Время инкубации не менее 1 ч.

2.2.4 Электрофоретическое фракционирование ДНК

Разделение фрагментов ДНК проводили методом электрофореза. Использовали 1,5-2,5 % агарозный гель, помещённый в ТАЕ-буфер. В каждый образец добавляли 1/10 объёма буфера для внесения. Для оценки длины фрагментов ДНК использовали коммерческие наборы ДНК-маркеров. Электрофорез проводили при показателях 10 В/см не менее 30 мин. Для создания электрического поля применяли источник постоянного тока. После разделения гель окрашивали в ТАЕ-буфере с содержанием бромистого этидия в концентрации около 0,5 мкг/мл. ДНК в геле детектировали по флуоресценции в проходящем ультрафиолетовом свете с длиной волны от 240 до 360 нм с использованием высокочувствительной гель-документирующей системы Amersham ImageQuant 800 («Cytiva», Швеция). При необходимости целевые фрагменты (бэнды) вырезали хирургическим скальпелем и переносили в чистые пробирки для дальнейшего использования.

2.2.5 Выделение фрагментов ДНК из агарозного геля

Выделение фрагментов ДНК из агарозного геля после ПЦР или ферментативных реакций проводили с использованием набора для очистки ДНК из агарозного геля и реакционных смесей Cleanup Mini («Евроген», РФ) согласно

рекомендациям производителя. Вкратце, в пробирку с гелем добавляли трёхкратный объём (но не менее 350 мкл) относительно массы геля связывающего раствора «S». Инкубировали при 55 °C до растворения геля. Добавляли равный объём изопропанола. Наносили образец в спин-колонку и центрифугировали 1 мин при 7000 g. Колонку промывали 700 мкл промывочного раствора и центрифугировали 1 мин при 7000 g. Проскок удаляли и повторно центрифугировали для полного удаления промывочного раствора. Для элюции ДНК на мембрану спин-колонки наносили 30 мкл элюирующего раствора, инкубировали 1 мин при комнатной температуре, помещали спин-колонку в чистую пробирку и центрифугировали 30 сек при 7000 g.

2.2.6 Очистка нуклеиновых кислот с помощью магнитных частиц

Реакционную смесь с ДНК смешивали с суспензией магнитных частиц в соотношении 1:1. Размешивали путём пипетирования. Инкубировали при комнатной температуре в течение 10 мин. Устанавливали пробирки/планшет с образцами на магнитный штатив и инкубировали в течение 5 мин. Не снимая с магнитного штатива, удаляли из образцов осветлённую надосадочную жидкость. Осадок из магнитных частиц дважды промывали 80 % этиловым спиртом, добавляя его по стенке и забирая через 30 сек, не снимая пробы с магнитного штатива. После подсушивали осадок при комнатной температуре. Снимали с магнитного штатива, сбрасывали остатки спирта на дно пробирки с помощью кратковременного центрифугирования и удаляли их. Сушили осадок до исчезновения характерного глянцевого блеска. После сушки магнитные частицы вне штатива перемешивали в 30–50 мкл свободной от нуклеаз воды путём пипетирования. Инкубировали в течение 10 мин и устанавливали пробы на магнитный штатив. Когда раствор становился прозрачным, забирали его и использовали для дальнейших приложений.

2.2.7 Клонирование фрагментов ДНК

Объём реакционной смеси составлял не более 15 мкл. В случае сборки генетических конструкций из 3 и более фрагментов в смесь вносили их

эквимолярное количество. Если предполагалось встраивание одного фрагмента в гидролизованный плазмидный вектор, то использовали пятикратный избыток фрагмента по отношению к вектору. Лигирование проводили в течение 30 мин для липких концов и не менее 1 ч для тупых концов при температуре 22 °С согласно рекомендациям производителя. На каждую реакцию использовали не менее 5 ед.ак. ДНК-лигазы бактериофага Т4 («Thermo Fisher Scientific», США). После лигирования реакционную смесь очищали путём осветления в растворе хлороформа. К реакционной смеси добавляли равный объём хлороформа, размешивали вортексированием и центрифугировали при 10000 g в течение 10 мин. Отбирали верхнюю водную фазу, содержащую плазмидную ДНК.

2.2.8 Трансформация клеток *E. coli* методом электропорации

Трансформацию проводили с использованием электрокомпетентных клеток *E. coli* линии NEB Stable («New England Biolabs», США). Трансформацию проводили на приборе Gene Pulser Xcell Microbal System («Bio-Rad», США). Вкратце, к 100 мкл компетентных клеток добавляли 5 мкл лигазной смеси, размешивали пипетированием. Наносили смесь в кювету для электропорации («Bio-Rad», США) между электродами (ширина зазора 1 мм). Помещали кювету в электропоратор. Показатели оптимизированной электропорации: напряжение – 1,9 кВ, время – 4-5 мс. После электропорации смесь извлекали и помещали в 1 мл жидкой среды LB для подращивания в течение 1 ч при температуре 30 °С. Затем клетки осаждали центрифугированием при 3000 g в течение 3 мин. Большую часть супернатанта убирали, а осадок из клеток суспендировали в остаточном объёме среды. После чего наносили на чашки Петри с агаризованной средой LB-агар путём нанесения одноразовым шпателем Дригальского. Инкубацию проводили при 30 °С не менее 18 ч.

2.2.9 ПЦР-скрининг целевых клонов-трансформантов *E. coli*

Селекцию целевых клонов-трансформантов *E. coli* на этапе конструирования плазмид проводили методом ПЦР-анализа с использованием специфических праймеров. Для этого микробиологической петлёй отбирали

отдельные колонии. Часть материала каждой колонии помещали в питательную среду, оставшуюся часть – в подготовленную ПЦР-смесь. Для ПЦР-анализа применяли HS-Taq ДНК-полимеразу. Амплификацию и анализ проводили по протоколу, указанному в пунктах 2.2.1. и 2.2.4. Позитивные клоны использовали для дальнейших приложений.

2.2.10 Выделение плазмидной ДНК

Целевые клоны культивировали в среде LB с содержанием селективного антибиотика ампициллина (50 мкг/мл) в течение 16-18 ч. Условия культивирования: 30 ºС, при постоянном помешивании 200-250 rpm в орбитальном шейкере-инкубаторе ES-20 («Biosan», Латвия). Культуру клеток осаждали центрифугированием в течение 5 мин при 3000 g. Плазмидную ДНК выделяли с использованием коммерческого набора Monarch Plasmid Miniprep Kit («New England Biolabs», США) согласно рекомендациям производителя.

2.2.11 Секвенирование ДНК

Определение нуклеотидных последовательностей проводили секвенированием по методу Сэнгера. Состав реакционной смеси: реактив BigDye Terminator v3.1 (1 мкл), 5-кратный буфер (2 мкл), праймер (100 пМ), матричная ДНК (вариативно, но не более 500 нг), деионизованная вода свободная от нуклеаз (до 10 мкл). Реакцию проводили по следующему температурно-временному профилю:

Параметры	**Стадия**				
	Инкубация	**40 циклов**			**Хранение**
		Денатурация	**Отжиг**	**Элонгация**	
Температура	96 ºС	96 ºС	57 ºС	60 ºС	4 ºС
Время	0:45	0:10	0:10	4:00	∞

Очистку реакционной смеси проводили осаждением в этаноле с использованием соосадителей. Для этого к каждому образцу добавляли 2 мкл

ацетата натрия (3 М, pH 5,2), 0,2 мкл Blue Dextran 2000 (20 мг/мл), 2,4 мкл LPA (20 мг/мл), 0,4 мкл ЭДТА (0,5 М, pH 5) и 60 мкл 96 % этанола. Центрифугировали при 2200 g в течение 45 мин. Удаляли надосадочную жидкость, к осадку добавляли 200 мкл 70 % этанола, повторно центрифугировали в течение 15 мин. После чего спирт удаляли и сушили осадок в вакуумном концентраторе «Concentrator plus» («Eppendorf», ФРГ) в течение 12 мин без нагревания. Сухой осадок растворяли в 12 мкл высокоочищенного формамида Hi-Di («Thermo Fisher Scientific», США). Непосредственный анализ реакции секвенирования проводили на автоматическом 24-х капиллярном генетическом анализаторе «ABI 3500/3500xl» («Applied Biosystems», США). Полученные хроматограммы секвенирования анализировали с использованием программного обеспечения SnapGene v 7.2.1.

2.2.12 Хранение и культивирование линий клеток эукариот

Длительное хранение линий эукариотических клеток осуществляли в криохранилище, содержащем жидкий азот. Клетки хранили в криопробирках («TPP», Швейцария) в стандартной для культуры ростовой среде без антибиотика с содержанием 40 % FBS и 10 % диметилсульфоксида (ДМСО). Итоговый объём для хранения не превышал 1 мл.

Размораживание криопробирок проводили при комнатной температуре. Далее суспензию клеток переносили в 10 мл питательной среды DMEM/F12, перемешивали пипетированием и центрифугировали при 300 g в течение 3 мин при комнатной температуре. Осадок клеток суспендировали в приемлемом объёме питательной среды и помещали в культуральные флаконы («TPP», Швейцария).

Линии клеток HEK293T и TZM-bl культивировали в среде DMEM/F12 с содержанием L-глутамина (до 0,3 мг/мл) и 10 % FBS. В питательную среду добавляли пенициллин (100 ед/мл) и стрептомицин (100 мкг/мл). Инкубацию проводили в CO_2-инкубаторе 8000 WJ («Thermo Fisher Scientific», США) при температуре 37 °C и содержании CO_2 5 %.

При достижении 80–90 % конфлюентности монослоя проводили пересев клеток. Для этого удаляли культуральную среду, монослой клеток промывали раствором DPBS. Затем на монослой наносили необходимое количество раствора Трипсина-Версена. Инкубировали в течение 10 мин при 37 °C. Клетки переносили в 15 мл или 50 мл пробирки и центрифугировали при 300 g в течение 3 мин при комнатной температуре. Удаляли надосадочную жидкость и суспендировали в необходимом объёме свежей питательной среды. Подсчёт клеток проводили с использованием автоматического счётчика клеток LUNA-II («Logos Biosystems», Республика Корея). Затем, исходя из поставленных задач, рассаживали клетки в необходимом количестве.

Перед проведением работ клетки проверяли на отсутствие контаминации микоплазмой с использованием коммерческого набора MycoReport согласно рекомендациям производителя («Евроген», РФ).

2.2.13 Трансфекция в монослой эукариотических клеток

Клетки (HEK293T или TZM-bl) рассаживали в 12-луночном планшете в количестве 180000 кл/лунку в 1 мл полной среды (DMEM/F12, 10 % FBS). Спустя 12 ч клетки трансфицировали необходимым количеством плазмидной ДНК с использованием реагента Lipofectamine 3000 («Thermo Fisher Scientific», США) в среде Opti-MEM без содержания FBS и антибиотика согласно рекомендациям производителя. Спустя 8–12 ч проводили замену среды на свежую.

2.2.14 Световая микроскопия

Для анализа монослоя клеток и фотофиксации флуоресцирующих клеток применяли систему визуализации Evos M5000 («Thermo Fisher Scientific», США). С использованием встроенных функций прибора получали фотоснимки клеточного монослоя при необходимости.

2.2.15 Проточная цитофлуорометрия

Клетки снимали с поверхности обработкой раствором Трипсина-Версена. После потери адгезии к суспензии клеток добавляли питательную среду без FBS.

Взвесь клеток в среде переносили в конические пробирки (Falcon). Затем осаждали центрифугированием при 300 g в течение 3 мин. Надосадочную жидкость удаляли. Осадок ресуспендировали в 500 мкл DPBS. Анализ доли клеток с флуоресценцией проводили с использованием проточного цитометра NovoCyte Quanteon 4025 («Agilent Technologies», США). Детекцию проводили с использованием канала FITC (экстракция 488 нм, эмиссия 530/30 нм). В качестве отрицательного контроля использовали нетрансфицированные клетки. Число анализируемых событий в каждом образце – не менее 15000. Работы по проточной цитометрии проводили совместно со старшим научным сотрудником отдела геномных исследований Еленой Юрьевной Прудниковой.

2.2.16 Выделение геномной ДНК из эукариотических клеток

Для снятия клеток с поверхности монослой обрабатывали раствором Трипсина-Версена. Клетки осаждали центрифугированием при 300 g в течение 3 мин. Максимально убирали надосадочную жидкость. Выделение геномной ДНК проводили с использованием набора РИБО-преп (ФБУН ЦНИИ Эпидемиологии Роспотребнадзора, РФ) согласно рекомендациям производителя. Чистоту препаратов геномной ДНК оценивали методом спектрофотометрии с использованием спектрофотометра NanoPhotometr NP80-Touch («Implen GmbH», ФРГ). Соотношение 260/280 в диапазоне 1,7-1,9 свидетельствовало о чистоте образцов геномной ДНК.

2.2.17 Выявление событий редактирования с использованием эндонуклеазы I бактериофага T7

Целевые ампликоны получали с использованием специфических праймеров (F-TZM-bl_T7EI и R-TZM-bl_T7EI) и очищали согласно пунктам 2.2.1., 2.2.4., 2.2.5. Перед ферментативной реакцией ампликоны подвергали денатурации путём нагрева до 95 °C в течение 10 мин с последующим восстановлением при постепенном охлаждении со скоростью 0,1 °C/сек до комнатной температуры. Затем ампликоны подвергали ферментативной обработке эндонуклеазой I бактериофага T7 («New England Biolabs», США) в

реакционной смеси на основе буфера NEBbuffer 3.1 согласно рекомендациям производителя. Фермент инактивировали нагреванием до 65 °C в течение 5 мин. Результат ферментативной обработки оценивали методом электрофореза в агарозном геле согласно пункту 2.2.4.

2.2.18 Продукция VLP

При продукции VLP за основу брали протокол, представленный в оригинальном исследовании Banskota и др. [163]. Для этого клетки HEK293T рассаживали в 10 см культуральные чашки («TPP», Швейцария) в количестве $3{,}2\times10^6$ кл/10 мл полной среды. Клетки инкубировали при 37 °C и 5 % CO_2. Спустя 15–16 ч клетки трансфицировали смесью плазмидной ДНК с использованием Lipofectamine 3000 («Thermo Fisher Scientific», США) согласно рекомендациям производителя. Для продукции VLP с MmCas12m-TadDE и dSpCas9-TadDE клетки трансфицировали смесью плазмид pLenti_Env (400 нг), pLenti_Pack (3375 нг), pGag-MmCas12m-TadDE или pGag-dSpCas9-TadDE (1125 нг), и psgRNA-MmCas12m или psgRNA-dSpCas9 (4400 нг) в среде Opti-MEM («Thermo Fisher Scientific», США) без содержания FBS и антибиотика. Спустя 8 ч среду заменяли на свежую поддерживающую среду: DMEM/F12 с добавлением 2 % FBS («БиолоТ», РФ). Концентрирование проводили по методу, описанному в оригинальной работе Borovikova и др. [164]. Для этого супернатант с VLP (около 10 мл) собирали спустя 48–50 ч после трансфекции, центрифугировали при 300 g в течение 3 мин для удаления крупных клеточных фрагментов. Дополнительно супернатант пропускали через 0,45 мкм PVDF фильтр («Millipore», США) и аликвотировали в 2 мл пробирки. Пробирки с очищенным супернатантом центрифугировали при 21000 g в течение 2,5 ч при 4 °C. Для центрифугирования использовали микроцентрифугу Eppendorf-5425 с фиксированным угловым ротором FA-24×2 («Eppendorf», ФРГ). Надосадочную жидкость удаляли. Осадок из VLP ресуспендировали в Opti-MEM среде («Thermo Fisher Scientific», США), объединяли в одну пробирку и доводили до итогового объёма в 200 мкл (объём уменьшен в ~50 раз относительно исходного

супернатанта согласно оригинальному исследованию). Аликвоту (20 мкл) каждого образца брали для вестерн-блот анализа. Оставшуюся часть каждого образца обрабатывали путём добавления 1,8 мкл ДНКазы I («SibEnzyme», РФ) и необходимого объёма 10-кратного буфера для ДНКазы I. Инкубировали при 37 °C в течение 10 мин. При необходимости образцы хранили при −80 °C.

2.2.19 Продукция лентивирусных частиц

При продукции лентивирусных частиц (LV) за основу брали протокол, представленный в оригинальном исследовании ter Brake и др. [165]. Для этого клетки HEK293T рассаживали в 10 см культуральные чашки («TPP», Швейцария) в количестве $3{,}2\times10^6$ кл/10 мл полной среды. Клетки инкубировали при 37 °C и 5 % CO_2. Спустя 15–16 ч клетки трансфицировали смесью из pLenti_Pack (1,3 пМ), pLenti_Env (0,72 пМ) и pLenti CMV GFP Puro (1,64 пМ) с использованием Lipofectamine 3000 («Thermo Fisher Scientific», США) в среде Opti-MEM («Thermo Fisher Scientific», США) без содержания FBS и антибиотика. Суммарное количество плазмидной ДНК на чашку составляло 27,8 мкг. Спустя 8 ч среду заменяли на свежую производственную среду: DMEM/F12 с добавлением 2 % FBS («БиолоТ», РФ) и инкубировали при 37 °C и 5 % CO_2 в течение 48–50 ч. Очистку и концентрирование проводили по протоколу, представленному для VLP. LV рутинной продукции служили контролем при электронной микроскопии для сравнения с VLP.

2.2.20 Электронная микроскопия

Для визуализации отпочковывания VLP при продукции использовали метод электронной микроскопии. Для этого монослой клеток покрывали раствором 4 % параформальдегида и инкубировали 24 ч при 4 °C. Затем отбирали параформальдегид и снимали культуру клеток с поверхности. Центрифугировали при 300 g в DPBS (pH 7,4) 3 раза по 15 мин. Затем клетки фиксировали в растворе OsO_4 в течение 2 ч и отмывали в DPBS 3 раза. Далее клеточные культуры дегидратировали в батарее этилового спирта возрастающей концентрации (30 %, 50 %, 70 %, 96 %) по 15 мин при каждой концентрации.

Затем материал помещали в смесь 96 % спирта и ацетона в соотношении 1:1 и инкубировали в течение 15 мин, после этого 15 мин в 100 % ацетоне. Клетки помещали в смесь ацетона и эпоксидной смолы (1:1) и оставляли на ночь для пропитывания образцов. На следующем этапе работы материал заключали в эпоксидную смолу (Embed+Araldite-502+DDSA+DMP-30) («Electron Microscopy Sciences», США) и инкубировали в термостате 2 дня: первые сутки – при 48 °С, последующие сутки – при 56 °С.

Ультратонкие срезы толщиной 40–60 нм изготавливали с помощью вольфрамовых ножей на ультрамикротоме Reichert-Jung Ultracut E («Reichert Jung», США). Полученные срезы помещали на медные сетки 100 mesh. Сетки помещали на фильтровальную бумагу и просушивали в течение суток. Далее образцы контрастировали по методике Рейнольдса: сетки помещали срезами на каплю спиртового раствора уранилацетата и выдерживали 15 мин, затем промывали в дистиллированной воде и помещали на каплю цитрата свинца на 15 мин. Визуализацию ультратонких образцов выполняли с помощью электронного микроскопа JEM-1400 («JEOL Ltd», Япония) при ускоряющем напряжении 80 кэВ с использованием программного обеспечения ITEM.

Также проводили электронную микроскопию концентрированных препаратов VLP и LV. Для этого 10 мкл каждого образца наносили на медные сетки 100 mesh. Позволяли препарату адсорбироваться на сетке в течение 10 сек, после чего удаляли излишки раствора. Затем сетку помещали на каплю цитрата свинца и инкубировали 20 сек. Сетку сушили и анализировали препараты при тех же настройках электронного микроскопа.

Физический титр определяли путём подсчета количества вирионов на медной сетке 100 mesh, локализованных в отдельной кубической ячейке. Объем каждой ячейки составляет около 27000000 мкм^3 или 0,000027 мл (~2,7 × 10^7 мкм^3 (2,7 × 10^{-5} мл), рассчитан как произведение площади ячейки сетки (0,254 × 0,254 мм^2) на толщину слоя препарата (~42 мкм)).

2.2.21 Вестерн-блот

Для выявления в образцах меченных химерных белков MmCas12m-TadDE и dSpCas9-TadDE клетки HEK293T снимали с поверхности спустя 48 ч после трансфекции и лизировали. Для этого клетки осаждали центрифугированием при 300 g в течение 3 мин. К клеточному осадку добавляли равный объём (около 200 мкл) лизирующего буфера (45 мМ Tris-HCl (pH 8,0), 160 мМ NaCl, 5 мМ ЭДТА, 1 % Triton X-100 и 1 мМ фенилметилсульфонил фторид – PMSF). Лизис проводили при 4 °C в течение 15 мин. Лизат каждого образца центрифугировали при 12000 g в течение 10 мин при 4 °C. Супернатант (по 15 мкл) смешивали с 5 мкл 4×SDS-PAGE буфера (250 мМ Tris-HCl (pH 6,8), 40 % глицерол, 8 % SDS, 4 % 2-меркаптоэтанол и 0,2 % бромфеноловый синий). Полученную смесь инкубировали при 95 °C в течение 3 мин. Для выявления химерных белков MmCas12m-TadDE и dSpCas9-TadDE в образцах VLP, сконцентрированные препараты смешивали с 4×SDS-PAGE буфером и инкубировали при 95 °C в течение 3 мин.

Разделение белков проводили с помощью электрофореза в 12 % полиакриламидном геле с додецилсульфатом натрия в буферной системе Лэммли. Электрофорез проводили при 150 В в течение 1,5 ч. Гель переносили на нитроцеллюлозную мембрану 0,45 мкм («Bio-Rad», США), используя систему полусухого переноса для блоттинга (#PB0012, «Thermo Fisher Scientific», США). Перенос проводили при напряжении 25 В и силе тока 1,3 А в течение 7 мин. Инкубировали мембрану в блокирующем буфере (0,5 % рыбьего желатина в буфере PBST) при комнатной температуре в течение 1 ч. Инкубировали мембрану в 0,5 % рыбьего желатина в буфере PBST с первичными антителами (разведение 1:5000) при комнатной температуре в течение 1 ч. Проводили три отмывки мембраны буфером PBST при комнатной температуре в течение 10–15 мин каждая. Инкубировали мембрану в 0,5 % рыбьего желатина в буфере PBST с конъюгатом пероксидазы хрена со вторичными антителами (разведение 1:5000) при комнатной температуре в течение 1 ч. Проводили три отмывки мембраны буфером PBST при комнатной температуре в течение 10–15 мин каждая. Не

допуская полного высыхания, помещали мембрану в проявляющий раствор, приготовленный непосредственно перед применением. Инкубировали 30 мин. Анализировали мембрану на наличие специфических белковых полос (бэндов) с помощью системы гель-документирования Amersham ImageQuant 800 («Cytiva», Швеция).

MmCas12m-TadDE и dSpCas9-TadDE, меченные меткой 6×His-tag, детектировали с использованием anti-6×His-tag моноклональных антител мыши (#J099B12, «BioLegend», США); бета-актин (контроль) детектировали с использованием специфичных моноклональных антител мыши (#GB15001, «Диаэм», РФ); p24 ВИЧ-1 (контроль) детектировали с использованием анти-p24 антител мыши (#MA5-44993, «Thermo Fisher Scientific», США). Антитела козы, конъюгированные с ферментом пероксидазой хрена (#1706516, «Bio-Rad», США), использовали в качестве вторичного антитела.

2.2.22 Определение трансдукционной способности клеток TZM-bl

К $3{,}5\times10^4$ клеток в 100 мкл Opti-MEM TZM-bl вносили по 0,1 мл последовательных десятикратных разведений препарата LV. Каждое разведение выполняли в трёх повторах. Поскольку LV несёт экспрессионную кассету с геном-репортером GFP, то событие заражения легко детектировать по сигналу флуоресценции. Спустя 72 ч оценивали долю GFP-позитивных клеток методом проточной цитометрии согласно пункту 2.2.15. Исходя из полученных значений рассчитывали $ИД_{50}$/0,1 мл методом Рида-Менча в модификации И.П. Ашмарина и А.А. Воробьева. Рассчитанный титр показывает количество инфекционных вирионов в 0,1 мл (100 мкл), способных к заражению половины клеток.

2.2.23 Обработка клеток TZM-bl вирусоподобными частицами

К $3{,}5\times10^4$ клеток TZM-bl в 100 мкл Opti-MEM среды вносили 50, 75 или 100 мкл каждого препарата VLP, перемешивали путём аккуратного пипетирования и культивировали в 96-луночном планшете в CO_2-инкубаторе при температуре 37 °C и содержании CO_2 5 %. Спустя 8 ч проводили замену

среды на свежую полную среду. Продолжали инкубацию клеток в прежних условиях. Спустя 72 ч после обработки проводили выделение геномной ДНК.

2.2.24 Секвенирование нового поколения (NGS)

Целевые ампликоны для NGS анализа получали с использованием специфических праймеров. Для оценки on-target эффекта использовали пару праймеров F-NGS_TZM-bl и R-NGS_TZM-bl. Для оценки off-target эффекта использовали ампликоны представленные далее в главе результаты. Ампликоны очищали согласно пунктам 2.2.1., 2.2.4., 2.2.5. Подтверждали приемлемую чистоту и концентрацию каждого образца на спектрофотометре NanoPhotometer NP80-Touch («Implen GmbH», ФРГ). Секвенирование образцов проводили независимо параллельно с использованием двух платформ: CycloneSEQ G100-ER («MGI Tech», КНР) и MiSeq («Illumina», США).

Секвенирование на платформе CycloneSEQ G100-ER

Подготовку библиотек проводили с использованием коммерческого набора реагентов CycloneSEQ 24 Barcode Library Prep Set (H940-000081) согласно рекомендациям производителя («MGI Tech», КНР). Для каждого образца (по 200 фмоль ДНК) проводили восстановление концов: 22 мкл образца, растворённого в воде, смешивали с 6 мкл буфера для восстановления (end repair buffer) концов и 2 мкл фермента для восстановления концов (end repair enzyme). Реакцию восстановления проводили при 20 °C в течение 5 мин. После чего проводили лигирование уникального баркода к каждому образцу. Для этого в реакционную смесь вносили 2,5 мкл баркода, 15 мкл лигазного буфера, 5 мкл ДНК-лигазы и 6,5 мкл свободной от нуклеаз воды. Суммарный объём реакции составил 60 мкл. Лигирование баркодов проводили при 25 °C в течение 10 мин. Затем фермент инактивировали добавлением 2 мкл 0,5 М ЭДТА. Все образцы объединяли в одну пробирку и чистили с использованием магнитных частиц. Магнитные частицы брали в соотношении 0,4 от объёма объединённых образцов. Общий пул элюировали в 32 мкл свободной от нуклеаз воды. Затем проводили лигирование адаптеров для секвенирования «barcode sequencing

adaptor». Для этого 31 мкл пула смешивали с 5 мкл «barcode sequencing adaptor», 50 мкл лигазного буфера, 20 мкл ДНК-лигазы и 94 мкл свободной от нуклеаз воды. Инкубировали при 25 °С в течение 10 мин. Реакционную смесь чистили с использованием магнитных частиц в соотношении 0,4 от объёма смеси. Готовую к секвенированию библиотеку элюировали в 42 мкл элюирующего буфера для коротких фрагментов. Измеряли концентрацию и брали суммарно 300 нг ДНК для анализа на секвенаторе. Смесь для загрузки в чип состояла из 150 мкл «working buffer», 3 мкл «incubation buffer», 60 мкл свободной от нуклеаз воды, 37 мкл библиотеки. Суммарный объём – 250 мкл. Перед загрузкой библиотеки чип помещали в секвенатор и проводили замер доступных пор. Затем переводили порт чипа в положение «On» и вносили в систему 1800 мкл «replacement buffer». Переводили порт чипа в положение «Off» и инкубировали 5 мин, после повторно измеряли количество доступных пор. Секвенирование проводили на чипах с количеством доступных пор не менее 1500. После проверки количества активных пор, вносили в чип 250 мкл смеси для загрузки, содержащую библиотеку. Пул баркодированных библиотек вносили в проточную ячейку CycloneSEQ G100-ER Flow Cell. Высокопроизводительное нанопоровое секвенирование в реальном времени выполняли на приборе CycloneSEQ G100-ER («MGI Tech», КНР). Первичный бейзколлинг и демультиплексирование данных осуществлялись автоматически в процессе запуска с использованием встроенных алгоритмов программного обеспечения прибора с генерацией файлов в формате FASTQ.

Секвенирование на платформе MiSeq

Подготовка библиотек коротких прочтений для верификации мутаций выполнялась по стандартной методике NEBNext Ultra II FS DNA Library Prep Kit («New England Biolabs», США). Качество подготовленных библиотек контролировали методом капиллярного электрофореза 4200 TapeStation («Agilent Technologies», США), а количественную оценку пула проводили флуориметрически на Qubit 4.0 («Thermo Fisher Scientific», США). Финальный

пул ДНК-библиотек денатурировали и разводили до рабочей концентрации 12 пМ. Парноконцевое секвенирование (paired-end, 2×300 п.н.) осуществляли на системе MiSeq («Illumina», США) с использованием картриджей реагентов MiSeq Reagent Kit v3.

Первичное качество прочтений, полученных с обеих платформ (Illumina и CycloneSEQ), оценивали независимо с помощью программ MultiQC 1.35 и NanoPlot 1.46.2. Фильтрацию по качеству (Q-score) и удаление адаптеров проводили с использованием Trimmomatic 0.39 для коротких прочтений и Chopper 0.12.0 для нанопоровых прочтений. Для точного картирования и идентификации точечных нуклеотидных замен (SNP) и индел-мутаций, индуцированных редакторами оснований, применяли стратегию независимого (негибридного) биоинформатического анализа. Короткие прочтения Illumina картировали на референсную последовательность мишени при помощи алгоритма BWA-MEM 7.19. Выравнивание прочтений, полученных на платформе CycloneSEQ G100-ER, осуществляли с помощью minimap2 2.31 (режим -ax map-ont). Последующую коррекцию ошибок нанопоровых прочтений проводили инструментом medaka 2.2.2 с использованием модели «r941_min_sup_g507». Выбор модели «r941_min_sup_g507» обоснован предварительными экспериментальными данными, подтвердившими её валидность для подобного типа данных. Поиск генетических вариантов и расчёт частоты целевого редактирования (модификации оснований A-в-G/C-в-T) осуществляли независимо для каждого датасета с использованием программного обеспечения BCFtools и VCFtools. Статистическую обработку данных, оценку достоверности частоты мутаций в экспериментальных группах по сравнению с интактным контролем и визуализацию распределения замен проводили с использованием среды Excel.

2.2.25 Количественная ПЦР в реальном времени

кПЦР проводили на приборе QuantStudio 5 Real-Time PCR System («Thermo Fisher Scientific», США) с использованием специфических праймеров.

Объём реакционной смеси составлял 20 мкл и включал: 10 мкл смеси БиоМастер HS-qPCR (2x) («Биолабмикс», РФ), 100 пМ каждого праймера, образец для анализа и дистиллированную воду. Реакцию проводили по следующему температурно-временному профилю:

Параметры	**Стадия**				
	Инкубация	**40 циклов**			**Хранение**
		Денатурация	**Отжиг**	**Элонгация**	
Температура	95 °C	95 °C	56 °C	72 °C	4 °C
Время	5:00	0:15	0:12	0:10	∞

2.2.26 Статистическая обработка

В экспериментах, предполагающих выявление статистической значимости, использовали по три биологических повтора (n = 3). Вычисляли среднее арифметическое значение и стандартную ошибку среднего. В случае более двух групп сравнение проводили с помощью дисперсионного анализа с post-hoc тестами Тьюки. Различия считали статистически значимыми при $p < 0,05$.

2.2.27 Моделирование пространственной структуры рибонуклеопротеиновых комплексов

Использовали модель искусственного интеллекта AlphaFold 3 («Google DeepMind» и «Isomorphic Labs», США), предназначенную для предсказания трёхмерной структуры биологических молекул и их взаимодействий. В программу загружали последовательности биополимеров и проводили анализ. Программа представляет результаты в виде пространственной модели комплекса и оценки достоверности предсказания структурных единиц комплекса между собой в виде двумерной тепловой карты (predicted aligned error plot). Для достоверности полученных результатов анализировали структуры с показателями ipTM не ниже 0,6 и pTM не ниже 0,5. ipTM является показателем

достоверности взаимодействий между компонентами (интерфейс белок-белок, белок-РНК и т.д.) в комплексе. pTM является показателем точности всей прогнозируемой структуры.

ГЛАВА 3. РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ

3.1 Создание системы скрининга эффективности связывания химерных Cas-белков и действия редакторов оснований

При наличии большого числа анализируемых sgRNA актуально иметь систему быстрого скрининга, которая помогла бы исключить низкоэффективные варианты без затрат на NGS. В связи с этим на первом этапе создана система скрининга, предназначенная для выявления взаимодействия химерных Cas-белков с целевой последовательностью ДНК и оценки действия редакторов оснований ДНК.

Принцип созданной системы заключается в инструментально детектируемом изменении фенотипа после редактирования (рис. 9). Перед репортерным геном помещается экзогенная последовательность для таргетирования тестируемого рибонуклеопротеинового комплекса. В текущей работе в качестве репортерного гена использовали ген зеленого флуоресцентного белка (mGFP), а в качестве целевой последовательности – экзогенный фрагмент провирусной ДНК ВИЧ-1. Стоит отметить, что целевая ДНК в данном случае содержит собственный старт-кодон, что позволяет инициировать синтез химерного флуоресцентного белка. При этом клонированная ДНК расположена перед геном-репортером двумя способами: в единой ОРС (pKW_R1) и со сдвигом ОРС репортерного гена (pKW_R3). Концептуально, наличие дополнительного старт-кодона в единой ОРС должно приводить к повышению экспрессии репортерного гена. И наоборот, наличие дополнительного старт-кодона, который инициирует трансляцию с альтернативной ОРС, должно приводить к снижению уровня флуоресценции, поскольку альтернативный старт-кодон инициирует синтез нефлуоресцентного белка. Сайт-направленный мутагенез старт-кодона целевой ДНК должен приводить к восстановлению исходного фенотипа, так как исключается возможность синтеза белка с альтернативного старт-кодона (pKW_R1m, pKW_R3m).

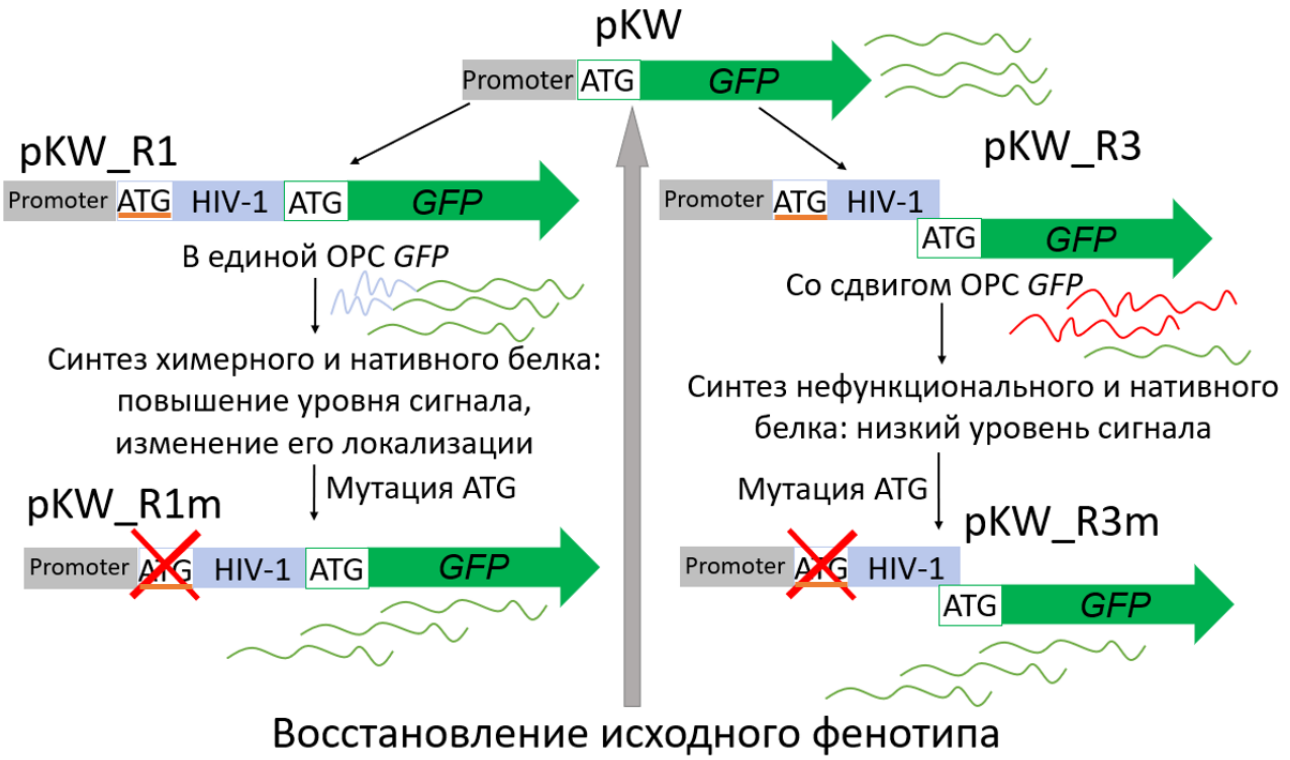


Рисунок 9. Схема принципа системы скрининга. Сайт-направленный мутагенез экзогенного старт-кодона воспроизводит действие редактора оснований и приводит к восстановлению исходного фенотипа. HIV-1 – фрагмент провирусной ДНК ВИЧ-1 (выделен синим); GFP – ген зеленого флуоресцентного белка (выделен зеленым); ATG – старт-кодон. Старт-кодон в контексте сильной последовательности Козак подчеркнут оранжевой линией. Подробное описание в тексте.

В качестве исходной генетической конструкции использовали плазмиду phMGFP, кодирующую ген *mGFP* под контролем CMV-промотора. В исходной плазмиде phMGFP репортерному гену предшествует сильная конститутивная последовательность Козак GCCACCATGG. На основе phMGFP получили новую плазмиду pKW, содержащую ОРС *GFP* со «слабой» последовательностью Козак под контролем CMV-промотора (рис. 10). Для этого в плазмиду phMGFP по сайтам *EcoRV* и *BmtI* клонировали ампликон, полученный с помощью пары праймеров F-mGFP_EcoRV и F-X_PA_7 при амплификации с этой же плазмиды. Праймер F-mGFP_EcoRV обеспечивает изменение последовательности Козак на «слабую» и дополнительно содержит сайт гидролиза *EcoRV*. Мы ослабили последовательность Козак в исходной плазмиде phMGFP путём замены С в

положении «-5» на A, C в положении «-4» на T и C в положении «-2» на T, получив мутантный вариант последовательности Козак GatAtCATG. Подобное изменение последовательности Козак должно обеспечить постепенное накопление сигнала, что позволит предотвратить избыточную экспрессию и ранний выход сигнала флуоресценции на плато. Помимо этого, дополнительно внесённый сайт эндонуклеазы рестрикции *EcoRV* позволит в дальнейшем легко клонировать последовательность ДНК для таргетирования.

На основе pKW получали ряд генетических конструкций серии pKW_R, имеющих общий принцип строения (рис. 10). В данных конструкциях перед ОРС *GFP* помещён фрагмент гена *gag* и сигнал упаковки провируса ВИЧ-1. Ампликон для клонирования данного фрагмента в pKW получали амплификацией геномной ДНК клеток линии TZM-bl с использованием пары праймеров F-partGAG_XbaI, R-partGAG_EcoRV_1 в случае pKW_R1 и F-partGAG_XbaI, R-partGAG_EcoRV_3 в случае pKW_R3.

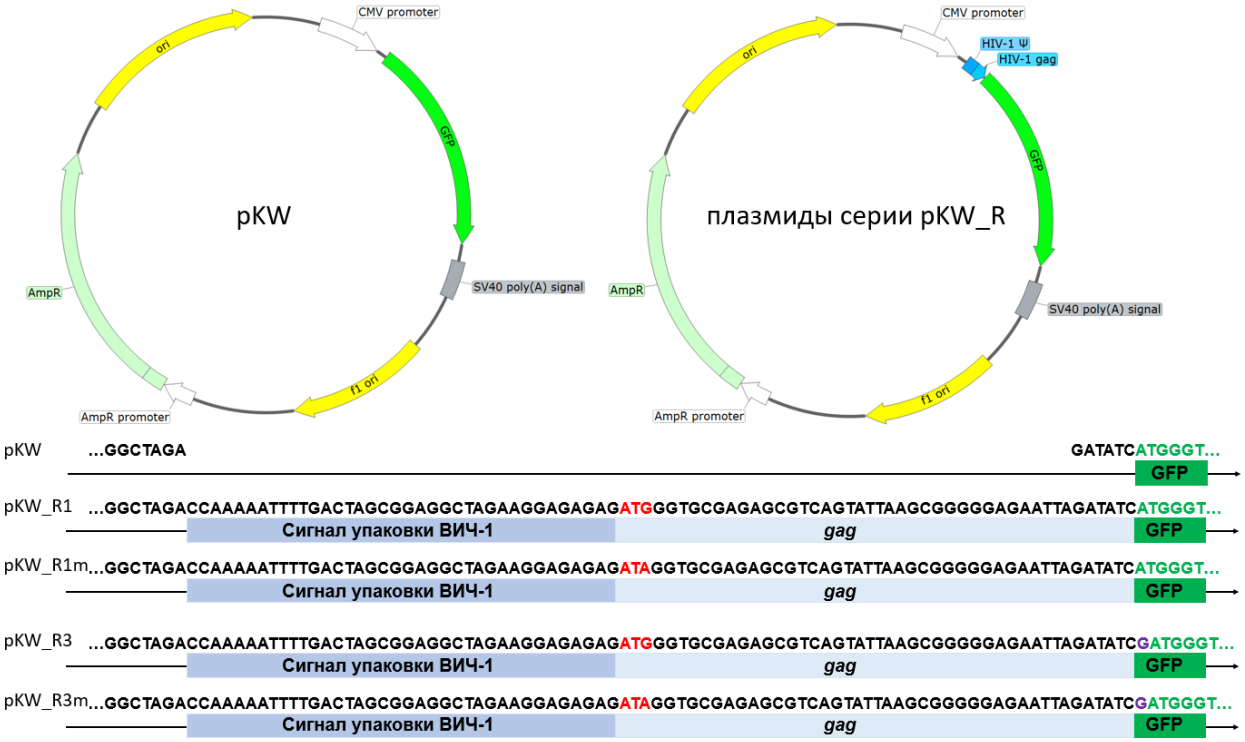


Рисунок 10. Схематичное изображение плазмиды pKW и плазмид серии pKW_R: CMV-promoter – CMV-промотор; HIV-1 Ψ – последовательность сигнала упаковки провируса ВИЧ-1; HIV-1 gag – последовательность фрагмента гена *gag* ВИЧ-1; GFP – ОРС гена *GFP*; SV40 poly(A) signal – сайт полиаденилирования вируса SV40; f1 ori – ориджин репликации бактериофага f1; AmpR promoter – промотор, управляющий экспрессией гена AmpR; AmpR –

ген бета-лактамазы, обеспечивающий резистентность к ампициллиновым антибиотикам; ori – высококопийный ориджин репликации ColE1; старт-кодон *gag* провируса ВИЧ-1 выделен красным. Нуклеотид G, приводящий к сдвигу ОРС *GFP*, выделен фиолетовым. Подробное описание в тексте. Сделано с помощью SnapGene.

Клонированный фрагмент гена *gag* представляет собой нативную последовательность провируса ВИЧ-1, кодирующую первые 14 аминокислот Gag белка, включая старт-кодон с «сильной» консенсусной последовательностью Козак. Стоит напомнить, что фрагмент провируса ВИЧ-1 клонировали в исходную плазмиду двумя способами: pKW_R1 содержит фрагмент провируса ВИЧ-1 в единой ОРС *GFP*, pKW_R3 – со сдвигом ОРС. Таким образом, трансляция белка GFP должна напрямую зависеть от фрагмента провируса ВИЧ-1 и состояния его старт-кодона. Основной принцип заключается в том, что именно данный фрагмент вирусного генома является целевым окном для редактирования. Таким образом, мутагенез старт-кодона *gag* ВИЧ-1 в конструкциях pKW_R1 и pKW_R3 должен приводить к восстановлению фенотипа, характерного при экспрессии конструкции pKW. С целью моделирования и подтверждения предположения в конструкциях pKW_R1 и pKW_R3 проводили сайт-направленный мутагенез старт-кодона *gag* с использованием пары праймеров F-dATG_GAG и R-dATG_GAG, кодирующих целевую мутацию. Таким образом получили pKW_R1m и pKW_R3m соответственно. Данные конструкции отличаются от своих исходных форм только мутацией в старт-кодоне фрагмента *gag* ВИЧ-1, что имитирует воздействие редактором оснований ДНК. Мутагенез проводили таким образом, чтобы заменить триплет ATG на ATA. Данная мутация имитирует воздействие редактора оснований на обратную цепь провируса ВИЧ-1. Корректность сборки всех плазмид системы скрининга подтверждали секвенированием по методу Сэнгера.

Всеми генетическими конструкциями по отдельности трансфицировали клетки линии HEK293T с использованием реагента Lipofectamine 3000 согласно

протоколу производителя (100 нг плазмидной ДНК/180000 клеток). Стоит отметить, что для данного эксперимента взято минимальное количество плазмидной ДНК. Поскольку на текущий момент задача заключается не во флуоресценции белка во всех клетках, а в определении влияния экзогенной последовательности на эффективность трансляции.

В случае трансфекции pKW_R1 выявлено изменение характера распределения сигнала флуоресценции. На рис. 11 видно, что флуоресцентный белок, трансляция которого велась с конструкции pKW_R1, изменил характер, характерный для мембранной локализации. Можно предположить, что это объясняется наличием слитого с GFP фрагмента Gag белка, имеющего сигнал миристоилирования, то есть ковалентного присоединения остатка миристиновой кислоты (14-углеродной насыщенной жирной кислоты) к N-концевому остатку глицина белка Gag после удаления стартового метионина. Сочетание гидрофобного миристоильного «якоря» и электростатического взаимодействия матриксного субдомена МА с отрицательно заряженными фосфолипидами обеспечивает специфическую локализацию Gag на внутренней поверхности плазматической мембраны, необходимую для нативной сборки вирусных частиц. [23, 24]. Полученный результат позволяет предположить, что трансляция инициирована с первого старт-кодона *gag* ВИЧ-1, что и вызвало локализацию сигнала предположительно на плазматической мембране.

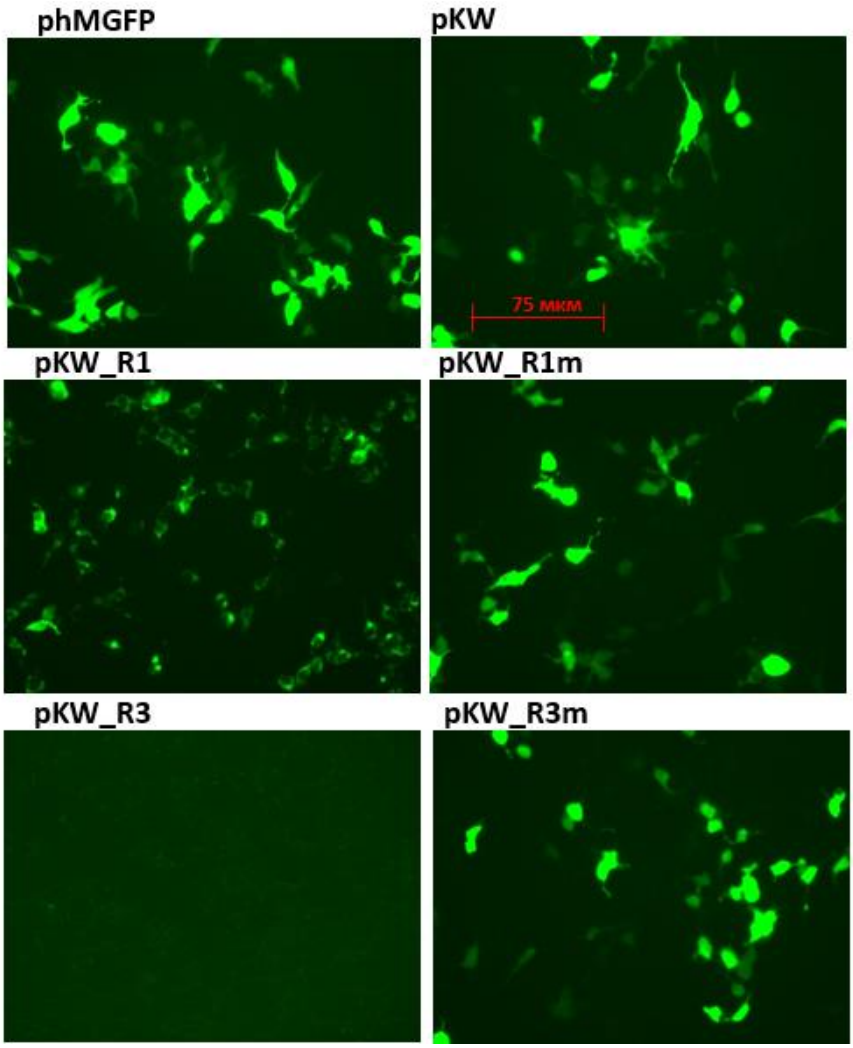


Рисунок 11. Характер распределения сигнала (локализации) флуоресценции клеток линии HEK293T, трансфицированных генетическими конструкциями разработанной системы скрининга.

Методом проточной цитометрии количественно оценивали долю клеток с флуоресценцией в каждой экспериментальной группе (рис. 12, приложение 1). Трансляция белка GFP с pKW_R1 привела к трёхкратному повышению доли клеток с флуоресценцией по сравнению с pKW: с 4 % до 12 %. Предположительно, это объясняется наличием дополнительного старт-кодона во фрагменте провируса ВИЧ-1 в единой ОРС *GFP*, которого нет в исходной плазмиде pKW. Фрагмент провируса ВИЧ-1, клонированный со сдвигом ОРС *GFP* (pKW_R3), способствовал резкому сокращению доли GFP-позитивных клеток до значения около 0,2 %. Остаточные эффекты флуоресценции в данном случае наблюдались по причине инициации трансляции со «слабого»

старт-кодона *GFP* и присутствия двух дополнительных старт-кодонов в ОРС *GFP* после первого в позициях 8-го и 14-го триплета. Стоит отметить, что второй старт-кодон ОРС *GFP* находится в хорошем контексте Козак, в то время как третий – в плохом контексте Козак, согласно основополагающей работе Мэрилин Козак [166]. Из представленных данных проточной цитометрии можно сделать вывод, что ключевую роль в инициации трансляции играет именно первый старт-кодон, локализованный в клонированном фрагменте провируса ВИЧ-1.

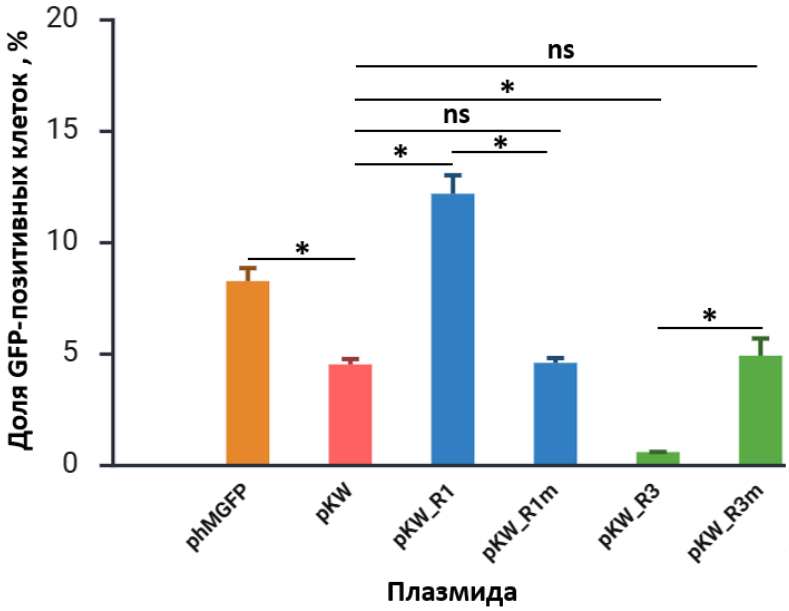


Рисунок 12. Сравнительный анализ доли флуоресцирующих клеток линии HEK293T, трансфицированных генетическими конструкциями системы скрининга. Представлены результаты проточной цитометрии. Размер выборки n = 3 (биологические повторы). *$p < 0{,}05$. ns – статистически незначимо.

Как и ожидалось, сайт-направленный мутагенез старт-кодона гена *gag* ВИЧ-1 в обеих конструкциях (pKW_R1 и pKW_R3) привел к восстановлению исходного фенотипа. Количество GFP-позитивных клеток в образцах, трансфицированных pKW_R1m и pKW_R3m, статистически не отличалось от контрольных образцов, трансфицированных pKW и составляло около 4–4,5 % (рис. 12). Таким образом, «имитация» успешного редактирования старт-кодона

gag ВИЧ-1 привела к восстановлению исходного фенотипа, характерного для pKW (ожидаемый положительный результат скрининга).

Созданные плазмиды применимы при подборе sgRNA для ДНК-специфичных Cas-белков и оценки редактирования старт-кодона *gag* ВИЧ-1. Данная система скрининга является универсальной, так как фрагмент гена *gag* ВИЧ-1 можно заменить на любую интересующую для таргетирования последовательность (*tat*, *rev*, *env*, *nef* и т.д.). При этом прослеживается возможность использования предложенной стратегии скрининга и этих плазмид для тестирования инструментов праймированного редактирования. Схожие скрининговые плазмиды с подобным принципом действия использовались и в других исследованиях для выявления эффективных sgRNA, хорошо себя зарекомендовав [9, 167–169]. Общий принцип заключается в том, что при взаимодействии с Cas-белком или редактировании происходит обнаруживаемое изменение сигнала флуоресценции.

3.2 Получение кодон-оптимизированных генов NLS-MmCas12m-NLS-TadDE и NLS-dSpCas9-NLS-TadDE

Двойные редакторы оснований применяются при необходимости множественного мутагенеза. В частности, они находят применения для нокаута генов при создании клеточных линий или трансгенезе лабораторных животных. Для апробации новых двойных редакторов оснований выбрана область старт-кодона гена *gag* провируса ВИЧ-1. Безусловно, как упоминалось в ходе обзора литературы, на текущий момент широко применяется антиретровирусная терапия, однако, нет ни одного препарата или комплексной стратегии для полного излечения от ВИЧ-инфекции. В связи с этим в рамках диссертационной работы выдвинута гипотеза о целесообразности применения двойных редакторов оснований для мутагенеза старт-кодона гена *gag* с целью нарушения морфогенеза вирионов ВИЧ-1. В данном случае задача как раз заключается во множественном мутагенезе с целью инактивировать трансляцию полипротеина Gag.

Поскольку двойной редактор оснований на базе MmCas12m получен нами впервые, то перед синтезом *de novo* искусственного химерного гена NLS-MmCas12m-NLS-TadDE рассчитывали его оптимизированную нуклеотидную последовательность с использованием программного обеспечения Codon Optimization Tool. Оптимизация гена направлена на нормализацию кодонного состава искусственного гена и исключение из последовательностей элементов, оказывающих негативное влияние на уровень экспрессии или стабильность мРНК. Олигонуклеотиды получали методом химического синтеза на базе лаборатории векторных систем на основе вирусных геномов отдела геномных исследований ГНЦ ВБ «Вектор» Роспотребнадзора с использованием оборудования для химического синтеза – универсального синтезатора ДНК/РНК ASM-32 («БИОССЕТ», РФ). Полный список олигонуклеотидов для сборки гена NLS-MmCas12m-NLS-TadDE представлен в приложениях 2 и 3.

На первом этапе проводили сборку и субклонирование генов MmCas12m и TadDE. Затем методом ПЦР с перекрывающимися концами проводили объединение обоих генов в единый химерный ген.

Для синтеза генов олигонуклеотиды смешивали в эквимолярном соотношении (по 0,1 пМ) в общий сет. Количество олигонуклеотидов в одном сете брали так, чтобы промежуточный продукт не превышал 1000 п.н. Затем брали 1 мкл из сета и использовали в качестве матрицы для постановки сборочной ПЦР. Для повышения точности синтеза и минимизации ошибок применяли высокоточную Q5 ДНК-полимеразу. Температура отжига составляла 57 °С, общее число циклов – 40. Объём реакционной смеси – 25 мкл. Затем из данной смеси после амплификации брали 1 мкл для постановки ПЦР с фланкирующими праймерами. Итоговый ПЦР-продукт анализировали в агарозном геле с последующим выделением. Поскольку ген MmCas12m превышает длину 1000 п.н., то его собирали из двух промежуточных продуктов длиной 877 и 930 п.н. путём постановки ПЦР с перекрывающимися концами. Ген TadDE длиной 501 п.н. собирали в один этап сборочной ПЦР.

Далее проводили клонирование генов в акцепторную плазмиду по сайтам эндонуклеаз рестрикции *KpnI, XbaI*. Для каждого гена анализировали по 3 клона-трансформанта. Клонов с полным отсутствием мутаций не наблюдалось. Самой распространённой мутацией были делеции, реже – однонуклеотидные замены. Один из клонов MmCas12m содержал 4 делеции и одну замену, приводящую к смене класса кодируемой аминокислоты. Второй клон – 2 делеции и 2 замены, приводящие к смене класса кодируемой аминокислоты. Третий клон – 3 делеции. В ходе секвенирования трёх клонов TadDE установлены исключительно делеции в количестве 2, 4 и 1 соответственно для каждого клона. Клоны с минимальным количеством мутаций отбирали для последующих этапов репарации мутаций. По итогам молекулярно-биологических работ получены две плазмиды pMmCas12m и pTadDE, несущие нуклеотидные последовательности генов MmCas12m и TadDE соответственно.

Данные плазмиды использовали для хранения кодирующих последовательностей генов MmCas12m и TadDE и в качестве матрицы для дальнейшей амплификации фрагментов при сборке единого гена NLS-MmCas12m-NLS-TadDE. Методом ПЦР с перекрывающимися концами и с использованием дополнительных олигонуклеотидов (F-SV40NLS_Cas12m_XmaI_1, F-SV40NLS_Cas12m_XmaI_2, F-Linker_1, F-Linker_2, 52bC12m и 53bC12m), кодирующих сигналы ядерной локализации и глицин-сериновый линкер, получали единый ген NLS-MmCas12m-NLS-TadDE.

Подробнее остановимся на составляющих искусственного гена NLS-MmCas12m-NLS-TadDE. Данный ген кодирует два сигнала ядерной локализации со стороны N- и C-конца от домена MmCas12m, что обеспечивает внутриклеточный транспорт редактора оснований. На N-конце располагается сигнал ядерной локализации большого Т-антигена вируса SV40. С-конец MmCas12m слит с сигналом ядерной локализации нуклеоплазмина. Между нуклеотидными последовательностями MmCas12m и TadDE локализована последовательность, кодирующая глицин-сериновый линкер длиной 11 амк. Данный тип линкера характеризуется повышенной гибкостью, что обеспечивает

свободное вращение и фолдинг доменов MmCas12m и TadDE относительно друг друга. Таким образом, нами впервые получен химерный кодон-оптимизированный ген NLS-MmCas12m-NLS-TadDE. Стоит отметить, что итоговая длина гена составляет 2409 п.н. Подобные компактные размеры упрощают инкапсулирование белка в вирусоподобные частицы. При этом появляется возможность закодировать ген NLS-MmCas12m-NLS-TadDE в одну экспрессионную кассету AAV.

Поскольку мы впервые описываем применение сверхкомпактного MmCas12m для создания двойного редактора оснований, то для корректного сравнения при проведении всех экспериментов мы создали двойной редактор оснований из хорошо аннотированного и широко применяемого каталитически неактивного Cas9, полученного из *Streptococcus pyogenes*, – dSpCas9 (D10A, H840A). Для этого в гене NLS-MmCas12m-NLS-TadDE фрагмент MmCas12m заменяли на нуклеотидную последовательность dSpCas9, полученную при постановке ПЦР с использованием пары праймеров F-SV40NLS_Cas9, R-nNLS_Cas9 и коммерческой плазмиды PX458-3×HA-dSpCas9 в качестве матрицы. Таким образом получали ген NLS-dSpCas9-NLS-TadDE. Для корректного сравнения выбрана полностью каталитически неактивная форма SpCas9, несущая две аминокислотные замены D10A и H840A. Выбор такого варианта для сравнения обусловлен тем, что MmCas12m не обладает нуклеазной активностью, и не способен вносить даже одноцепочечные разрывы в ДНК.

3.3 Подтверждение функциональной активности MmCas12m и dSpCas9 в комплексе с отдельными sgRNA, обеспечивающими таргетирование на область старт-кодона гена *gag* ВИЧ-1

Следующим этапов стало подтверждение способности MmCas12m dSpCas9 взаимодействовать с нуклеотидной последовательностью в области старт-кодона гена *gag*. Для этого созданы две плазмиды pMmCas12m-sgRNA и pdSpCas9-sgRNA, которые обеспечивают синтез соответствующего Cas-белка в комплексе с отдельно взятой sgRNA.

Плазмиды pMmCas12m-sgRNA и pdSpCas9-sgRNA получали на основе трансферной плазмиды pLenti CMV GFP Puro. Для этого из исходной плазмиды удаляли последовательности гена *GFP*, CMV-промотора, CMV-энхансера и регуляторных элементов лентивирусной системы с заменой на нуклеотидную последовательность гена MmCas12m или dSpCas9 под контролем PGK-промотора (рис. 13). Дополнительно в pMmCas12m-sgRNA и pdSpCas9-sgRNA клонировали последовательности специфичной tracrRNA MmCas12m и SpCas9 под контролем U6-промотора соответственно. Сборка tracrRNA MmCas12m проведена синтезом *de novo* из трёх олигонуклеотидов: F-U6p_crR_C12m_1, R-U6p_crR_C12m_2 и R-U6p_crR_C12m_3. В дальнейшем данный фрагмент соединяли с U6-промотором с использованием пары праймеров F-U6p_Cas9 и R-U6p_crR_C12m_fused в ПЦР с перекрывающимися концами. Амплификацию единой последовательности U6-промотора и tracrRNA SpCas9 проводили с использованием пары праймеров F-U6p_Cas9, R-U6p_Cas9 из коммерческой плазмиды PX458-3×HA-dSpCas9 в качестве матрицы. Два сайта эндонуклеазы рестрикции *BsaI* закладывали в pMmCas12m-sgRNA и pdSpCas9-sgRNA непосредственно рядом со специфичной tracrRNA для клонирования целевой последовательности crRNA. Полная нуклеотидная последовательность экспрессионных кассет для синтеза sgRNA представлена в приложении 4.

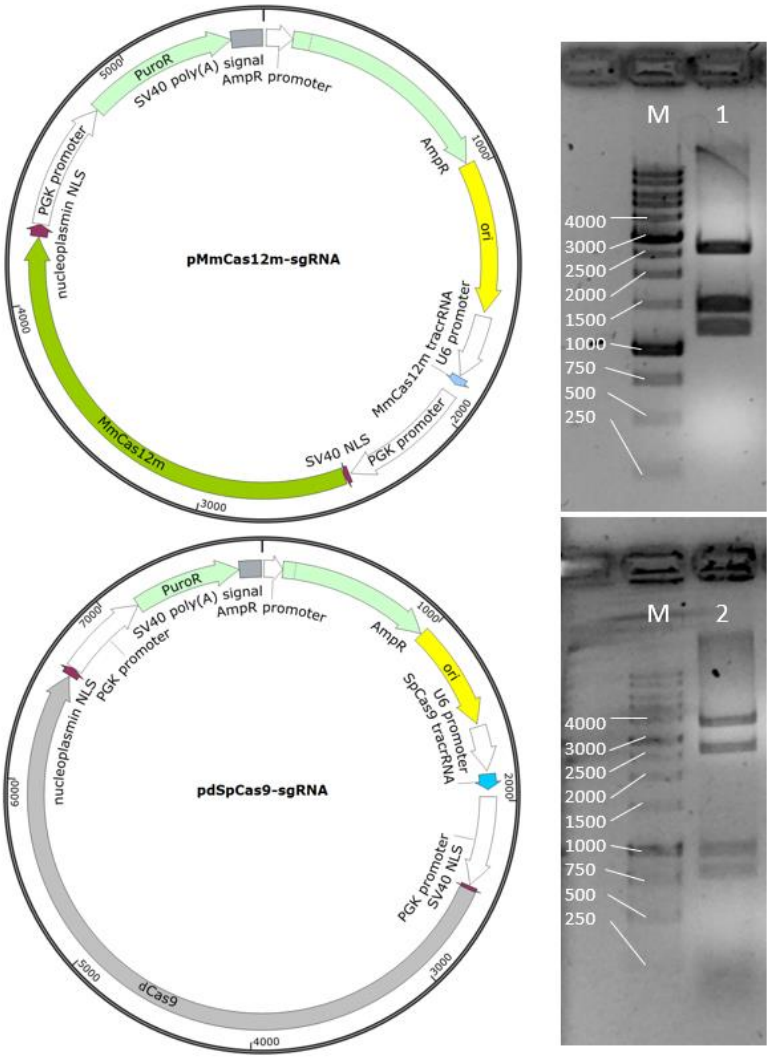


Рисунок 13. Схематичное изображение генетических конструкций pMmCas12m-sgRNA и pdSpCas9-sgRNA и их рестрикционный анализ ферментами эндонуклеаз рестрикции *HindIII* и *NheI*. M – ДНК-маркер 1 kb; 1 – pMmCas12m-sgRNA; 2 – pdSpCas9-sgRNA. U6 promoter – U6-промотор человека; tracrRNA – последовательность, кодирующая tracrRNA для MmCas12m или dSpCas9; PGK promoter – промотор фосфоглицераткиназы человека; SV40 NLS – сигнал ядерной локализации большого Т-антигена вируса SV40; MmCas12m – ген MmCas12m, dSpCas9 – ген каталитически неактивной формы SpCas9; nucleoplasmin NLS – сигнал ядерной локализации нуклеоплазмина; PuroR – ген пуромицин N-ацетилтрансферазы; SV40 poly(A) signal – сигнал полиаденилирования вируса SV40; AmpR promoter – промотор, управляющий экспрессией гена AmpR; AmpR – ген бета-лактамазы; ori – высококопийный ориджин репликации ColE. Сделано с помощью SnapGene.

На этапе дизайна эксперимента провели подбор 9 sgRNA, обеспечивающих таргетирование MmCas12m и dSpCas9 на область старт-кодона гена *gag*

(приложение 5). Поскольку в нужном для таргетинга участке провируса ВИЧ-1 оптимальные PAM для MmCas12m (5′-TTN-3′) встречаются редко, часть подобранных sgRNA примыкала к неоптимальной последовательности PAM. Из 9 sgRNA к неоптимальной последовательности PAM примыкают 6 в случае MmCas12m и 7 в случае dSpCas9. Проводили синтез олигонуклеотидов, кодирующих соответствующие sgRNA, с последующим клонированием их в акцепторные плазмиды pMmCas12m-sgRNA и pdSpCas9-sgRNA. Таким образом получали серию из 9 плазмид pMmCas12m-sgRNA1-9 и 9 плазмид pdSpCas9-sgRNA1-9, каждая из которых обеспечивает синтез рибонуклеопротеинового комплекса из MmCas12m или dSpCas9 и одной уникальной sgRNA1-9 (табл 3).

Таблица 3. Список генетических конструкций серий pMmCas12m-sgRNA*N* и pdSpCas9-sgRNA*N* с последовательностью sgRNA и указанием оптимальности PAM

Название плазмиды	**Описание**	**PAM, оптимальность (N/3)**
pMmCas12m-sgRNA1	sgRNA1 – AGCGGGGGAAAAUUGGAUAU	TTA (3/3)
pMmCas12m-sgRNA2	sgRNA2 – UAGCCUCCGCUAGUCAAAAU	TTC (3/3)
pMmCas12m-sgRNA3	sgRNA3 – GACUAGCGGAGGCUAGAAGGA	TTT (3/3)
pMmCas12m-sgRNA4	sgRNA4 – GCGGAGGCUAGAAGGAGAGA	CTA (2/3)
pMmCas12m-sgRNA5	sgRNA5 – GCACCCAUCUCUCUCCUUCU	CTC (2/3)
pMmCas12m-sgRNA6	sgRNA6 – UCUCUCCUUCUAGCCUCCGC	ATC (2/3)
pMmCas12m-sgRNA7	sgRNA7 – GGUGCGAGAGCGUCAGUAUU	ATG (2/3)
pMmCas12m-sgRNA8	sgRNA8 – CGAGAGCGUCAGUAUUGAGC	GTG (2/3)
pMmCas12m-sgRNA9	sgRNA9 – GAAGGAGAGAGAUGGGUGCG	CTA (2/3)
pdSpCas9-sgRNA1	sgRNA1 – AGCGGGGGAAAAUUGGAUAU	GGG (3/3)
pdSpCas9-sgRNA2	sgRNA2 – UAGCCUCCGCUAGUCAAAAU	TTG (2/3)
pdSpCas9-sgRNA3	sgRNA3 – GACUAGCGGAGGCUAGAAGGA	GAG (2/3)
pdSpCas9-sgRNA4	sgRNA4 – GCGGAGGCUAGAAGGAGAGA	GAT (1/3)
pdSpCas9-sgRNA5	sgRNA5 – GCACCCAUCUCUCUCCUUCU	AGC (2/3)
pdSpCas9-sgRNA6	sgRNA6 – UCUCUCCUUCUAGCCUCCGC	TAG (2/3)

pdSpCas9-sgRNA7	sgRNA7 – GGUGCGAGAGCGUCAGUAUU	AAG (2/3)
pdSpCas9-sgRNA8	sgRNA8 – CGAGAGCGUCAGUAUUGAGC	GGG (3/3)
pdSpCas9-sgRNA9	sgRNA9 – GAAGGAGAGAGAUGGGUGCG	AGA (2/3)

С использованием данных плазмидных конструкций оценена способность комплексов MmCas12m/dSpCas9 с индивидуальными sgRNA индуцировать стерическую репрессию транскрипции и последующей трансляции репортерного гена GFP (в рамках экспериментального подхода, основанного на принципах интерференции CRISPRi). Для этого каждую из плазмид pMmCas12m-sgRNA*N* (где N – порядковый номер sgRNA) котрансфицировали с описанной ранее плазмидой pKW_R1 в молярном соотношении 1:1 в клетки линии HEK293T. Для конструкций серии pdSpCas9-sgRNA*N* параллельно проводили аналогичный эксперимент. Напомним, pKW_R1 кодирует последовательность гена *GFP*, слитую с фрагментом ДНК-мишени (область старт-кодона *gag*). Концептуально, если комплекс MmCas12m-sgRNA или dSpCas9-sgRNA блокирует транскрипцию, сигнал флуоресценции будет снижаться. В таком случае, ингибирование транскрипции объясняется тем, что рибонуклеопротеиновый комплекс взаимодействует с последовательностью ДНК, локализованной между промотором и GFP-репортером, тем самым блокируя движение РНК-полимеразы II. Спустя 36 ч после котрансфекции оценивали число GFP-позитивных клеток методом проточной цитометрии (рис. 14, приложение 6).

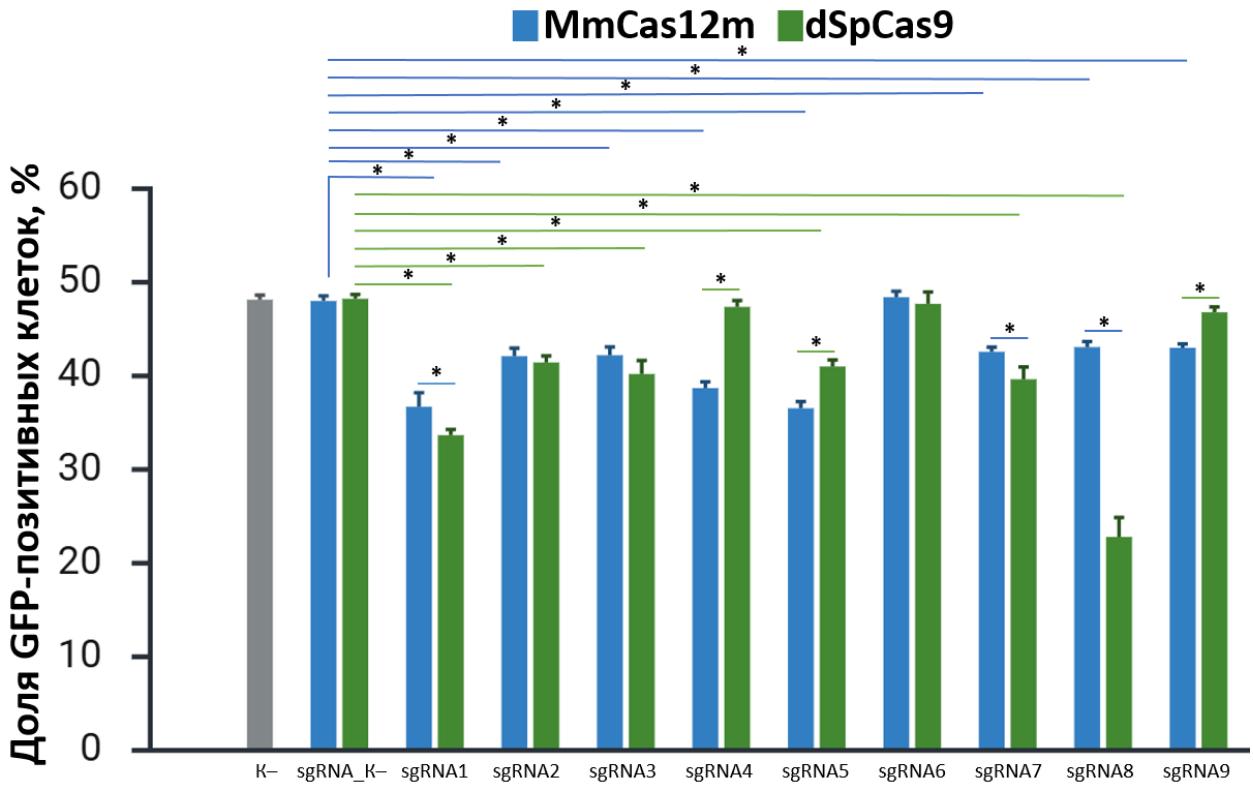


Рисунок 14. Применение системы скрининга для тестирования sgRNA при таргетировании MmCas12m и dSpCas9 на область старт-кодона гена *gag* провируса ВИЧ-1. Сравнение доли GFP-позитивных клеток при котрансфекции pKW_R1 и pMmCas12m-sgRNA*N* или pdSpCas9-sgRNA*N*. (K-) – группа клеток, трансфицированная исключительно pKW_R1; (sgRNA_K-) – группа клеток, трансфицированная pKW_R1 и pMmCas12m-sgRNA или pdSpCas9-sgRNA без sgRNA. В виде графика представлены результаты проточной цитометрии. Размер выборки n = 3 (биологические повторы). *p < 0,05. Для облегчения графика представлены сравнения только для статистически значимых различий.

Применение sgRNA5 в случае MmCas12m привело к снижению доли GFP-позитивных клеток с около 49 % (группы контроля) до около 37 %, показав наиболее выраженный результат для данного варианта Cas-белка. В группах, где рибонуклеопротеиновый комплекс MmCas12m содержал sgRNA1, 2 и 3 (оптимальные PAM) также наблюдалось снижение доли GFP-позитивных клеток до около 40–41 %. При sgRNA4 – до 39,5 % и при sgRNA7, 8 и 9 – до около 43 %. Применение sgRNA6 не обеспечило статистически значимого изменения доли GFP-позитивных клеток по сравнению с контрольной группой. На данном этапе удалось детектировать воспроизводимое ингибирование флуоресценции

для некоторых sgRNA в комплексе с MmCas12m, что обусловлено CRISPR-интерференцией. В случае dSpCas9 наибольшая эффективность интерференции, до уровня около 24 % GFP-позитивных клеток, наблюдалась при использовании sgRNA8 (оптимальный PAM). dSpCas9 в комплексе с sgRNA1 (оптимальный PAM) обеспечил снижение доли GFP-позитивных клеток до около 37–38 %. Наименьшая эффективность ингибирования флуоресценции (до около 40–42 %) наблюдалась при использовании sgRNA2, 3, 5, 7, примыкающих к неоптимальному PAM. Стоит отметить, что при использовании sgRNA4, 6, 9 не наблюдалось статистически значимого снижения доли GFP-позитивных клеток.

В качестве критерия ингибирования было выбрано наличие/отсутствие флуоресценции в клетках (бинарное деление). Измерение интегральной флуоресценции не проводилось. Первая причина заключается в том, что сигнал флуоресценции может в разы превысить максимальное значение, которое способен регистрировать прибор. Все значения выше порога насыщения регистрируются одним значением. Таким образом, даже если ингибитор условно снизил количество GFP на 30–40%, но сигнал всё ещё выше порогового значения, то интегральное значение флуоресценции не изменится вовсе. Следовательно, прибор не отобразит снижение ингибирования, даже если оно было. Вторая причина – это повышенная стабильность белка GFP. Ингибирование транскрипции останавливает синтез новых молекул мРНК, но уже синтезированный GFP способен к длительной персистенции в течение недель. Поэтому интегральная интенсивность будет падать медленно и преимущественно за счёт деления клеток, что делает этот параметр не совсем удобным для проведения быстрых скрининговых экспериментов.

Примечательно, что в ходе ингибирования эписомальной транскрипции гена-репортера MmCas12m продемонстрировал меньшую зависимость от оптимальной последовательности PAM. Важно отметить, что связывание в условиях неоптимальной последовательности PAM свойственно рекомбинантным Cas-белкам. Так, например, выявлено сохранение нуклеазной активности SpCas9 при применении sgRNA, примыкающих к неоптимальной

последовательности NAG и NGA, при условии, что PAM у SpCas9 является тринуклеотид 5′-NGG-3′ [108]. Ортологи Cas12a (AsCpf1, LbCpf1), имеющие в качестве PAM последовательность 5′-TTTV-3′, способны узнавать и такие неоптимальные PAM, как CTTV, TCTV и TTCV [109].

На данном этапе нам удалось детектировать статистически значимое снижение флуоресценции, которое, по всей видимости, вызвано взаимодействием Cas-белков с последовательностью ДНК-мишени. Однако важно проверить, способен ли MmCas12m не только ингибировать транскрипцию по механизму CRISPR-интерференции, но и таргетировать слитые с ним белковые домены, сохраняя их функциональность. В связи с этим следующим этапом стало подтверждение функциональности новой рекомбинантной дезаминазы TadDE в составе единого редактора оснований на базе MmCas12m.

3.4 Редактирование генома TZM-bl с помощью двойных редакторов оснований MmCas12m-TadDE и dSpCas9-TadDE

Для подтверждения функциональности сконструированных двойных редакторов оснований созданы две плазмиды pMmCas12m-TadDE-sgRNA и pdSpCas9-TadDE-sgRNA, которые обеспечивают синтез полного комплекса редактирования, состоящего из соответствующего редактора оснований и отдельно взятой sgRNA.

На основе плазмид с прошлого эксперимента (pMmCas12m-sgRNA и pdSpCas9-sgRNA) получали плазмиды pMmCas12m-TadDE-sgRNA и pdSpCas9-TadDE-sgRNA соответственно путём добавления нуклеотидной последовательности дезаминазы TadDE к гену каждого Cas-белка. Дополнительно на N-конец от каждого Cas белка добавлен флуоресцентный белок, который обеспечит визуальный контроль трансфекции и трансляции. При синтезе единого полипептида происходит отщепление GFP, поскольку перед ним располагается 2A-пептид тешовируса-1 (P2A), обеспечивающий саморасщепление полипротеина на этапе трансляции. Таким образом,

присутствие GFP не препятствует работе комплекса редактирования. Для получения ОРС *GFP*, слитого с P2A, применяли пару праймеров F-P2A_GFP, R-P2A_GFP и плазмиду pGFP_P2A, имеющуюся в коллекции лаборатории. Для получения последовательности дезаминазы TadDE применяли пару праймеров F-SV40NLS_TadDE, R-TadDE_2×STOP и ранее полученную плазмиду pTadDE. Три фрагмента (GFP, Cas-белок, дезаминаза TadDE) объединяли в один в ходе ПЦР с перекрывающимися концами с использованием пары праймеров F-P2A_GFP и R-TadDE_2×STOP (приложение 7) и клонировали в pMmCas12m-sgRNA и pdSpCas9-sgRNA с получением pMmCas12m-TadDE-sgRNA и pdSpCas9-TadDE-sgRNA соответственно (рис. 15). Полные нуклеотидные последовательности, кодирующие MmCas12m-TadDE и dSpCas9-TadDE, представлены в приложении 8 и 9 соответственно.

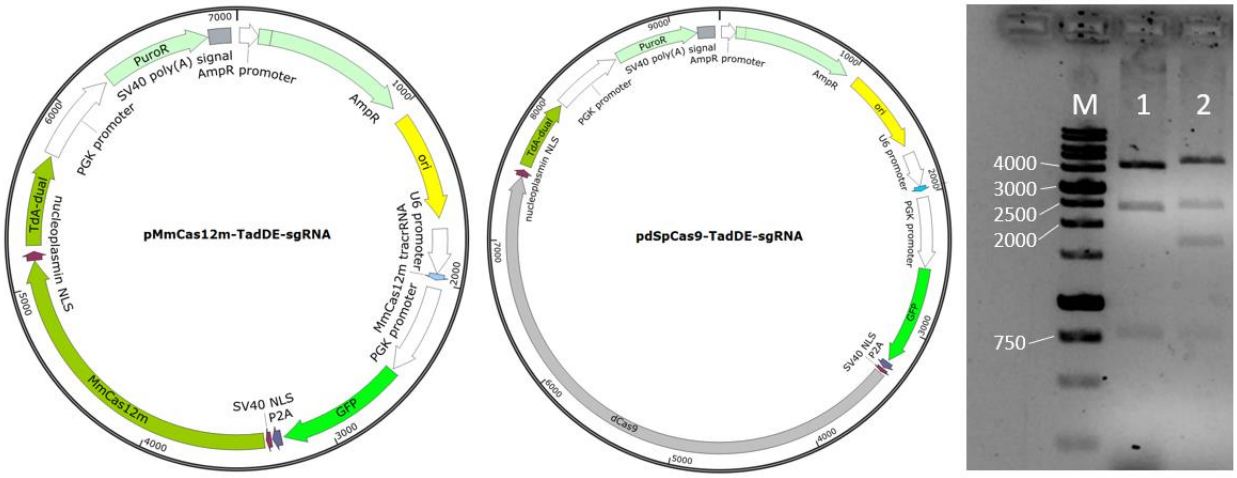


Рисунок 15. Схематичное изображение строения генетических конструкций pMmCas12m-TadDE-sgRNA и pdSpCas9-TadDE-sgRNA и их рестрикционный анализ с использованием эндонуклеаз рестрикции *EcoRI* (1), *EcoRI* и *PvuI* (2). M – ДНК-маркер 1 kb; 1 – pMmCas12m-TadDE-sgRNA; 2 – pdSpCas9-TadDE-sgRNA. U6 promoter – U6-промотор человека; tracrRNA – последовательность, кодирующая tracrRNA для MmCas12m или dSpCas9; PGK promoter – промотор фосфоглицераткиназы человека; GFP – ген *GFP*; P2A – 2A-пептид тешовируса-1; SV40 NLS – сигнал ядерной локализации большого Т-антигена вируса SV40; MmCas12m – ген MmCas12m, dSpCas9 – каталитически неактивная форма SpCas9; nucleoplasmin NLS SV40 – сигнал ядерной локализации нуклеоплазмина; TdA-dual – ген рекомбинантной дезаминазы TadDE; PuroR – ген пуромицин N-ацетилтрансферазы; SV40 poly(A) signal – сигнал полиаденилирования вируса SV40; AmpR promoter – промотор,

управляющий экспрессией гена AmpR; AmpR – ген бета-лактамазы; ori – высококопийный ориджин репликации ColE. Сделано с помощью SnapGene.

По аналогии с предыдущей парой акцепторных плазмид, в плазмиды pMmCas12m-TadDE-sgRNA и pdSpCas9-TadDE-sgRNA клонировали последовательности, ответственные за синтез 9 различных sgRNA. Таким образом получен набор из 18 плазмид – pMmCas12m-TadDE-sgRNA1-9 и pdSpCas9-TadDE-sgRNA1-9. Плазмиды серии pMmCas12m-TadDE-sgRNA*N* обеспечивают синтез двойного редактора оснований MmCas12m-TadDE в комплексе с одной из sgRNA. Плазмиды серии pdSpCas9-TadDE-sgRNA*N* – двойного редактора оснований dSpCas9-TadDE в комплексе с одной из sgRNA. Несмотря на верифицированные ранее эффективные sgRNA, принято решение использовать все варианты. Поскольку применение 9 различных sgRNA позволит в условиях оптимального и неоптимального PAM выявить особенности редактирования, в частности, ширину окна редактирования и пик эффективности редактирования. Это важно учитывать при работе, так как данные показатели специфичны для каждого редактора оснований.

В качестве модели *in vitro* выбрали клетки линии TZM-bl. Геном клеток TZM-bl содержит гены-репортеры люциферазы и β-галактозидазы под контролем 5’-LTR ВИЧ-1. Данная интеграция получена в ходе лентивирусной трансдукции. При этом интеграция содержит область старт-кодона гена *gag*, искомую для таргетирования. Присутствие целевого локуса для редактирования дополнительно подтверждали секвенированием по методу Сэнгера. Концептуально, наличие фрагмента провирусной ДНК ВИЧ-1 безопасно моделирует состояние генома инфицированной клетки после интеграции.

Монослой клеток TZM-bl отдельно трансфицировали каждой из плазмид pMmCas12m-TadDE-sgRNA1-9 и pdSpCas9-TadDE-sgRNA1-9 (1000 нг плазмидной ДНК/180000 кл) с использованием Lipofectamine 3000 согласно рекомендациям производителя. В качестве отрицательных контролей выступали группы клеток, трансфицированные акцепторными плазмидами pMmCas12m-

TadDE-sgRNA и pdSpCas9-TadDE-sgRNA без sgRNA. Спустя 8 ч трансфекционную среду заменяли на свежую среду с добавлением пуромицина для отбора исключительно трансфицированных клеток. Такой подход позволяет оценить эффективность редактирования только в клетках, трансфицированных плазмидой. Спустя 72 ч удаляли мёртвые клетки путём трёхкратного промыва клеточного монослоя раствором DPBS и проводили выделение геномной ДНК живых флуоресцирующих клеток. Флуоресценция в данном случае указывает на синтез GFP, экспрессируемый из единого транскрипта с редактором оснований через P2A-пептид. Перед NGS анализом проводили предварительное выявление событий редактирования с использованием эндонуклеазы I бактериофага T7 (T7EI) исключительно для образцов, где редактирование опосредовано новым редактором оснований MmCas12m-TadDE. Для этого ПЦР-продукты каждой экспериментальной группы по исследуемой sgRNA подвергали денатурации и отжигу, а затем обрабатывали эндонуклеазой T7EI. При наличии в анализируемом гетеродуплексе хотя бы одного неспаренного основания (мисматча), возникшего в результате направленного редактирования, эндонуклеаза T7EI специфически распознает данный сайт и внесёт двухцепочечный разрыв в ДНК в непосредственной близости от него. Таким образом, исходный ампликон будет гидролизован на два меньших фрагмента, что можно детектировать при проведении стандартного электрофореза в агарозном геле. В ходе проведенного анализа события редактирования детектировали для всех ранее верифицированных sgRNA в рамках нашей системы скрининга, исключая sgRNA6 (рис. 16).

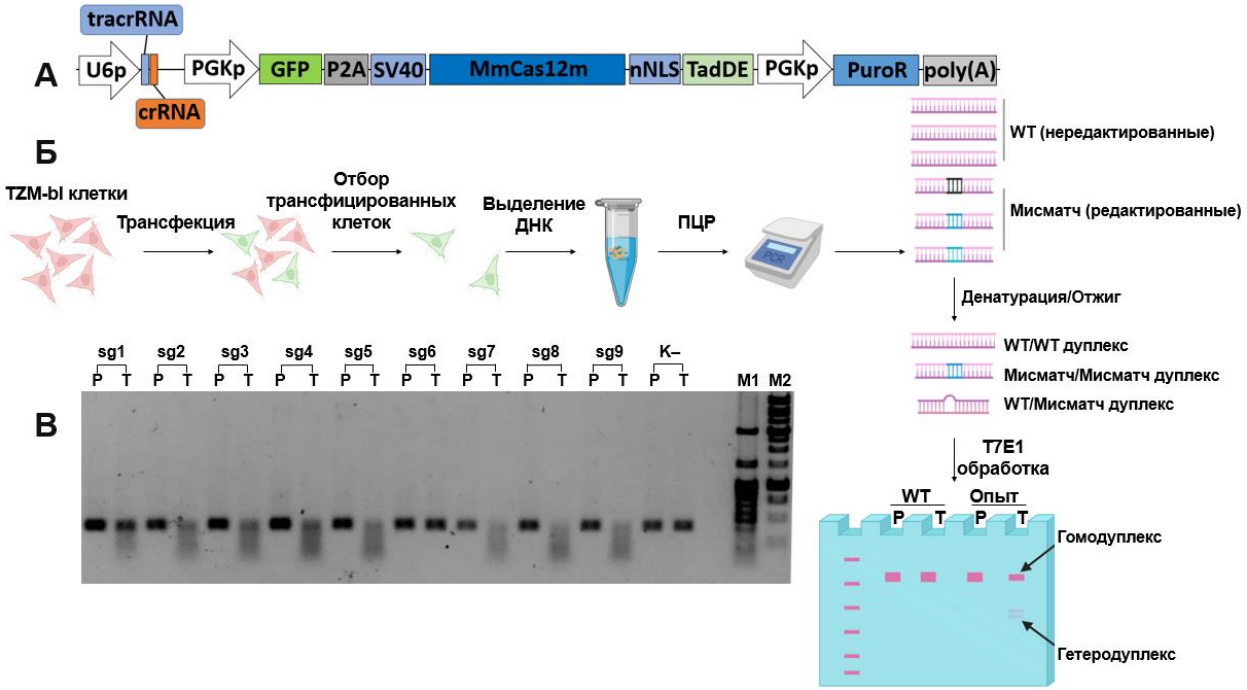


Рисунок 16. Выявление событий редактирования с использованием эндонуклеазы I бактериофага T7. Схематичное изображение экспрессионной кассеты, переносимой генетическими конструкциями серии pMmCas12m-TadDE-sgRNA (А); Схема шагов эксперимента по выявлению событий редактирования с использованием эндонуклеазы I бактериофага T7 (Б); Результат выявления событий редактирования, представленный на гель-электрофорезе (В). P – ПЦР-продукт без обработки T7EI, T – ПЦР-продукт, обработанный T7EI, sg1-9 – опытные образцы, редактированные с использованием sgRNA1-sgRNA9 соответственно, WT – отрицательный контроль без редактирования. M1 – 100+ bp ДНК-маркер, M2 – 1kb ДНК-маркер. U6p – U6-промотор человека; tracrRNA и crRNA – последовательности, кодирующая единую sgRNA для MmCas12m; PGKp – промотор фосфоглицераткиназы человека; GFP – ген *GFP*; P2A – 2A-пептид тешовируса-1; SV40 – сигнал ядерной локализации большого T-антигена вируса SV40; MmCas12m – ген MmCas12m, nNLS – сигнал ядерной локализации нуклеоплазмина; TadDE – ген рекомбинантной дезаминазы TadDE; PuroR – ген пуромицин N-ацетилтрансферазы; poly(A) – сигнал полиаденилирования вируса SV40. Подробное описание в тексте. Сделано с помощью BioRender.

Примечательно, что ампликоны гидролизованы с разной эффективностью в зависимости от применяемой sgRNA. Большая часть исходного ампликона гидролизована в образце после редактирования с использованием MmCas12m-TadDE-sgRNA5. Напомним, sgRNA5 в рамках системы скрининга стала самой

эффективной для MmCas12m. При этом в образцах, где редактирование проводили с использованием MmCas12m-TadDE-sgRNA7, 8, 9, большая часть исходного ампликона также гидролизована, хотя в рамках системы скрининга они показывали меньшую эффективность. В образцах, где редактирование проводили с использованием MmCas12m-TadDE-sgRNA1, 2, 3, 4, интенсивность бэнда исходного ампликона снизилась незначительно. При этом данные sgRNA показали меньшую эффективность по сравнению с sgRNA7, 8, 9 в рамках нашей системы скрининга. События редактирования не удалось детектировать в образце, где редактирование проводили с использованием MmCas12m-TadDE-sgRNA6. Поскольку детекция неспаренных оснований с использованием T7EI является качественным методом, а не количественным, и призвана лишь выявить события редактирования, то определить точную эффективность не представляется возможным. С целью точного определения эффективности и выявления целевых мутаций в старт-кодоне гена *gag* в зависимости от применяемой sgRNA и редактора оснований проводили NGS анализ ампликонов каждой экспериментальной группы. В ходе анализа NGS данных выявили, что искомый мутагенез старт-кодона *gag* наблюдается только в случае применения sgRNA5 для обоих редакторов оснований (рис. 17).

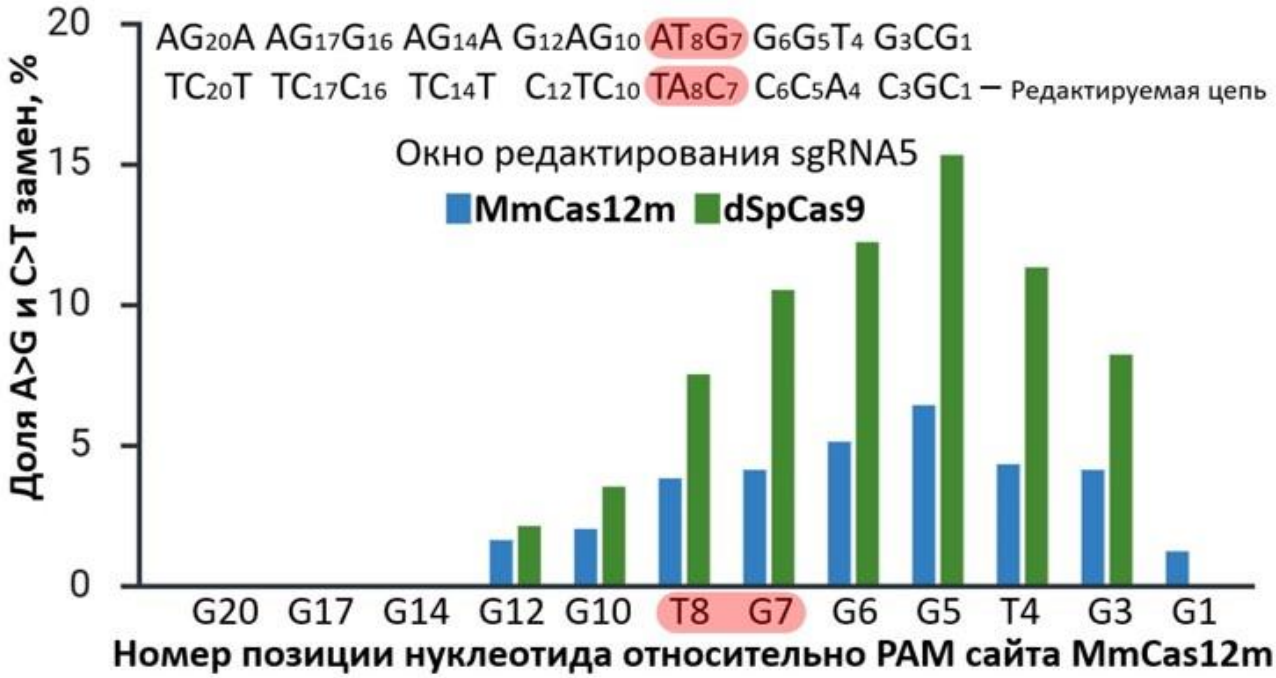


Рисунок 17. Эффективность редактирования с использованием sgRNA5. Редактирование опосредовано трансфекцией плазмид pMmCas12m-TadDE-sgRNA5 (выделено синим) и pdSpCas9-TadDE-sgRNA5 (выделено зеленым). Старт-кодон гена *gag* выделен красным. Размер выборки n = 3 (биологические повторы).

Поскольку sgRNA5 комплементарна прямой цепи, то при взаимодействии с редактором оснований в одноцепочечное состояние переходит обратная цепь геномной ДНК. Следовательно, редактированию в случае использования sgRNA5 подвергается именно обратная (матричная) цепь, по которой синтезируется молекула РНК вируса. В таком случае синтезированная молекула РНК будет нести мутации в старт-кодоне. Старт-кодону ATG комплементарна последовательность 5′-CAT-3′ в редактируемой цепи. Мутации в позиции нуклеотида C_7 выявлены в 4,25 % прочтений, а в позиции A_8 – в 3,78 % в случае MmCas12m-TadDE. Суммарная эффективность мутагенеза старт-кодона составила 8,03 %. Стоит отметить, что в результате трансфекции плазмид, кодирующих MmCas12m-TadDE, эффективность редактирования не превышала 10–11 % для всех рассмотренных sgRNA. При использовании dSpCas9-TadDE суммарная эффективность мутагенеза старт-кодона составила 17,9 % (7,4 % – в позиции A_8, 10,5 % – в позиции C_7).

Стоит обсудить полученные результаты для каждой sgRNA и отметить ряд отличий между редактированием в случае применения каждого редактора оснований (рис. 18).

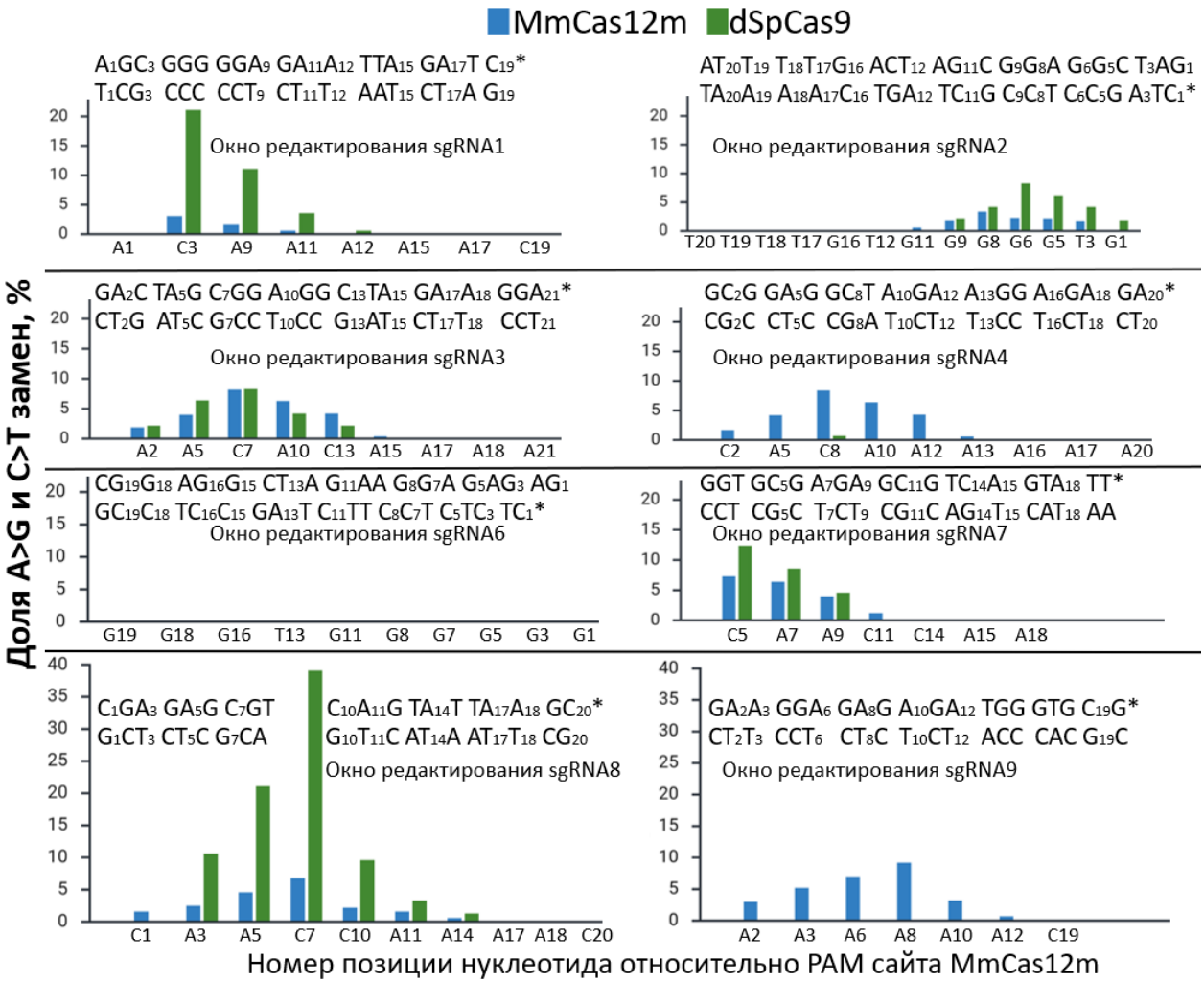


Рисунок 18. Эффективность редактирования с использованием sgRNA 1-4, 6-9. Редактирование опосредовано трансфекцией соответствующих плазмид pMmCas12m-TadDE-sgRNA*N* и pdSpCas9-TadDE-sgRNA*N*. Знаком “*” отмечена редактируемая цепь ДНК. Размер выборки n = 3 (биологические повторы).

Анализируя результаты, можно отметить, что эффективность редактирования при использовании MmCas12m-TadDE ниже, чем при dSpCas9-TadDE, и достигает уровня около 10–11 %. В то время как dSpCas9 обеспечил эффективность редактирования в некоторых случаях до 40 %. При этом замечено, что dSpCas9 оказался более чувствителен к неоптимальному PAM, чем MmCas12m. Видно, что MmCas12m-TadDE в комплексе с sgRNA4, 7, 8 и 9 (примыкают к неоптимальному PAM) обеспечивал эффективность редактирования на уровне sgRNA1-3 (примыкают к оптимальному PAM) или даже лучше, что видно при сравнении, например, sgRNA3 и 9. Таким образом

MmCas12m-TadDE обеспечил редактирование во всех случаях кроме sgRNA6. При этом, применение sgRNA6 также не обеспечило редактирование в комплексе с dSpCas9-TadDE. Эффективность редактирования с использованием dSpCas9-TadDE была значительно выше только в тех случаях, где dSpCas9 примыкал к оптимальному PAM – sgRNA1 (пик эффективности около 20 %) и sgRNA8 (около 40 %). А при условиях неоптимального PAM наблюдали потерю до около 75 % от максимально достигнутой эффективности или её полную утрату. Так, например, dSpCas9-TadDE-sgRNA2 (неоптимальный PAM) обеспечил эффективность редактирования около 8 % в позиции 6 протоспейсера. При использовании sgRNA3 MmCas12m-TadDE и dSpCas9-TadDE продемонстрировали схожую эффективность редактирования с пиком около 9 %. Заметим, что dSpCas9-TadDE не проявил способности к редактированию в комплексе с sgRNA4 и sgRNA9, хотя MmCas12m обеспечил редактирование в данных случаях со средней для своего уровня эффективностью (пик до 8–10 %). И это при условии, что, например, sgRNA9 примыкает к неоптимальным PAM (2/3) для обоих Cas-белков. Видно, что при неоптимальном PAM способность редактирования dSpCas9-TadDE снижается в 2–3 раза (sgRNA2, 3, 7) или вовсе утрачивается (sgRNA4, 9). Напротив, не наблюдалось значимых различий в эффективности редактирования при использовании MmCas12m-TadDE в условиях оптимальных и некоторых неоптимальных PAM. Стоит отметить, что оба редактора оснований имеют схожие размеры окон редактирования, а пик эффективности редактирования приходится на позиции 5–6 протоспейсера. Доля прочтений, в которых наблюдали сразу оба типа замен (A>G, C>T), составляла менее 1 % от общего числа. Таким образом локус после редактирования содержит преимущественно один тип мутации: A>G или C>T, что характерно для двойных редакторов оснований.

В продолжение обсуждения приведём данные недавних исследований, в рамках которых применяли различные Cas-белки для создания редакторов оснований. В одной из таких работ для таргетирования редактора оснований аденина применяли другой компактный вариант Cas12m (GoCas12m под

контролем CAG-промотора), который обеспечил эффективность редактирования на уровне 9–12 % в клетках линии HEK293T при условии оптимального PAM. Эффективность согласуется с полученной нами. Но данные значения ниже эффективности редактирования, которую смог обеспечить dSpCas9-TadDE (до около 20–40 %). Стоит отметить, что редакторы на основе никазной формы SpCas9 в условиях неоптимального PAM демонстрируют схожую с нашей эффективность редактирования в неоптимальных условиях [167]. Другим широко применяемым вариантом является SaCas9. Согласно некоторым исследованиям редакторы на основе никазной формы SaCas9 обеспечивают эффективность редактирования на уровне 20–58 % в условиях оптимального PAM [170, 171]. Недавно открытый вариант SauriCas9 (никазная форма) также применялся для редактирования оснований, показав эффективность около 40 % в условиях оптимального PAM [172]. Cas12a (Cpf1) является представителем Cas12-белков, которым требуется T-богатый PAM (TTTV). В связи с этим Cas12a изначально использовался для расширения целей редактирования [173]. Впрочем, Cas12a остается крупноразмерным эффектором (1307 амк), чьи размеры близки к параметрам SpCas9 (1368 амк), что накладывает аналогичные ограничения на его доставку. При этом эффективность редактирования с использованием никазных форм Cas12a-редакторов оснований составляет около 40 % в условиях оптимального PAM [168]. В некоторых случаях удавалось достичь повышенной эффективности до 70 % путём модификации домена дезаминазы или сверхэкспрессии гена редактора оснований [174].

В ходе проведённых экспериментов мы применяли дезаминазу TadDE, которая ранее получена группой Neugebauer и др. путём направленной эволюции [10]. В ходе оригинального исследования эффективность редактирования с использованием данной дезаминазы составляла до 73–80 % при оптимальных PAM. Впрочем, Neugebauer и др. применили никазную форму SpCas9. Стоит отметить, что при создании редакторов оснований преимущественно применяют никазные формы Cas-белков для повышения эффективности редактирования. При взаимодействии с таким редактором оснований дезаминирование

сопровождается внесением одноцепочечного разрыва в нередактированную цепь, что вызывает её репарацию по матрице редактированной цепи, опережая механизм репарации дезаминированного основания. Однако одноцепочечный разрыв, который вносят никазные формы Cas-белков, может быть сопряжен с повышением генотоксичности. Обнаружено, что при использовании никазных форм Cas-белков в некоторых случаях окна редактирования содержат вставки или удаления нескольких нуклеотидов (инделы) с частотой до 6–17 % [175, 176]. Таким образом, немалая часть отредактированных локусов (на фоне пика эффективности 40–80 %) содержит нежелательные индел-мутации. Следовательно, использование dCas-белков и изначально каталитически неактивного MmCas12m является взвешенным решением, направленным на снижение генотоксичности и исключение инделов при редактировании.

3.5 Создание плазмид для продукции вирусоподобных частиц VLP-MmCas12m-TadDE и VLP-dSpCas9-TadDE

Выбор платформы для доставки инструмента геномного редактирования требует особого внимания. Применение распространённых платформ на основе вирусных векторов, переносящих экспрессионную кассету, сопряжено с проникновением чужеродной ДНК в ядро клетки. При использовании AAV векторов экспрессионные кассеты способны к длительному персистированию в эписомальной форме, тогда как применение лентивирусных векторов приводит к их стабильной интеграции в геном клетки-хозяина. Оба сценария сопряжены с рисками: пролонгированная экспрессия экзогенного нуклеазного комплекса может вызывать клеточную токсичность, а случайный характер интеграции лентивирусной провирусной ДНК повышает риск инсерционного мутагенеза, что существенно снижает профиль безопасности данных систем доставки.

Неконтролируемое во времени или избыточное присутствие эффекторного комплекса в клетке нарушает кинетику взаимодействия с ДНК, что повышает риск нецелевой активности (off-target). В поиске наиболее безопасной и

оптимальной платформы мы выбрали стратегию использования вирусоподобных частиц – VLP (рис. 19).

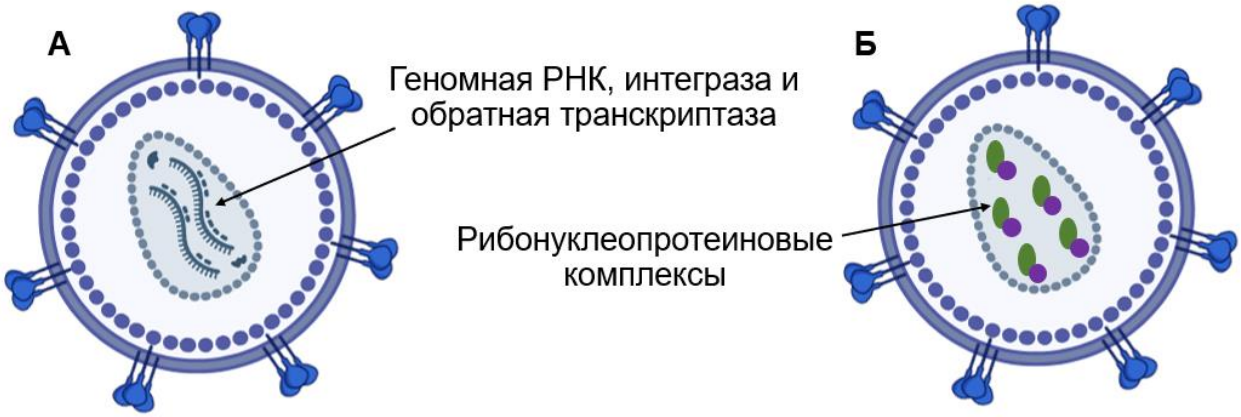


Рисунок 19. Схематичное изображение строения лентивирусной частицы и созданной на её основе вирусоподобной частицы – VLP. Подробное описание в тексте. Сделано с помощью BioRender.

По морфологии VLP напоминают вирионы нативных вирусов, на основе которых получены, но являются репликативно-дефектными, поскольку не содержат геномных нуклеиновых кислот. Это не только исключает риск инсерционного мутагенеза, но и делает невозможной рекомбинацию генома с вирусами «дикого типа». В ряде недавних научных публикаций представлены результаты успешного применения VLP для доставки Cas-белков, редакторов оснований и даже крупных молекул праймированного редактирования [163, 177–179]. С течением времени белковые комплексы, принесённые в клетку посредством VLP, подвергаются деградации ферментами клетки-хозяина. Таким образом, после редактирования не остаётся остатков самого инструмента редактирования. Это позволяет существенно снизить нецелевой эффект. За основу оригинальных VLP мы выбрали репликативно-дефектный лентивирус, псевдотипированный поверхностным гликопротеином G вируса везикулярного стоматита (VSV). Поверхностный гликопротеин G VSV обеспечит широкий тропизм, что оправдано на начальном этапе создания и тестирования новых VLP. Перспектива дальнейших работ включает применение различных гликопротеинов для изменения тропизма VLP в зависимости от поставленной

задачи. Для продукции оригинальных VLP-MmCas12m-TadDE и VLP-dSpCas9-TadDE получали два набора генетических конструкций. Ниже приведено их описание и процесс создания (рис. 20).

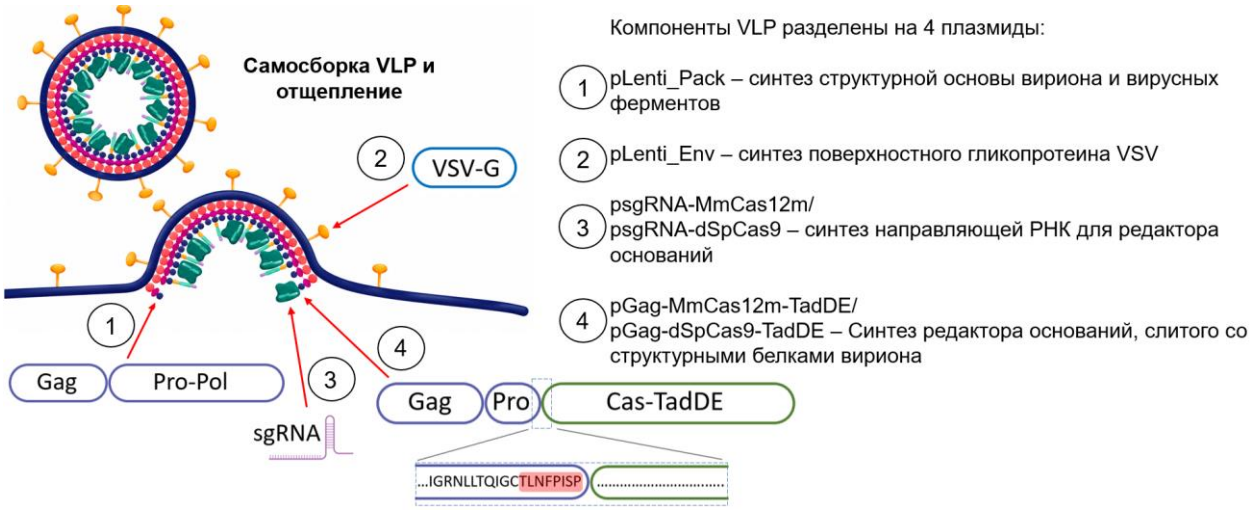


Рисунок 20. Схематичное изображение компонентов VLP с указанием плазмид, которые обеспечивают их синтез. VSV-G – поверхностный гликопротеин VSV; Gag – структурный полипротеин ВИЧ-1; Pro-Pol – полипротеин протеазы, обратной транскриптазы и интегразы ВИЧ-1; Pro – вирусная протеаза; Cas-TadDE – редактор оснований. Красным выделена аминокислотная последовательность сайта гидролиза вирусной протеазы, обеспечивающая отщепление редактора оснований Cas-TadDE от полипротеина Gag-Pro.

Плазмидные векторы pGag-MmCas12m-TadDE и pGag-dSpCas9-TadDE обеспечивают экспрессию химерных белков, формирующих структурный каркас собираемых VLP. Их экспрессионная кассета кодирует единый полипептид-прекурсор, состоящий из последовательно расположенных доменов структурных белков ВИЧ-1 MA, CA, NC, а также неструктурных белков вирусной протеазы и редактора оснований Cas-TadDE, слитого с меткой 6×His-tag. Мы сохранили сайты гидролиза вирусной протеазы между доменами, что обеспечит созревание VLP и отщепление редакторов оснований от структурных компонентов вириона. Полная нуклеотидная последовательность для кодирования данных полипептидов-прекурсоров представлена в приложениях 10 и 11. Последовательности получали объединением двух фрагментов в ходе постановки ПЦР с перекрывающимися концами (приложение

12). Для амплификации фрагмента, кодирующего структурные белки и вирусную протеазу, использовали три праймера F-GAGmod_1, F-GAGmod_2, R-Gag_Cas_uni и плазмиду pLenti_Pack в качестве матрицы. Поскольку нуклеотидная последовательность 5’-области гена *gag* в плазмиде pLenti_Pack идентична таковой с провирусом ВИЧ-1, то при котрансфекции может происходить ингибирование продукции VLP редакторами оснований. Чтобы этого избежать, нуклеотидную последовательность 5’-области гена *gag* в плазмидах pGag-MmCas12m-TadDE, pGag-dSpCas9-TadDE и pLenti_Pack изменяли с сохранением кодируемой аминокислотной последовательности. Для амплификации фрагмента, кодирующего двойной редактор оснований MmCas12m-TadDE или dSpCas9-TadDE, использовали универсальную пару праймеров F-Gag_Cas_fused, R-TadDE_XcmI и плазмиду pMmCas12m-TadDE-sgRNA или pdSpCas9-TadDE-sgRNA в качестве матрицы соответственно. Полученные фрагменты клонировали в плазмиду pLenti_Pack по сайтам эндонуклеаз рестрикции *SphI* и *XcmI* с удалением исходной последовательности. Таким образом получали генетические конструкции pGag-MmCas12m-TadDE и pGag-dSpCas9-TadDE.

Конструкции psgRNA-MmCas12m и psgRNA-dSpCas9 кодируют специфичную sgRNA под контролем U6-промотора. Конструкции psgRNA-MmCas12m и psgRNA-dSpCas9 получали на основе плазмиды pSpCas9_sgRNA путём замены исходной последовательности на целевую (приложение 4) по сайтам эндонуклеаз рестрикции *PciI* и *KpnI*. Для амплификации фрагментов для клонирования использовали пару праймеров F-U6p_PciI, R-U6p_KpnI и плазмиду pMmCas12m-TadDE-sgRNA или pdSpCas9-TadDE-sgRNA в качестве матрицы соответственно.

Помимо созданных нами оригинальных конструкций при продукции VLP использовали коммерческие плазмиды pLenti_Pack и pLenti_Env. pLenti_Pack является стандартной упаковочной плазмидой второго поколения, используемой в молекулярной биологии для создания лентивирусных векторов. pLenti_Pack обеспечивает синтез исходных компонентов лентивирусной частицы, входящих

в полипептид-прекурсор Gag/Pro-Pol. Присутствие нативного полипептида-прекурсора Gag-Pro-Pol к химерному Gag-Pro-Cas в молярном соотношении составляет 2:1 является стандартом для сохранения пространственной конформации вириона [163]. pLenti_Env, как указывалось ранее, применяется для псевдотипирования вирусных векторов поверхностным гликопротеином G VSV.

3.6 Продукция, морфологическое описание и подтверждение функциональности вирусоподобных частиц VLP-MmCas12m-TadDE и VLP-dSpCas9-TadDE

Перед применением VLP для редактирования генома клеток линии TZM-bl мы поставили задачу предварительного подтверждения их способности проникать в клетки *in vitro* и сохранять функциональность доставляемых редакторов оснований.

Продукцию VLP-MmCas12m-TadDE и VLP-dSpCas9-TadDE, переносящих двойной редактор оснований MmCas12m-TadDE-sgRNA5 и dSpCas9-TadDE-sgRNA5 соответственно, проводили в клетках линии HEK293T. Выбор sgRNA5 для продукции VLP обусловлен тем, что именно данная sgRNA обеспечивала мутагенез старт-кодона *gag* в ходе предыдущих экспериментов. В качестве контроля параллельно проводили продукцию лентивирусных частиц (LV) в клетках HEK293T.

В ходе продукции VLP и LV компоненты вирионов самособираются на плазматической мембране с дальнейшим отпочковыванием. Для выявления событий отпочковывания спустя 48 ч инкубации каждый из экспериментальных образцов анализировали методом электронной микроскопии. Удалось выявить учащённые события отпочковывания везикулярных структур с поверхности клеток, в которых производилась продукция VLP и LV (рис. 21). В клетках отрицательного контроля, которые трансфицировали смесью из pLenti_Env и pLenti CMV GFP Puro, учащённых событий отпочковывания не обнаружено.

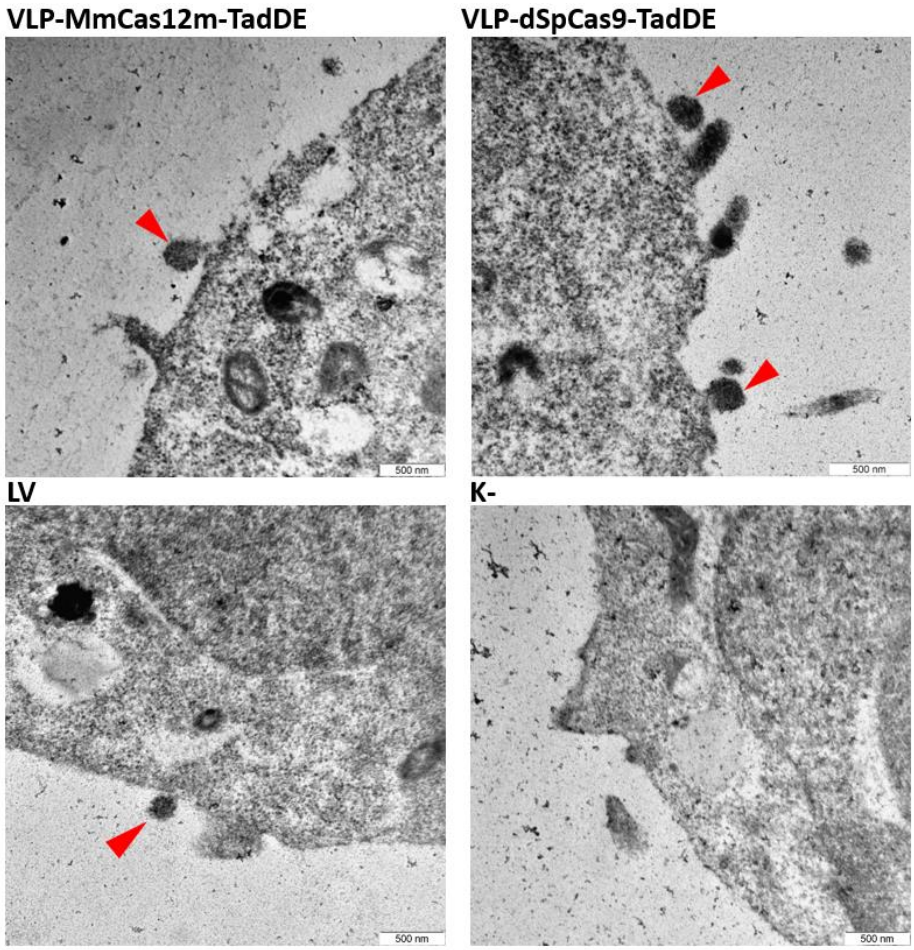


Рисунок 21. Отпочковывание везикулярных структур (указаны стрелками) от плазматической мембраны клеток-продуцентов HEK293T при продукции соответствующих VLP и LV. (K-) – клетки отрицательного контроля, трансфицированные исключительно pLenti_Env и pLenti CMV GFP. 48 ч продукции. Получено методом электронной микроскопии.

Сбор и концентрирование VLP и LV проводили спустя 48–72 ч после трансфекции. Кондиционированную среду центрифугировали для удаления крупного клеточного дебриса. Очистку и концентрирование проводили согласно протоколу, представленному в работе Borovikova и др., и подробно описанному в разделе «Методы». Сконцентрированные VLP анализировали методом электронной микроскопии для подтверждения успешного морфогенеза VLP и LV. В ходе анализа методом электронной микроскопии установлено, что полученные VLP и LV рутинной продукции имеют не только сопоставимые между собой размеры около 130–150 нм, но и схожий рисунок оптической плотности вириона, что указывает на общность строения (рис. 22). При этом

размеры VLP (130–150 нм), несколько превышающие таковые для нативных вирионов (~100–120 нм), согласуются с наличием белковой нагрузки (MmCas12m-TadDE (~90 кДа) и dSpCas9-TadDE (~185 кДа)).

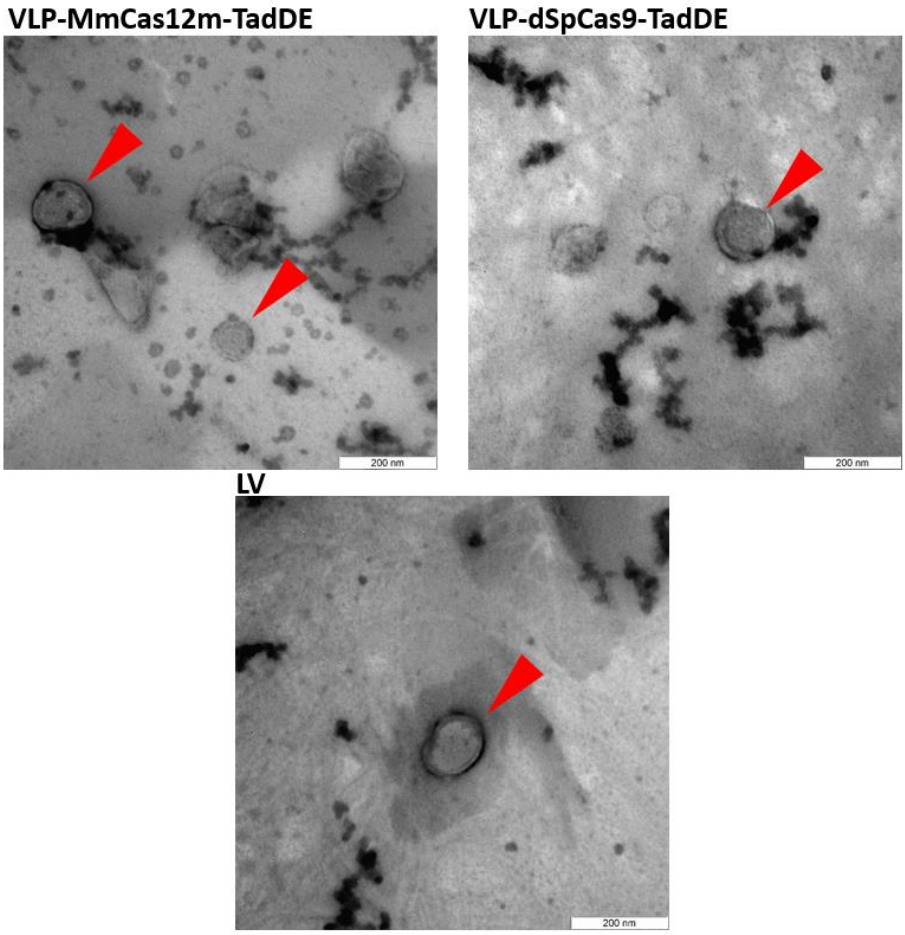


Рисунок 22. Результаты электронной микроскопии сконцентрированных препаратов VLP и LV. Стрелками указаны вирионы VLP и LV.

Поскольку VLP не содержат геномных нуклеиновых кислот, то определить титр путём кПЦР не представляется возможным. В связи с этим мы оценили физический титр VLP и LV. Для каждого препарата VLP и LV подсчитывали количество вирионов в 5 случайных ячейках (табл. 4). Вычисляли среднее значение и исходя из полученных данных пересчитывали количество вирионов в 1 мл препарата.

Таблица 4. Количество вирионов каждого препарата в ячейках медной сетки 100 mesh.

	Наименование препарата		
Номер ячейки	**VLP-MmCas12m-TadDE-sgRNA5**	**VLP-dSpCas9-TadDE-sgRNA5**	**LV**
1	192	201	245
2	148	152	198
3	153	129	174
4	161	143	159
5	197	175	163

Приблизительный физический титр VLP-MmCas12m-TadDE-sgRNA5 составил не менее 6,25(±0,27)×10^6 частиц/мл, VLP-dSpCas9-TadDE-sgRNA5 – не менее 5,94(±0,35)×10^6 частиц/мл, а для LV – не менее 6,9×10^6 частиц/мл. Стоит отметить, что данные показатели физического титра являются приемлемыми при продукции лентивирусных частиц [180, 181].

Результаты электронной микроскопии подтверждают успешный морфогенез, но не дают представлений об инкапсулировании молекул редакторов оснований в VLP. Для точного подтверждения присутствия полноразмерных молекул каждого редактора оснований проводили вестерн-блот с использованием специфичных антител. Вестерн-блот проводили для образцов лизата клеток-продуцентов и очищенных препаратов VLP. Для детекции редакторов оснований применяли антитела к метке 6×His-tag, которая, напомним, присоединена к С-концу каждого редактора оснований. В ходе анализа удалось выявить полосы задержки (бэнды), соответствующие приблизительной молекулярной массе ~ 90 кДа в образцах с MmCas12m-TadDE и ~ 185 кДа в образцах с dSpCas9-TadDE (рис. 23А). Бэнды располагаются вблизи значений, которые соответствуют рассчётным молекулярным массам полноразмерных молекул новых редакторов оснований MmCas12m-TadDE (88,6 кДа) и dSpCas9-TadDE (181,6 кДа). Детекция бета-актина и белка p24 ВИЧ-1 с использованием специфичных антител является стандартным контролем и способом нормализации. В ходе анализа бэнды бета-актина (42 кДа)

и белка p24 ВИЧ-1 (24 кДа) соответствовали ожидаемым молекулярным массам данных белков. Исходные необработанные снимки результатов вестерн-блота с указанием маркера молекулярного веса представлены для обозрения в приложении 13.

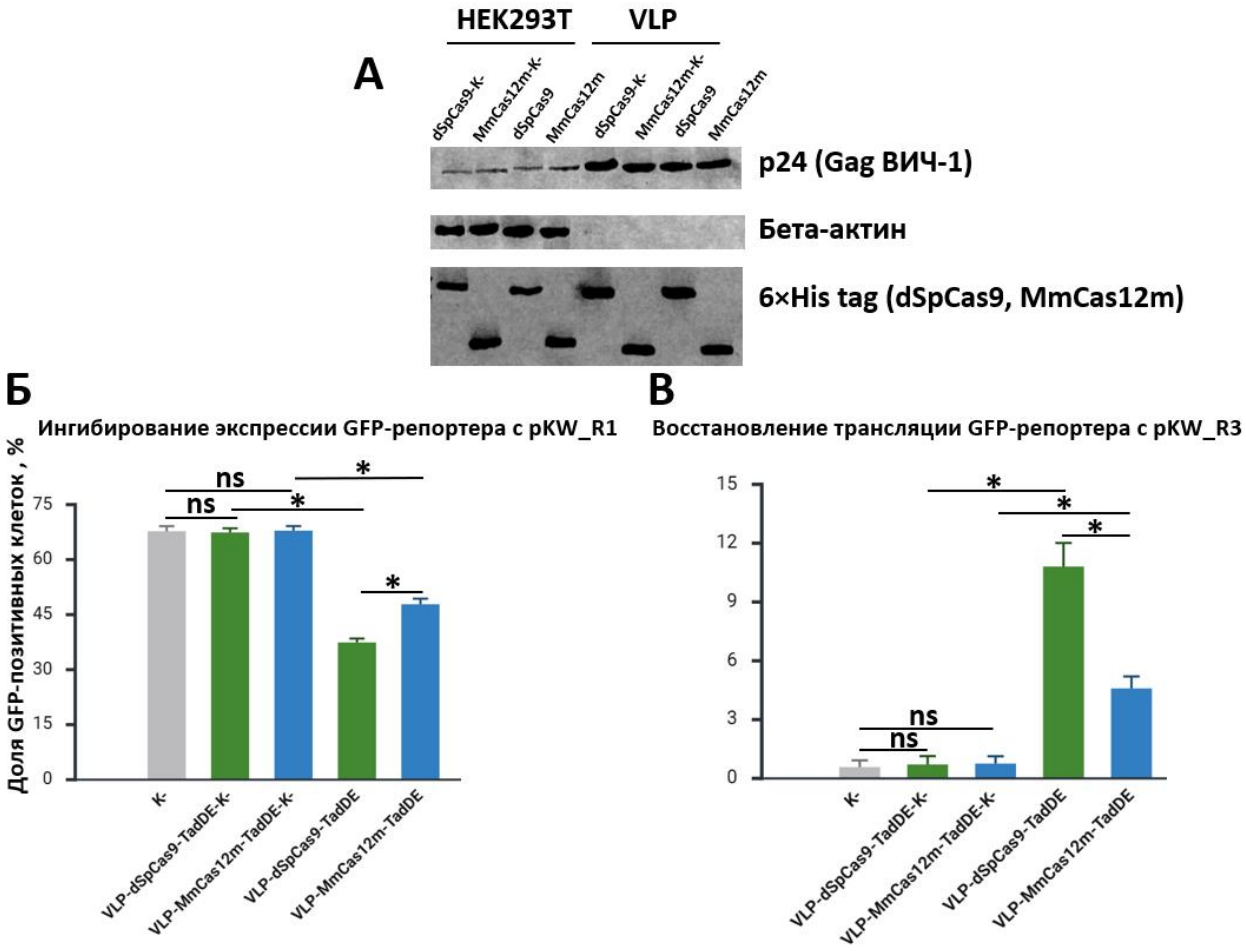


Рисунок 23. Результаты вестерн-блота, подтверждающие наличие молекул редакторов оснований в клетках-продуцентов HEK293T и очищенных препаратах VLP (А); Сравнение доли GFP-позитивных клеток при трансфекции pKW_R1 и обработке VLP (Б); Сравнение доли GFP-позитивных клеток при трансфекции pKW_R3 и обработке VLP (В). (К-) – клетки, исключительно трансфицированные pKW_R1 или pKW_R3; VLP-dSpCas9-TadDE-K- и VLP-MmCas12m-TadDE-K- – клетки, трансфицированные pKW_R1 или pKW_R3 и обработанные соответствующими VLP без sgRNA; VLP-dSpCas9-TadDE и VLP-MmCas12m-TadDE – клетки, трансфицированные pKW_R1 или pKW_R3 и обработанные соответствующими VLP с sgRNA5. Представлены результаты проточной цитометрии. Размер выборки n = 3 (биологические повторы). *$p < 0,05$. ns – статистически незначимо.

Следующим этапом стало подтверждение способности VLP проникать в клетки *in vitro* и доставлять функционально активный рибонуклеопротеиновый комплекс. Концептуально, эксперимент включает два этапа: трансфекция клеток плазмидами скрининга и обработка препаратами VLP, которые переносят редакторы оснований. При успешном проникновении в клетки редакторы оснований должны вызвать изменение фенотипа путём редактирования трансфицированных ранее плазмид (приложение 14).

Для этого клетки линии HEK293T ($3{,}5\times10^4$ кл/лунку) по отдельности трансфицировали плазмидой pKW_R1 или pKW_R3 (по 100 нг). Спустя 8 ч клетки обрабатывали сконцентрированными препаратами VLP-MmCas12m-TadDE или VLP-dSpCas9-TadDE (объём препарата – 50 мкл). Спустя 40 ч после обработки проводили оценку доли GFP-позитивных клеток методом проточной цитометрии. Обработка препаратами привела к устойчивому снижению доли GFP-позитивных клеток HEK293T, трансфицированных pKW_R1, с около 68 % (контрольные группы) до около 47 % в случае VLP-MmCas12m-TadDE и до около 37 % в случае VLP-dSpCas9-TadDE (рис. 23Б, приложение 15). Мы также наблюдали устойчивое увеличение доли GFP-позитивных клеток HEK293T, трансфицированных pKW_R3, с около 1 % (контрольные группы) до около 4,5 % в случае воздействия VLP-MmCas12m-TadDE и до около 11 % в случае VLP-dSpCas9-TadDE (рис. 23В, приложение 16).

Таким образом, ингибирование транскрипции с pKW_R1 свидетельствует об успешном взаимодействии редакторов оснований с целевой ДНК. Однако данный эффект может быть обусловлен CRISPR-интерференцией, а не целевым мутагенезом старт-кодона. Функциональность дезаминазы TadDE подтверждается увеличением доли GFP-позитивных клеток HEK293T, трансфицированных pKW_R3, после обработки VLP. Напомним, исходно pKW_R3 обеспечивает крайне низкий уровень флуоресценции, поскольку перед GFP-репортером располагается фрагмент гена *gag*, приводящий к сдвигу ОРС GFP. Трансляция функционального GFP, и как следствие повышение доли GFP-позитивных клеток, возможны только в случае мутагенеза старт-кодона гена *gag*,

что и наблюдалось в ходе эксперимента. Контрольные группы содержали клетки исключительно после трансфекции pKW_R1 или pKW_R3 (K-), и клетки после трансфекции pKW_R1 или pKW_R3 и обработки соответствующими VLP без sgRNA (VLP-Cas9-K- и VLP-Cas12m-K-). Такой подход позволяет оценить, влияет обработка VLP (без sgRNA) на уровень экспрессии гена-репортера. Впрочем, подобного эффекта не наблюдали, что свидетельствует о том, что изменение доли GFP-позитивных клеток вызвано именно воздействием исследуемых редакторов оснований на область редактирования. Таким образом, применяя систему скрининга и оригинальные VLP, удалось подтвердить функциональность системы доставки и редактирования. Следующим шагом стало количественное определение эффективности редактирования целевого локуса в геноме клеток линии TZM-bl.

3.7 Редактирование генома TZM-bl с использованием двойных редакторов оснований MmCas12m-TadDE и dSpCas9-TadDE, доставляемых вирусоподобными частицами

На текущем этапе задача заключалась в редактировании генома клеток линии TZM-bl с использованием VLP, доставляющих редакторы оснований MmCas12m-TadDE или dSpCas9-TadDE. Важной составляющей являлось выявление способности к редактированию генома, который обладает сложной пространственной конформацией по сравнению с плазмидной ДНК.

Предварительно оценили эффективность трансдукции клеток TZM-bl с использованием сконцентрированного препарата LV, полученного на предыдущих этапах. Титрование проводили заражением клеток ($3{,}5\times10^4$ кл/лунку) 0,1 мл последовательных десятикратных разведений (от 10^{-1} до 10^{-9}). Методом Рида-Менча установили, что титр LV, приводящий к 50 % эффективности трансдукции, составил около $4{,}1\times10^4$ ($ИД_{50}$/0,1 ФОЕ/мл).

Редактирование генома клеток линии TZM-bl проводили путём обработки монослоя сконцентрированными препаратами VLP. Для оценки дозозависимого эффекта использовали фиксированные объёмы препаратов в трёх вариантах: 50,

75 и 100 мкл. Ввиду отсутствия инфекционности получаемых вирусоподобных частиц, их дозирование проводилось на основе финального объёма очищенного концентрата. Исходя из результатов физического титра и трансдукционной способности клеток TZM-bl, объём препарата в 100 мкл должен приводить к «заражению» более половины $3{,}5\times10^4$ клеток. Спустя 72 ч инкубации после обработки клетки собирали и выделяли геномную ДНК. Геномную ДНК использовали в качестве матрицы для амплификации редактируемого локуса. Полученные ампликоны анализировали методом NGS для количественного определения событий редактирования. При использовании VLP-MmCas12m-TadDE целевые мутации выявляли в позиции C_7 в 2,35–3,24 % всех прочтений, и в позиции A_8 в 2,1–2,91 % прочтений. Суммарная эффективность мутагенеза старт-кодона гена *gag* с использованием VLP-MmCas12m-TadDE-sgRNA5 составила 4,35–6,15 % в зависимости от вносимого объёма препарата (рис. 24).

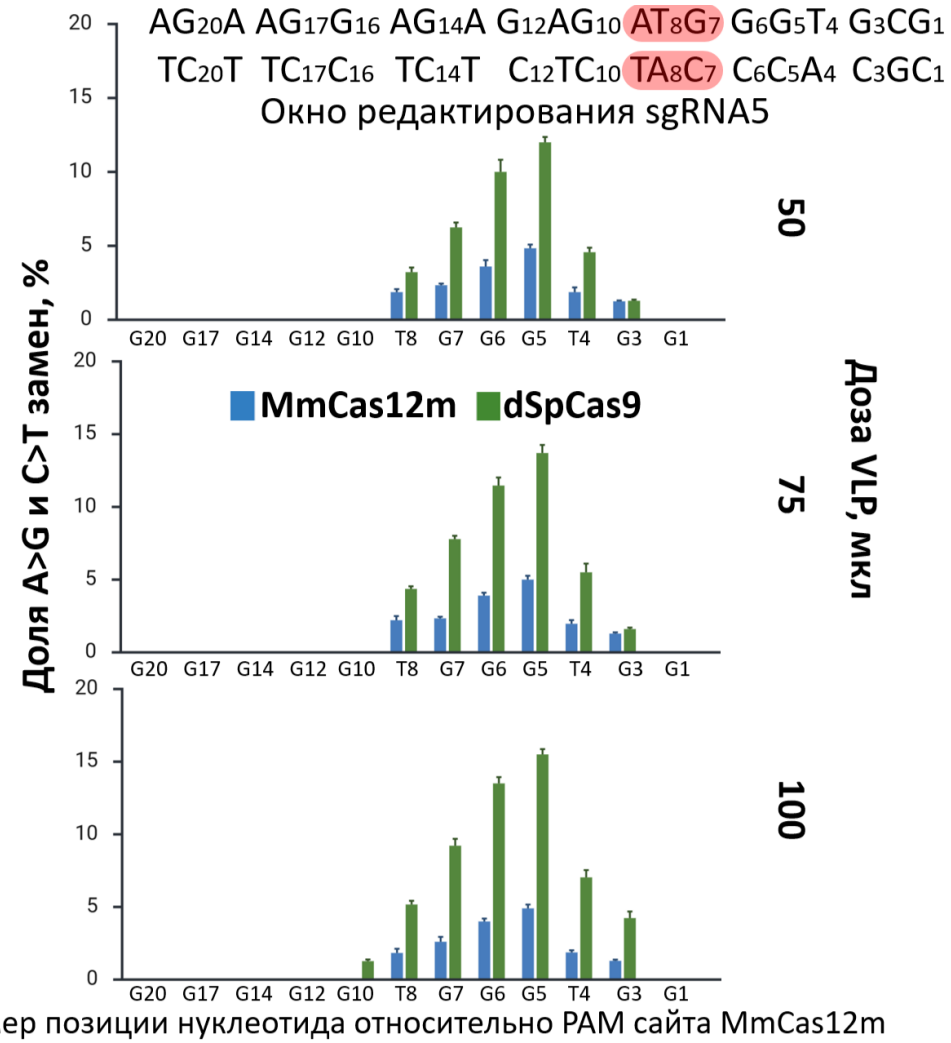


Рисунок 24. Эффективность редактирования с использованием sgRNA5. Редактирование опосредовано обработкой сконцентрированными препаратами VLP различного объема. Нуклеотиды старт-кодона гена *gag* выделены красным. Размер выборки n = 3 (биологические повторы). *$p < 0,05$.

Обработка клеток препаратом VLP-dSpCas9-TadDE-sgRNA5 обеспечила повышенную эффективность редактирования: 6,21–9,67 % прочтений имели мутации в позиции C_7 и 2,91–5,45 % – в позиции A_8. Суммарная эффективность мутагенеза старт-кодона гена *gag* при использовании VLP-dSpCas9-TadDE-sgRNA5 составила 9,12–15,12 % в зависимости от вносимого объёма VLP (табл. 5). Примечательно, что dSpCas9-TadDE также обеспечивал лучшую эффективность по сравнению с MmCas12m-TadDE, что сопоставляется с результатами редактирования при прямой трансфекции плазмидами. Стоит отметить, что редакторы оснований, доставляемые в составе VLP, сохранили прежние размеры окна редактирования и пик эффективности редактирования. Данные результаты будут учтены при дизайне новых sgRNA, чтобы пик редактирования приходился на нуклеотиды старт-кодона гена *gag*.

Таблица 5. Эффективность мутагенеза старт-кодона *gag* в зависимости от применяемого редактора оснований и дозы VLP. Размер выборки n = 3 (биологические повторы).

	Нуклеотид старт-кодона *gag*		
Название препарата	**T8**	**G7**	**Объём VLP, мкл**
VLP-MmCas12m-TadDE-sgRNA5	2,04; 2,04; 2,11	2,55; 2,71; 2,44	50
VLP-dSpCas9-TadDE-sgRNA5	3,2; 2,91, 3,54	6,21; 5,93; 6,56	
VLP-MmCas12m-TadDE-sgRNA5	2,28; 2,1; 2,45	2,35; 3,21; 2,44	75
VLP-dSpCas9-TadDE-sgRNA5	4,21; 4,56; 4,32	7,55; 8,01; 7,79	
VLP-MmCas12m-TadDE-sgRNA5	2,52; 2,1; 2,91	2,91; 2,65; 3,24	100

VLP-dSpCas9-TadDE-sgRNA5	5,1; 5,45; 4,95	9,25; 8,72; 9,67	
	Эффективность редактирования, %		

В ходе дополнительных работ выявлено, что доля отредактированного основания относительно референса в некоторых позициях может достигать 24 % при использовании VLP-dSpCas9-TadDE с другими ранее рассмотренными sgRNA. Эффективность редактирования с применением VLP-MmCas12m-TadDE и других ранее рассмотренных sgRNA способна достигать максимальной отметки 10,9–11,5 % в некоторых позициях окон редактирования. При этом повышение объёма VLP-MmCas12m-TadDE-sgRNA5 до 150 и 200 мкл не приводило к статистически значимому повышению эффективности мутагенеза старт-кодона *gag* (данные не приведены). Мы предполагаем, что насыщение эффективности указывает на то, что ограничение на величину редактирования определяется не количеством доставленных частиц, а внутренними параметрами системы: вероятностью связывания комплекса с целевым локусом, эффективностью внутриядерного транспорта, временем жизни редактора в ядре и/или скоростью каталитического цикла TadDE. Таким образом, дальнейшее увеличение дозы не повышает эффективность, что является благоприятным фактором с точки зрения безопасности: не существует «дозового окна», в котором избыточная активность редактора приводила бы к росту нецелевого эффекта.

3.8 Повышение эффективности мутагенеза старт-кодона гена *gag* путём дизайна оптимальных sgRNA

В ходе предыдущих экспериментов мы установили, что наибольшая эффективность редактирования наблюдается в позициях 5–6 протоспейсера sgRNA. Учитывая предыдущие результаты, мы провели дизайн двух новых sgRNA таким образом, чтобы старт-кодон располагался в позициях 5–7 протоспейсера (рис. 25). Оптимизация выбора sgRNA должна обеспечить повышенную эффективность мутагенеза старт-кодона гена *gag*. sgRNA

подбирали на обратную цепь геномной ДНК, поскольку именно она является матричной и содержит два нуклеотида C и A, которые комплементарны старт-кодону.

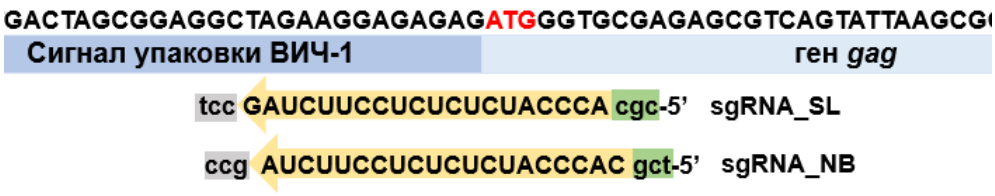


Рисунок 25. Схематичное изображение фрагмента провируса ВИЧ-1, на который предполагается таргетирование редакторов оснований MmCas12m-TadDE и dSpCas9-TadDE с использованием дополнительных sgRNA. Жёлтым выделены использованные sgRNA. Красными буквами выделен старт-кодон *gag*. Зёленым выделены PAM MmCas12m в случае каждой sgRNA, серым выделены PAM dSpCas9 в случае каждой sgRNA.

Продукцию VLP-MmCas12m-TadDE и VLP-dSpCas9-TadDE с дополнительными sgRNA, обработку клеток TZM-bl, анализ редактирования проводили по прежним протоколам.

В результате анализа методом NGS установлено, что sgRNA_SL не обеспечила редактирование для обоих редакторов оснований. Вероятно, это связано с тем, что sgRNA_SL примыкает к неоптимальному PAM (совпадение 1/3) для обоих Cas-белков. В случае MmCas12m sgRNA_SL примыкает к CGC (TTN – оптимальный PAM), в случае dSpCas9 sgRNA_SL примыкает к CCT (NGG – оптимальный PAM).

В то же время применение sgRNA_NB обеспечило редактирование только в комплексе с MmCas12m-TadDE. В случае dSpCas9-TadDE sgRNA_NB примыкает к GCC (совпадение 1/3 с оптимальным PAM), что, вероятно, отразилось на таргетировании. Средняя эффективность (с учётом стандартной ошибки среднего) мутагенеза старт-кодона *gag* при использовании препарата VLP-MmCas12m-TadDE-sgRNA_NB составила 19,55 (±1,91) % (табл. 6). Таким образом, отсутствие событий редактирования с sgRNA_SL является ожидаемым результатом и подтверждает PAM-зависимость таргетирования обоих Cas-

белков. Для будущих работ целесообразно подбирать sgRNA с PAM-совпадением не менее 2/3 для MmCas12m и dSpCas9.

Таблица 6. Эффективность редактирования области старт-кодона гена *gag* провируса ВИЧ-1 при обработке клеток TZM-bl препаратами VLP, несущими редакторы оснований. Подчеркнуты нуклеотиды старт-кодона гена *gag* обратной (матричной) цепи ДНК. В таблице указаны средние значения. n = 3 (биологические повторы).

Наименование препарата	Позиция нуклеотида в протоспейсере и эффективность его редактирования: A>G, C>T								Оптимальность PAM (N/3)
VLP-MmCas12m-TadDE-sgRNA_NB	C_{10}	C_8	$\underline{A_6}$	$\underline{C_5}$	C_4	C_3	A_2	C_1	TCG (2/3)
	0,52	3,53	<u>8,24</u>	<u>11,31</u>	6,26	3,92	2,16	0	

Стоит отметить, что в области старт-кодона гена *gag* оптимальные PAM для широко применяемого SpCas9 встречаются редко, что вероятно затрудняет редактирование. Поэтому большую часть редактирования мы проводили в условиях неоптимального PAM. Это ещё раз подчёркивает необходимость в создании новых инструментов, расширяющих спектр доступных локусов редактирования. Мы впервые представили результаты успешного мутагенеза старт-кодона гена *gag* провируса ВИЧ-1 в модельных клетках линии TZM-bl. В ходе применения дополнительных sgRNA выявлено, что MmCas12m-TadDE обеспечил мутагенез старт-кодона гена *gag* с эффективностью, превосходящей таковую при использовании dSpCas9-TadDE. Таким образом, MmCas12m-TadDE может применяться для мутагенеза в тех случаях, где применение dSpCas9-TadDE невозможно или сопряжено с низкой эффективностью, что и было показано в ходе нашей работы.

3.9 Оценка безопасности применения вирусоподобных частиц VLP-MmCas12m-TadDE и VLP-dSpCas9-TadDE при мутагенезе старт-кодона *gag* в клетках линии TZM-bl

Важным условием применения инструментов геномного редактирования является безопасность и минимизация нецелевых эффектов. Чужеродная ДНК может персистировать, вызывая сверхэкспрессию или интегрироваться в геном клетки. VLP не содержат геномных нуклеиновых кислот, что является их преимуществом. Однако во время продукции VLP в клетках-продуцентах используется плазмидная ДНК. В теории, молекулы плазмидной ДНК могут быть случайно инкапсулированы в VLP в ходе морфогенеза, что приведёт к их попаданию в клетки на этапе обработки. Чтобы выявить остаточное количество плазмидной ДНК в образцах геномной ДНК клеток линии TZM-bl, мы провели кПЦР с использованием специфичных праймеров к каждой из 4 плазмид, которые использовали для продукции VLP. Однако выявить остаточную ДНК выше фонового уровня не удалось, что свидетельствует о чистоте обработанных клеток линии TZM-bl от чужеродной ДНК. кПЦР с использованием тех же праймеров для выявления плазмидной ДНК в очищенных и сконцентрированных образцах каждого типа VLP также не показала присутствие остаточной плазмидной ДНК.

С использованием программы Cas-OFFinder выявляли вероятные нецелевые сайты редактирования для двух наиболее эффективных препаратов: VLP-MmCas12m-TadDE-sgRNA_NB и VLP-dSpCas9-TadDE-sgRNA5. Анализ ДНК последовательностей, схожих с целевым сайтом редактирования, проводили в полном геноме человека и провирусе ВИЧ-1. Количество несоответствий в sgRNA и геномной ДНК устанавливали равным «3 или менее». Данное значение является стандартным при анализе нецелевых сайтов, поскольку несоответствия в 4 и более нуклеотидов между sgRNA и целевой ДНК, как правило, исключает взаимодействие Cas-белка с ДНК. В ходе анализа выявили по 4 вероятных нецелевых сайта для MmCas12m и dSpCas9 (все – в геноме человека; в провирусе ВИЧ-1 таковые отсутствуют) (табл. 7). В ходе NGS

анализа вероятных нецелевых сайтов редактирования каких-либо мутаций *de novo* не выявили.

Таблица 7. Список потенциальных нецелевых сайтов редактирования при использовании sgRNA_NB для MmCas12m и sgRNA5 для SpCas9.

Номер нецелевого сайта	**Последовательность геномной ДНК (несоответствия с sgRNA выделены прописными буквами)**	**Хромосома**	**Позиция**	**Праймеры для амплификации вероятного нецелевого сайта**
OT1-Cas12m-NB	CACCCATCTCTCTCCcTCTt	15	26058353	OT1-12-F, OT1-12-R
OT2-Cas12m-NB	CcCCCATCTCTCTCCTTCTc	1	159257342	OT2-12-F, OT2-12-R
OT3-Cas12m-NB	CACCCcTCTCTCTCCTgCTA	16	56835631	OT3-12-F, OT3-12-R
OT4-Cas12m-NB	tcCCCATCTCTCTCCTTCTA	X	20956413	OT4-12-F, OT4-12-R
OT1-Cas9-5	GCAgCCATCTCTCTCCTgCT	9	115090378	OT1-9-F, OT1-9-R
OT2-Cas9-5	GCAtCCtTCTCcCTCCTTCT	15	75682079	OT2-9-F, OT2-9-R
OT3-Cas9-5	GCAtCCtTCTCcCTCCTTCT	15	84401603	OT3-9-F, OT3-9-R
OT4-Cas9-5	GCcCCCATCTtTgTCCTTCT	5	33536463	OT4-9-F OT4-9-R

3.10 Моделирование пространственной структуры редакторов оснований MmCas12m-TadDE и dSpCas9-TadDE

Представляет интерес воссоздать пространственную структуру двойных редакторов оснований MmCas12m-TadDE и dSpCas9-TadDE с использованием современного компьютерного моделирования, что позволило бы подтвердить обособление доменов и сохранение их третичной структуры.

Для предсказания пространственной структуры редакторов оснований в процессе таргетирования использовали AlphaFold3, алгоритмы которого позволяют моделировать взаимодействие различных биополимеров, в том числе белков, ДНК и РНК [182]. Провели моделирование таргетирования MmCas12m-TadDE в комплексе с sgRNA_NB и dSpCas9-TadDE в комплексе с

sgRNA5 на целевые последовательности ДНК. Моделирование проведено с sgRNA, обеспечившими самую высокую эффективность мутагенеза старт-кодона гена *gag*.

В ходе моделирования установлено строгое пространственное разграничение дезаминазы TadDE от Cas-белков, которое обеспечивает глицин-сериновый линкер. Обособление двух доменов видно на двумерных тепловых картах, представленных в приложении 17.

Пространственная структура TadDE имеет пять бета-складок, окружённых четырьмя альфа-спиралями, что характерно для группы аденозиндезаминаз тРНК бактерий, включая *E. coli* [183]. Смоделированная пространственная конформация обоих Cas-белков в составе редактора оснований совпадает с описанной в литературе. Для MmCas12m определены характерные два пучка из альфа-спиралей, которые соответствуют частям REC1 и REC2 домена REC (рис. 26). REC1 и REC2 являются некаталитическими частями, они связывают молекулу ДНК, стабилизируя её относительно sgRNA. Точно моделируется структура из бета-складок, формирующая часть WED домена REC. WED состоит из нескольких бета-складок, окруженных альфа-спиралями с каждой стороны, которые формируют положительно заряженную область. WED играет ключевую роль в образовании ДНК-sgRNA дуплекса, связывая и удерживая sgRNA [184].

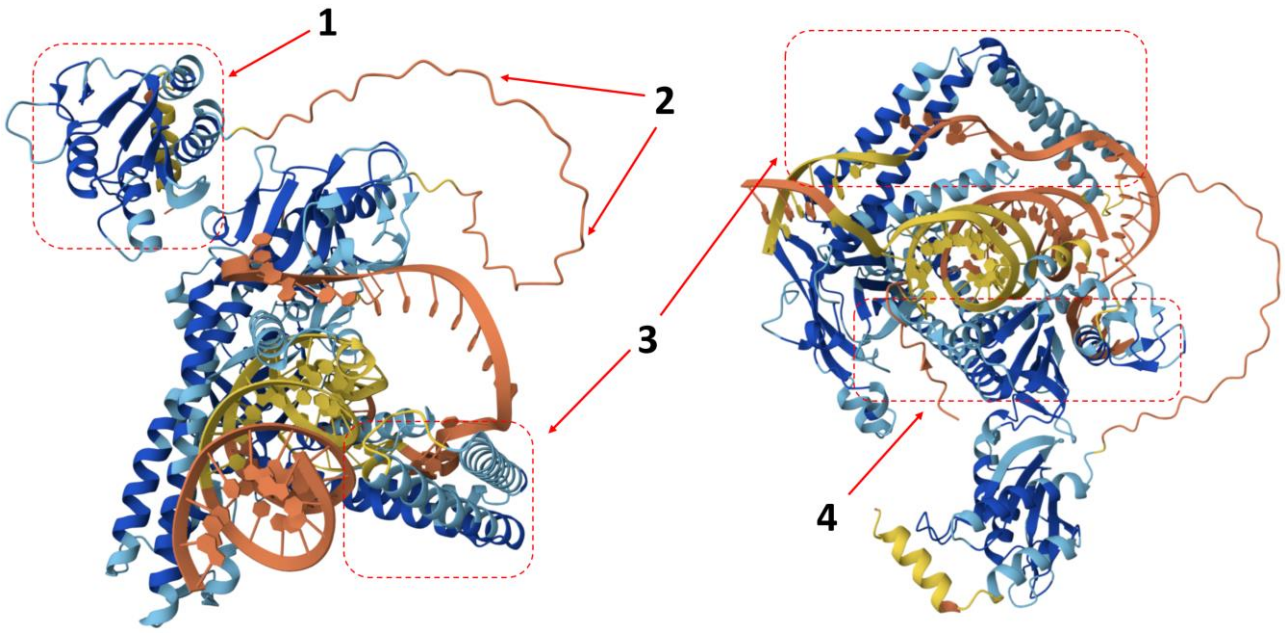


Рисунок 26. Моделирование пространственной структуры редактора оснований MmCas12m-TadDE в момент таргетирования на область старт-кодона *gag*. 1 – TadDE домен; 2 – глицин-сериновый линкер; 3 – домен REC; 4 – домен NUC. ipTM > 0,6; pTM > 0,5. Получено с помощью AlphaFold3.

Части RuvC и TNB домена NUC формируют бета-складки и окружают молекулу ДНК с противоположной стороны от альфа-спиралей REC домена. В каталитически активных вариантах Cas-белков RuvC часть обеспечивает внесение двухцепочечных разрывов. Однако в случае каталитически неактивных Cas12m белков RuvC обеспечивает усиление связывания с ДНК после формирования гетеродуплекса ДНК-sgRNA. TNB часть содержит домен цинковых пальцев типа HCCC и также обеспечивает связывание с ДНК в районе PAM [184].

При моделировании пространственной структуры dSpCas9-TadDE также отличимы характерные альфа-спирали домена REC. Домен NUC на пространственной модели взаимодействует с ДНК с противоположной стороны от домена REC. Таким образом, два домена окружают молекулу ДНК с обеих сторон, что характерно для SpCas9 [185]. Дезаминаза TadDE пространственно вынесена от dSpCas9, как и в случае MmCas12m-TadDE (рис. 27).

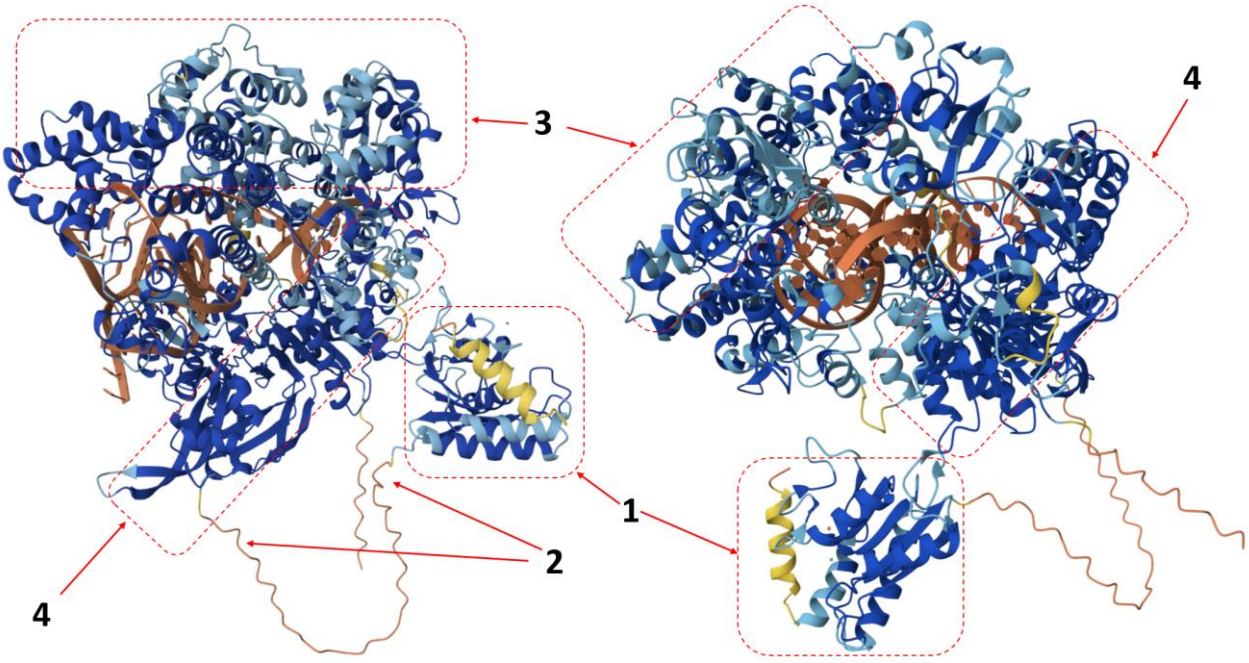


Рисунок 27. Моделирование пространственной структуры редактора оснований dSpCas9-TadDE в момент таргетирования на область старт-кодона *gag*. 1 – TadDE домен; 2 – глицин-сериновый линкер; 3 – домен REC; 4 – домен NUC. ipTM > 0,6; pTM > 0,5. Получено с помощью AlphaFold3.

Согласно модели AlphaFold3, глицин-сериновый линкер (11 амк) обеспечивает пространственное обособление домена TadDE от обоих Cas-белков: на картах предсказанного выравнивания ошибок (PAE) значения для интерфейса между TadDE и REC/NUC доменами максимальны, а показатель достоверности интерфейса (ipTM) минимален. Это свидетельствует об отсутствии устойчивых конформационных контактов и, соответственно, о стерической независимости каталитического центра дезаминазы от белкового каркаса нуклеаз. Расстояние от каталитического центра TadDE до оси ДНК-дуплекса в MmCas12m-TadDE находится вблизи верхней границы эффективного радиуса действия дезаминазы TadA-8e (оценочно ~15 Å, исходя из глубины субстратного кармана и геометрии активного центра, описанных для семейства тРНК-аденозиндезаминаз [145, 183]). Структурно это объясняется следующим образом: при связывании Cas-белка с целевым локусом и формировании R-петли (R-loop) целевая цепь ДНК спаривается с sgRNA, тогда как противоположная, нецелевая цепь (non-target strand), вытесняется из дуплекса и переходит в одноцепочечное состояние. Именно на этом одноцепочечном участке дезаминаза получает доступ к экзоциклическим аминным группам азотистых оснований. Ширина и профиль раскрытия R-петли варьируют вдоль протоспейсера. В PAM-проксимальной области (позиции 1–4 от PAM) ДНК подвергается локальной деспирализации, тогда как в PAM-дистальном направлении (позиции 9–12) нецелевая цепь термодинамически стремится к ренатурации. Максимальная амплитуда раскрытия и доступность цепи приходятся на центральную часть протоспейсера (позиции 5–8), где одноцепочечный участок наиболее стабилен, а основания максимально открыты для вхождения в субстратный карман дезаминазы.

Данная модель полностью согласуется с наблюдаемой шириной окна редактирования (позиции 1–12 от PAM) и пиком эффективности в позициях 5–6 (рис. 17–18, табл. 6). Нуклеотиды в позициях 1–4 модифицируются субоптимально, поскольку степень их вытеснения из дуплекса ограничена, а геометрия активного центра TadDE не позволяет эффективно дезаминировать основания, расположенные в непосредственной близости к регулярной двойной спирали. Аналогично, позиции 9–12 находятся на дистальном «хвосте» R-петли, где вероятность ренатурации ДНК возрастает, а время пребывания основания в одноцепочечной форме сокращается.

Примечательно, что при принципиально различной архитектуре исследуемых Cas-белков (MmCas12m, 596 амк, с характерной для семейства Cas12m геометрией R-петли; dSpCas9, 1368 амк, с архитектурой (REC + NUC), типичной для CRISPR-Cas типа II) экспериментальные границы окна и позиция пика относительно PAM практически совпадают. Это указывает на то, что лимитирующим фактором геометрии редактирования является не индивидуальная структура Cas-белка как таковая, а консервативная протяжённость вытесненной цепи (~20 нуклеотидов протоспейсера) в сочетании с: (1) длиной и гибкостью линкера, определяющего пространственную свободу каталитического центра TadDE, (2) радиусом действия дезаминазы TadDE; (3) локальным профилем доступности одноцепочечной ДНК в пределах сформированной R-петли.

Дополнительным преимуществом каталитически неактивных Cas-белков (MmCas12m с заменой H-D-D в RuvC; dSpCas9 D10A/H840A) является повышенная кинетическая стабильность R-петли. Отсутствие гидролиза ДНК исключает преждевременный распад рибонуклеопротеинового комплекса и увеличивает время контакта дезаминазы с одноцепочечным субстратом. Это вносит существенный вклад в профиль редактирования по сравнению с традиционными никазными формами (nCas), где одноцепочечный разрыв, стимулирующий репарационные системы клетки, одновременно ограничивает

«окно возможностей» дезаминазы временным горизонтом до момента репарации повреждения [175, 176].

ЗАКЛЮЧЕНИЕ

Развитие инструментов прецизионного геномного редактирования сопровождается переходом от систем, вносящих двухцепочечные разрывы в ДНК, к инструментам, действующим через точечное дезаминирование оснований без разрыва фосфодиэфирного остова. В связи с этим поиск компактных, безопасных и функционально универсальных редакторов оснований остаётся приоритетным направлением, а их валидация в прикладных моделях – обязательным этапом. Результаты настоящей работы демонстрируют системный подход на стыке этих задач: от дизайна химерных генов (MmCas12m-TadDE и dSpCas9-TadDE) до экспериментального подтверждения индуцированного мутагенеза в провирусной ДНК ВИЧ-1 в геноме модельной клеточной линии.

На первом этапе создана репортерная система быстрого скрининга для оценки связывания химерных Cas-белков и действия редакторов оснований в области старт-кодона гена *gag* провируса ВИЧ-1. Система скрининга реализована в рамках двух плазмид. Плазмида pKW_R1 успешно применена для выявления эффекта MmCas12m и dSpCas9 на подавление экспрессии гена-репортера по механизму CRISPR-интерференции, pKW_R3 – для выявления целевого редактирования с использованием редакторов оснований. Преимуществом использования плазмид серии pKW_R является технологическая доступность и высокая скорость получения предварительных результатов для подбора sgRNA, верификации взаимодействий Cas-белков с ДНК *in cellulo*, а также оценки функциональной активности редакторов оснований.

Впервые получен химерный ген NLS-MmCas12m-NLS-TadDE, кодирующий новый двойной редактор оснований MmCas12m-TadDE. В ходе последовательных экспериментов подтверждена его функциональность. Компактные размеры в 2409 п.н. позволяют при необходимости закодировать данный ген в одиночный AAV вектор, что невозможно в случае широко применяемых SpCas9 или SaCas9. В то же время в качестве группы сравнения на

протяжении всего исследования выступал аналог MmCas12m-TadDE – dSpCas9-TadDE, который продемонстрировал в некоторых случаях повышенную эффективность (до 40 %), но был более зависим от последовательности PAM. MmCas12m-TadDE при этом проявлял себя как вариант, менее восприимчивый к неоптимальному PAM. Эффективность редактирования при использовании MmCas12m-TadDE достигала около 10–11 % в условиях оптимального и неоптимального PAM. dSpCas9-TadDE демонстрировал 2–3 кратное снижение эффективности в условиях неоптимального PAM или полную утрату способности к таргетному редактированию.

Разработанная система доставки на основе VLP функциональна: эффективность редактирования при обработке клеток концентрированными препаратами VLP-MmCas12m-TadDE и VLP-dSpCas9-TadDE сопоставима с таковой при прямой трансфекции плазмидами, кодирующими соответствующие редакторы оснований. При этом клетки, обработанные VLP, не содержали остаточной чужеродной плазмидной ДНК и не имели мутаций *de novo* в анализируемых нецелевых сайтах редактирования. Совокупность этих данных подтверждает, что VLP представляют собой эффективную и генотоксично безопасную платформу для внутриклеточной доставки комплексов геномного редактирования.

Результаты, полученные в ходе выполнения диссертационной работы, опубликованы в четырёх статьях в рецензируемых журналах, рекомендованных ВАК, и представлены на ряде международных конференций. На химерный ген NLS-MmCas12m-NLS-TadDE и оригинальные генетические конструкции получен патент РФ. Представленные данные создают экспериментальную основу для дальнейшей оптимизации редакторов оснований MmCas12m-TadDE и dSpCas9-TadDE, расширения набора sgRNA на другие ключевые гены провируса ВИЧ-1 и адаптации VLP-платформы для *in vivo* применения. В более широком контексте работа подтверждает, что стратегия комбинирования каталитически неактивных Cas-белков с компактной двойной дезаминазой и

доставки в составе VLP является рабочим подходом к точному мутагенезу генов и геномов с минимальной генотоксичностью.

ВЫВОДЫ

1. Впервые осуществлен биоинформатический дизайн и синтез *de novo* кодон-оптимизированного химерного гена NLS-MmCas12m-NLS-TadDE, кодирующего новый двойной редактор оснований MmCas12m-TadDE. Получен контрольный редактор-аналог dSpCas9-TadDE на основе каталитически неактивной формы dSpCas9 (D10A/H840A). Оба гена клонированы в генетические конструкции для экспрессии в эукариотических клетках;

2. Разработана и апробирована репортерная система быстрого скрининга, позволяющая эффективно оценивать ДНК-связывающую способность различных химерных Cas-белков, а также каталитическую активность редакторов оснований в отношении целевого локуса петли SL4 сигнала упаковки PSI провируса ВИЧ-1;

3. Пиковая эффективность редактирования таргетного региона генома клеток линии TZM-bl при использовании MmCas12m-TadDE составила 10,9–11,5 % в позициях протоспейсера для ряда отобранных sgRNA. Контрольный редактор-аналог dSpCas9-TadDE продемонстрировал при оптимальном PAM повышенную эффективность редактирования до ~40 %. Вместе с тем dSpCas9-TadDE проявляет выраженную зависимость от последовательности PAM – для sgRNA, примыкающих к неоптимальному PAM, эффективность снижается в 2–3 раза или утрачивается полностью. Установлено, что оба редактора оснований обладают сопоставимой шириной окна редактирования в позициях 1–12 протоспейсера. Пик эффективности редактирования для обоих редакторов оснований приходится в позиции 5–6 протоспейсера. Доля прочтений, в которых наблюдались сразу оба типа замен (A>G, C>T), составляла менее 1 % от общего числа. Таким образом локус после

редактирования содержит преимущественно один тип мутации: A>G или C>T, что характерно для двойных редакторов оснований;

4. Разработана система на основе вирусоподобных частиц для доставки двойных редакторов оснований MmCas12m-TadDE и dSpCas9-TadDE в эукариотические клетки *in vitro*. Полученные вирусоподобные частицы проявляют способность к проникновению и доставке редакторов оснований с сохранением их функциональности в клетки линий HEK293T и TZM-bl;

5. Установлено, что обработка клеток линии TZM-bl с использованием VLP-MmCas12m-TadDE и VLP-dSpCas9-TadDE приводит к мутагенезу старт-кодона гена *gag* с эффективностью, сопоставимой с таковой при прямой трансфекции плазмидами, кодирующими соответствующий редактор оснований. Суммарная эффективность мутагенеза старт-кодона *gag* при использовании VLP-MmCas12m-TadDE-sgRNA_NB достигает 19,5 (±1,9) %. А при использовании VLP-dSpCas9-TadDE-sgRNA5 – 15,3 (±2,3) %. Мутации *de novo* в вероятных нецелевых локусах не обнаружены. Остаточная чужеродная плазмидная ДНК, использованная при продукции VLP, отсутствует как в очищенных препаратах частиц, так и в образцах геномной ДНК обработанных клеток.

ДАЛЬНЕЙШИЕ НАПРАВЛЕНИЯ ИССЛЕДОВАНИЙ

Созданные нами редакторы оснований функциональны и обеспечивают воспроизводимое редактирование. Редактор оснований MmCas12m-TadDE обеспечивает эффективность редактирования в некоторых позициях протоспейсера до ~10–11 %. Схожая эффективность, как описывалось ранее, характерна и для другого представителя группы Cas12m – GoCas12m. Пиковая эффективность dSpCas9-TadDE достигает ~40 %, что является высоким показателем для каталитически неактивной формы Cas-белка, не несущей риска одноцепочечных и двухцепочечных разрывов. Вместе с тем текущая эффективность MmCas12m-TadDE, по-видимому, не исчерпывает потенциал системы, что подтверждается сравнительно немногочисленными, но обнадеживающими литературными данными по группе Cas12m. Дальнейшие исследования будут направлены на её повышение путём модификации компонентов редактора. Среди вероятных путей модификации текущего варианта MmCas12m-TadDE мы рассматриваем вариации линкеров (XTEN, пролиновые, глицин-сериновые) и изменение их длины, для оптимизации пространственного положения каталитического центра TadDE относительно оси ДНК. Имеет определённый смысл рассмотреть вариант с изменением количества сигналов ядерной локализации и их расположением в комплексе редактирования. Предварительно полагаем, что компактная доменная организация MmCas12m может ограничивать доступность NLS для распознавания транспортными белками клетки, в связи с чем вынос NLS на гибкий линкер может улучшить внутриядерный транспорт комплекса. Все высказанные предположения будут проверены в ходе последующих исследований. Параллельно планируется подбор sgRNA для таргетирования редакторов оснований на консервативные регионы ОРС других ключевых генов ВИЧ-1: *tat*, *rev*, *env*. Перспективным направлением является продукция VLP, несущих несколько молекул редактора оснований с различными sgRNA для одновременного множественного мутагенеза нескольких генов провируса ВИЧ-1 в рамках единой дозы препарата.

СПИСОК ЛИТЕРАТУРЫ

1. Gaj T., Gersbach C.A., Barbas C.F. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering // Trends Biotechnol. – 2013. – V. 31. – P. 397-405.

2. Gupta R.M., Musunuru K. Expanding the genetic editing tool kit: ZFNs, TALENs, and CRISPR-Cas9 // J Clin Invest. – 2014. – V. 124. – P. 4154-4161.

3. Cradick T.J., Ambrosini G., Iseli C., Bucher P., McCaffrey A.P. ZFN-Site searches genomes for zinc finger nuclease target sites and off-target sites // BMC Bioinformatics. – 2011. – V. 12. – P. 152.

4. Mushtaq M., Bhat J.A., Mir Z.A., Sakina A., Ali S., Singh A.K., Tyagi A., Salgotra R.K., Dar A.A., Bhat R. CRISPR/Cas approach: A new way of looking at plant-abiotic interactions // J Plant Physiol. – 2018. – V. 224-225. – P. 156-162.

5. Sakovina L., Vokhtantsev I., Vorobyeva M., Vorobyev P., Novopashina D. Improving Stability and Specificity of CRISPR/Cas9 System by Selective Modification of Guide RNAs with 2′-fluoro and Locked Nucleic Acid Nucleotides // Int. J. Mol. Sci. – 2022. – V. 23. – P. 13460.

6. Jeong Y.K., Lee S., Hwang G.H., Hong S.A., Park S.E., Kim J.S., Woo J.S., Bae S. Adenine base editor engineering reduces editing of bystander cytosines // Nat Biotechnol. – 2021. – V. 39. – P. 1426-1433.

7. Brezgin S., Kostyusheva A., Kostyushev D., Chulanov V. Dead Cas Systems: Types, Principles, and Applications // Int J Mol Sci. – 2019. – V. 20. – P. 6041.

8. Wu W.Y., Mohanraju P., Liao C., Adiego-Pérez B., Creutzburg S.C.A., Makarova K.S., Keessen K., Lindeboom T.A., Khan T.S., Prinsen S., Joosten R., Yan W.X., Migur A., Laffeber C., Scott D.A., Lebbink J.H.G., Koonin E.V., Beisel C.L., van der Oost J. The miniature CRISPR-Cas12m effector binds DNA to block transcription // Mol Cell. – 2022. – V. 82. – P. 4487-4502.e7.

9. Bigelyte G., Duchovska B., Zedaveinyte R., Sasnauskas G., Sinkunas T., Dalgediene I., Tamulaitiene G., Silanskas A., Kazlauskas D., Valančauskas L., Madariaga-Marcos J., Seidel R., Siksnys V., Karvelis T. Innate programmable DNA binding by CRISPR-Cas12m effectors enable efficient base editing // Nucleic Acids Res. – 2024. – V. 52. – P. 3234-3248.

10. Neugebauer M.E., Hsu A., Arbab M., Krasnow N.A., McElroy A.N., Pandey S., Doman J.L., Huang T.P., Raguram A., Banskota S., Newby G.A., Tolar J., Osborn M.J., Liu D.R. Evolution of an adenine base editor into a small, efficient cytosine base editor with low off-target activity // Nat Biotechnol. – 2023. – V. 41. – P. 673–685.

11. Tian X., Chen J., Wang X., Xie Y., Zhang X., Han D., Fu H., Yin W., Wu N. Global, regional, and national HIV/AIDS disease burden levels and trends in 1990-2019: A systematic analysis for the global burden of disease 2019 study // Front Public Health. – 2023. – V. 94. – № 10. – P. 5061-5065.

12. О состоянии санитарно-эпидемиологического благополучия населения в Российской Федерации в 2025 году: Государственный доклад. Москва : Федеральная служба по надзору в сфере защиты прав потребителей и благополучия человека, 2026. – 407 с.

13. Li Q.H., Zhang Y.Q., Li E.L., Guo Q., Chen X.H., Wang F.X., Wang J.Y. Characterization of an imported HIV-1 A1/A7/G recombinant in China // Virol J. – 2024. – V. 21. – P. 3.

14. Montavon C., Toure-Kane C., Nkengasong J.N., Vergne L., Hertogs K., Mboup S., Delaporte E., Peeters M. CRF06-cpx: a new circulating recombinant form of HIV-1 in West Africa involving subtypes A, G, K, and J // J Acquir Immune Defic Syndr. – 2002. – V. 29. – P. 522-530.

15. Thomson M.M., Casado G., Posada D., Sierra M., Nájera R. Identification of a novel HIV-1 complex circulating recombinant form (CRF18_cpx) of Central African origin in Cuba // AIDS. – 2005. – V. 19. – P. 1155-1163.

16. Sakuragi J. Morphogenesis of the Infectious HIV-1 Virion // Front Microbiol. – 2011. – V. 2. – P. 242.

17. Rossi E., Meuser M.E., Cunanan C.J., Cocklin S. Structure, Function, and Interactions of the HIV-1 Capsid Protein // Life. – 2021. – V. 11. – P. 100.

18. Blumenthal R., Durell S., Viard M. HIV entry and envelope glycoprotein-mediated fusion // J Biol Chem. – 2012. – V. 287. – P. 40841-40849.

19. Stano A., Leaman D.P., Kim A.S., Zhang L., Autin L., Ingale J., Gift S.K., Truong J., Wyatt R.T., Olson A.J., Zwick M.B. Dense Array of Spikes on HIV-1 Virion Particles // J Virol. – 2017. – V. 91. – P. e00415-17.

20. Li W., Qin Z., Nand E., Grunst M.W., Grover J.R., Bess J.W.Jr., Lifson J.D., Zwick M.B., Tagare H.D., Uchil P.D., Mothes W. HIV-1 Env trimers asymmetrically engage CD4 receptors in membranes // Nature. – 2023. – V. 623. – P. 1026–1033.

21. Glatzová D., Cebecauer M. Dual Role of CD4 in Peripheral T Lymphocytes // Front. Immunol. – 2019. – V. 10. – P. 618.

22. Caillat C., Guilligay D., Torralba J., Friedrich N., Nieva J.L., Trkola A., Chipot C.J., Dehez F.L., Weissenhorn W. Structure of HIV-1 gp41 with its membrane anchors targeted by neutralizing antibodies // Elife. – 2021. – V. 10. – P. e65005.

23. Marie V., Gordon M.L. The HIV-1 Gag Protein Displays Extensive Functional and Structural Roles in Virus Replication and Infectivity // Int J Mol Sci. – 2022. – V. 23. – P. 7569.

24. Bussienne C., Marquet R., Paillart J.C., Bernacchi S. Post-Translational Modifications of Retroviral HIV-1 Gag Precursors: An Overview of Their Biological Role // Int J Mol Sci. – 2021. – V. 22. – P. 2871.

25. Qu K., Ke Z., Zila V., Anders-Össwein M., Glass B., Mücksch F., Müller R., Schultz C., Müller B., Kräusslich H.G., Briggs J.A.G. Maturation of the matrix and viral membrane of HIV-1 // Science. – 2021. – V. 373. – P. 700-704.

26. Larijani M.S., Sadat S.M., Bolhassani A., Pouriayevali M.H., Bahramali G., Ramezani A. In Silico Design and Immunologic Evaluation of HIV-1 p24-Nef Fusion Protein to Approach a Therapeutic Vaccine Candidate // Curr HIV Res. – 2018. – V. 16. – P. 322-337.

27. Duchon A., Hu W-S. HIV-1 RNA genome packaging: it's G-rated // mBio. – 2024. – V. 15. – P. e0086123.

28. Li G., Piampongsant S., Faria N.R., Voet A., Pineda-Peña A.C., Khouri R., Lemey P., Vandamme A.M., Theys K. An integrated map of HIV genome-wide variation from a population perspective // Retrovirology. – 2015. – V. 12. – P. 18.

29. Wang Y., Guo C., Wang X., Xu L., Li R., Wang J. The Zinc Content of HIV-1 NCp7 Affects Its Selectivity for Packaging Signal and Affinity for Stem-Loop 3 // Viruses. – 2021. – V. 13. – P. 1922.

30. Stacey J.C.V., Hrebík D., Nand E., Shetty S.D., Qu K., Boicu M., Anders-Össwein M., Uchil P.D., Dick R.A., Mothes W., Kräusslich H.G., Müller B., Briggs J.A.G. The conserved HIV-1 spacer peptide 2 triggers matrix lattice maturation // Nature. – 2025. – V. 640. – P. 258–264.

31. Chen X., Wang X. The HIV-1 gag p6: a promising target for therapeutic intervention // Retrovirology. – 2024. – V. 21. – P. 1.

32. Liu S., Koneru P.C., Li W., Pathirage C., Engelman A.N., Kvaratskhelia M., Musier-Forsyth K. HIV-1 integrase binding to genomic RNA 5′-UTR induces local structural changes in vitro and in virio // Retrovirology. – 2021. – V. 18. – P. 37.

33. Burdick R.C., Pathak V.K. Efficient HIV-1 in vitro reverse transcription: optimal capsid stability is required // Sig Transduct Target Ther. – 2021. – V. 6. – P. 13.

34. Tabler C.O., Tilton J.C. Mechanism and Kinetics of HIV-1 Protease Activation // Viruses. – 2024. – V. 16. – P. 1826.

35. Ning J., Erdemci-Tandogan G., Yufenyuy E.L., Wagner J., Himes B.A., Zhao G., Aiken C., Zandi R., Zhang P. In vitro protease cleavage and computer simulations reveal the HIV-1 capsid maturation pathway // Nat Commun. – 2016. – V. 7. – P. 13689.

36. Fan T., Xie C., Li L., Jin X., Cui J., Wang J. HIV-1 Nef activates proviral DNA transcription by recruiting Src kinase to phosphorylate host protein Nef-associated factor 1 to compromise its viral restrictive function // J Virol. – 2025. – V.99. – P. e00280-25.

37. Li Y.L., Langley C.A., Azumaya C.M., Echeverria I., Chesarino N.M., Emerman M., Cheng Y., Gross J.D. The structural basis for HIV-1 Vif antagonism of human APOBEC3G // Nature. – 2023. – V. 615. – P. 728–733.

38. González M.E. Vpu Protein: The Viroporin Encoded by HIV-1 // Viruses. – 2015. – V. 7. – P. 4352-4368.

39. Khan N., Geiger J.D. Role of Viral Protein U (Vpu) in HIV-1 Infection and Pathogenesis // Viruses. – 2021. – V. 13. – P. 1466.

40. Waheed A., MacDonald S., Khan M., Mounts M., Swiderski M., Xu Y., Ye Y., Freed E.O. The Vpu-interacting Protein SGTA Regulates Expression of a Non-glycosylated Tetherin Species // Sci Rep. – 2016. – V. 6. – P. 24934.

41. Xiao T., Cai Y., Chen B. HIV-1 Entry and Membrane Fusion Inhibitors // Viruses. – 2021. – V. 13. – P. 735.

42. Xavier Ruiz F., Arnold E. Evolving understanding of HIV-1 reverse transcriptase structure, function, inhibition, and resistance // Curr Opin Struct Biol. – 2020. – V. 61. – P. 113-123.

43. Tu N.Q., Richetta C., Putzu F., Delelis O., Ahmed K., Masand V.H., Schobert R., Tramontano E., Corona A., Biersack B. Identification of HIV-1 Reverse

Transcriptase-Associated Ribonuclease H Inhibitors Based on 2-Hydroxy-1,4-naphthoquinone Mannich Bases // Molecules. – 2025. – V. 30. – P. 495.

44. Tramontano E., Di Santo R. HIV-1 RT-associated RNase H function inhibitors: Recent advances in drug development // Curr Med Chem. – 2010. – V. 17. – P. 2837-2853.

45. Rose M., Bai B., Tang M., Cheong C.M., Beard S., Burgess J.T., Adams M.N., O'Byrne K.J., Richard D.J., Gandhi N.S., Bolderson E. The Impact of Rare Human Variants on Barrier-To-Auto-Integration Factor 1 (Banf1) Structure and Function // Front. Cell Dev. Biol. – 2021. – V. 9. – P. 775441.

46. Burgess J.T., Cheong C.M., Suraweera A., Sobanski T., Beard S., Dave K., Rose M., Boucher D., Croft L.V., Adams M.N., O'Byrne K., Richard D.J., Bolderson E. Barrier-to-autointegration-factor (Banf1) modulates DNA double-strand break repair pathway choice via regulation of DNA-dependent kinase (DNA-PK) activity // Nucleic Acids Res. – 2021. – V. 49. – P. 3294-3307.

47. Gu W.G., Yao Z.Y. Host proteins interacting with integrase of human immunodeficiency virus type 1 // Biochem Pharmacol. – 2025. – V. 20. – P. 117254.

48. Sapp N., Burge N., Cox K., Prakash P., Balasubramaniam M., Thapa S., Christensen D., Li M., Linderberger J., Kvaratskhelia M., Pandhare J., Craigie R., Poirier M.G., Dash C. HIV-1 Preintegration Complex Preferentially Integrates the Viral DNA into Nucleosomes Containing Trimethylated Histone 3-Lysine 36 Modification and Flanking Linker DNA // J Virol. – 2022. – V. 96. – P. e01011-22.

49. Engelman A.N., Singh P.K. Cellular and molecular mechanisms of HIV-1 integration targeting // Cell. Mol. Life Sci. – 2018. – V. 75. – P. 2491–2507.

50. Feng L., Larue R.C., Slaughter A., Kessl J.J., Kvaratskhelia M. HIV-1 integrase multimerization as a therapeutic target // Curr Top Microbiol Immunol. – 2015. – V. 389. – P. 93-119.

51. Suzuki Y., Craigie R. The road to chromatin — nuclear entry of retroviruses // Nat Rev Microbiol. – 2017. – V. 5. – P. 187–196.

52. Emerman M. HIV-1, Vpr and the cell cycle // Curr Biol. – 1996. – V. 6. – P. 1096-103.

53. Nitahara-Kasahara Y., Kamata M., Yamamoto T., Zhang X., Miyamoto Y., Muneta K., Iijima S., Yoneda Y., Tsunetsugu-Yokota Y., Aida Y. Novel Nuclear Import of Vpr Promoted by Importin α Is Crucial for Human Immunodeficiency Virus Type 1 Replication in Macrophages // J Virol. – 2007. – V. 81. – P. 5284-5293.

54. Wu V.H., Nobles C.L., Kuri-Cervantes L., McCormick K., Everett J.K., Nguyen S., Del Rio Estrada P.M., González-Navarro M., Torres-Ruiz F., Ávila-Ríos S., Reyes-Terán G., Bushman F.D., Betts M.R. Assessment of HIV-1 integration in tissues and subsets across infection stages // JCI Insight. – 2020. – V. 5. – P. e139783.

55. Marini B., Kertesz-Farkas A., Ali H., Lucic B., Lisek K., Manganaro L., Pongor S., Luzzati R., Recchia A., Mavilio F., Giacca M., Lusic M. Nuclear architecture dictates HIV-1 integration site selection // Nature. – 2015. – V. 521. – P. 227–231.

56. Margolis D.M., Archin N.M. Proviral Latency, Persistent Human Immunodeficiency Virus Infection, and the Development of Latency Reversing Agents // J Infect Dis. – 2017. – V. 215. – P. S111-S118.

57. Kibe A., Buck S., Gribling-Burrer A.S., Gilmer O., Bohn P., Koch T., Mireisz C.N., Schlosser A., Erhard F., Smyth R.P., Caliskan N. The translational landscape of HIV-1 infected cells reveals key gene regulatory principles // Nat Struct Mol Biol. – 2025. – V. 32. – P. 841–852.

58. Ohlmann T., Mengardi C., López-Lastra M. Translation initiation of the HIV-1 mRNA // Translation (Austin). – 2014. – V. 2. – P. e960242.

59. Asamitsu K., Fujinaga K., Okamoto T. HIV Tat/P-TEFb Interaction: A Potential Target for Novel Anti-HIV Therapies // Molecules. – 2018. – V. 23. – P. 933.

60. Gu L., Tsuji T., Jarboui M.A., Yeo G.P., Sheehy N., Hall W.W., Gautier V.W. Intermolecular masking of the HIV-1 Rev NLS by the cellular protein HIC: Novel insights into the regulation of Rev nuclear import // Retrovirology. – 2011. – V. 8. – P. 17.

61. Bai Y., Tambe A., Zhou K., Doudna J.A. RNA-guided assembly of Rev-RRE nuclear export complexes // Elife. – 2014. – V. 3. – P. e03656.

62. Freed E. HIV-1 assembly, release and maturation // Nat Rev Microbiol. – 2015. – V. 13. – P. 484–496.

63. Sundquist W.I., Kräusslich H.G. HIV-1 assembly, budding, and maturation // Cold Spring Harb Perspect Med. – 2012. – V. 2. – P. a006924.

64. McBride M.S., Schwartz M.D., Panganiban A.T. Efficient encapsidation of human immunodeficiency virus type 1 vectors and further characterization of cis elements required for encapsidation // J Virol. – 1997. – V. 71. – P. 4544-4554.

65. Murphy R.E., Saad J.S. The Interplay between HIV-1 Gag Binding to the Plasma Membrane and Env Incorporation // Viruses. – 2020. – V. 12. – P. 548.

66. Hurley J.H., Cada A.K. Inside job: how the ESCRTs release HIV-1 from infected cells // Biochem Soc Trans. – 2018. – V. 46. – P. 1029-1036.

67. Weber I.T., Wang Y.F., Harrison R.W. HIV Protease: Historical Perspective and Current Research // Viruses. – 2021. – V. 13. – P. 839.

68. Ghosh A.K. Four decades of continuing innovations in the development of antiretroviral therapy for HIV/AIDS: Progress to date and future challenges // Glob Health Med. – 2023. – V. 5. – P. 194-198.

69. Carr A., Mackie N.E., Paredes R., Ruxrungtham K. HIV drug resistance in the era of contemporary antiretroviral therapy: A clinical perspective // Antivir Ther. – 2023. – V. 28. – P. 13596535231201162.

70. Yeo J.Y., Goh G.R., Su C.T., Gan S.K. The Determination of HIV-1 RT Mutation Rate, Its Possible Allosteric Effects, and Its Implications on Drug Resistance // Viruses. – 2020. – V. 12. – P. 297.

71. Brégeon D., Doetsch P.W. Transcriptional mutagenesis: causes and involvement in tumour development // Nat Rev Cancer. – 2011. – V. 11. – P. 218-27.

72. Rawson J.M.O., Nikolaitchik O.A., Keele B.F., Pathak V.K., Hu W.S. Recombination is required for efficient HIV-1 replication and the maintenance of viral genome integrity // Nucleic Acids Res. – 2018. – V. 46. – P. 10535-10545.

73. Desimmie B.A., Burdick R.C., Izumi T., Doi H., Shao W., Alvord W.G., Sato K., Koyanagi Y., Jones S., Wilson E., Hill S., Maldarelli F., Hu W.S., Pathak V.K. APOBEC3 proteins can copackage and comutate HIV-1 genomes // Nucleic Acids Res. – 2016. – V. 44. – P. 7848-7865.

74. Ophinni Y., Inoue M., Kotaki T., Kameoka M. CRISPR/Cas9 system targeting regulatory genes of HIV-1 inhibits viral replication in infected T-cell cultures // Sci Rep. – 2018. – V. 8. – P. 7784.

75. Ophinni Y., Miki S., Hayashi Y., Kameoka M. Multiplexed tat-Targeting CRISPR-Cas9 Protects T Cells from Acute HIV-1 Infection with Inhibition of Viral Escape // Viruses. – 2020. – V. 12. – P. 1223.

76. Wang G., Zhao N., Berkhout B., Das A.T. CRISPR-Cas9 Can Inhibit HIV-1 Replication but NHEJ Repair Facilitates Virus Escape // Mol Ther. – 2016. – V. 24. – P. 522-526.

77. Letchumanan P., Theva Das K. The role of genetic diversity, epigenetic regulation, and sex-based differences in HIV cure research: a comprehensive review // Epigenetics Chromatin. – 2025. – V. 18. – P. 1.

78. Yoder K.E., Bundschuh R. Host Double Strand Break Repair Generates HIV-1 Strains Resistant to CRISPR/Cas9 // Sci Rep. – 2016. – V. 6. – P. 29530.

79. Isaguliants M., Bayurova E., Avdoshina D., Kondrashova A., Chiodi F., Palefsky J.M. Oncogenic Effects of HIV-1 Proteins, Mechanisms Behind // Cancers Basel. – 2021. – V. 13. – P. 305.

80. Langer S., Hammer C., Hopfensperger K., Klein L., Hotter D., De Jesus P.D., Herbert K.M., Pache L., Smith N., van der Merwe J.A., Chanda S.K., Fellay J., Kirchhoff F., Sauter D. HIV-1 Vpu is a potent transcriptional suppressor of NF-κB-elicited antiviral immune responses. // Elife. – 2019. – V. 8. – P. e41930.

81. Hu W., Kaminski R., Yang F., Zhang Y., Cosentino L., Li F., Luo B., Alvarez-Carbonell D., Garcia-Mesa Y., Karn J., Mo X., Khalili K. RNA-directed gene editing specifically eradicates latent and prevents new HIV-1 infection // Proc Natl Acad Sci U S A. – 2014. – V. 111. – P. 11461-11466.

82. Canver M.C., Bauer D.E., Dass A., Yien Y.Y., Chung J., Masuda T., Maeda T., Paw B.H., Orkin S.H. Characterization of genomic deletion efficiency mediated by clustered regularly interspaced short palindromic repeats (CRISPR)/Cas9 nuclease system in mammalian cells // J Biol Chem. – 2017. – V. 292. – P. 2556.

83. Lai M., Maori E., Quaranta P., Matteoli G., Maggi F., Sgarbanti M., Crucitta S., Pacini S., Turriziani O., Antonelli G., Heeney J.L., Freer G., Pistello M. CRISPR/Cas9 Ablation of Integrated HIV-1 Accumulates Proviral DNA Circles with Reformed Long Terminal Repeats // J Virol. – 2021. – V. 95. – P. e0135821.

84. Hultquist J.F., Schumann K., Woo J.M., Manganaro L., McGregor M.J., Doudna J., Simon V., Krogan N.J., Marson A. A Cas9 Ribonucleoprotein Platform for

Functional Genetic Studies of HIV-Host Interactions in Primary Human T Cells // Cell Rep. – 2016. – V. 17. – P. 1438-1452.

85. Li C., Guan X., Du T., Jin W., Wu B., Liu Y., Wang P., Hu B., Griffin G.E., Shattock R.J., Hu Q. Inhibition of HIV-1 infection of primary CD4+ T-cells by gene editing of CCR5 using adenovirus-delivered CRISPR/Cas9 // J Gen Virol. – 2015. – V. 96. – P. 2381-2393.

86. Xu L., Wang J., Liu Y., Xie L., Su B., Mou D., Wang L., Liu T., Wang X., Zhang B., Zhao L., Hu L., Ning H., Zhang Y., Deng K., Liu L., Lu X., Zhang T., Xu J., Li C., Wu H., Deng H., Chen H. CRISPR-Edited Stem Cells in a Patient with HIV and Acute Lymphocytic Leukemia // N Engl J Med. – 2019. – V. 381. – P. 1240-1247.

87. Nie Y., Han Y.C., Zou Y.R. CXCR4 is required for the quiescence of primitive hematopoietic cells // J Exp Med. – 2008. – V. 205. – P. 777-783.

88. Zhang F., Wen Y., Guo X. CRISPR/Cas9 for genome editing: progress, implications and challenges // Hum Mol Genet. – 2014. – V. 23. – P. 40-46.

89. Knipping F., Newby G.A., Eide C.R., McElroy A.N., Nielsen S.C., Smith K., Fang Y., Cornu T.I., Costa C., Gutierrez-Guerrero A., Bingea S.P., Feser C.J., Steinbeck B., Hippen K.L., Blazar B.R., McCaffrey A., Mussolino C., Verhoeyen E., Tolar J., Liu D.R., Osborn M.J. Disruption of HIV-1 co-receptors CCR5 and CXCR4 in primary human T cells and hematopoietic stem and progenitor cells using base editing // Mol Ther. – 2022. – V. 30. – P. 130-144.

90. Nyerges Á., Bálint B., Cseklye J., Nagy I., Pál C., Fehér T. CRISPR-interference-based modulation of mobile genetic elements in bacteria // Synth Biol (Oxf). – 2019. – V. 4. – P. ysz008.

91. Koonin E.V., Makarova K.S. CRISPR in mobile genetic elements: counter-defense, inter-element competition and RNA-guided transposition // BMC Biol. – 2024. – V. 22. – P. 295.

92. Panahi B., Dehganzad B., Nami Y. CRISPR-Cas systems feature and targeting phages diversity in Lacticaseibacillus rhamnosus strains // Front. Microbiol. – 2023. – V. 14. – P. 1281307.

93. Swarts D.C., Mosterd C., van Passel M.W., Brouns S.J. CRISPR interference directs strand specific spacer acquisition // PLoS ONE. – 2012. – V. 7. – P. e35888.

94. Datsenko K.A., Pougach K., Tikhonov A., Wanner B.L., Severinov K., Semenova E. Molecular memory of prior infections activates the CRISPR/Cas adaptive bacterial immunity system // Nat. Commun. – 2012. – V. 3. – P. 945.

95. Deveau H., Barrangou R., Garneau J.E., Labonte J., Fremaux C., Boyaval P., Romero D.A., Horvath P., Moineau S. Phage response to CRISPR-encoded resistance in Streptococcus thermophilus // J. Bacteriol. – 2008. – V. 190. – P. 1390–1400.

96. Mojica F.J., Diez-Villasenor C., Garcia-Martinez J., Almendros C. Short motif sequences determine the targets of the prokaryotic CRISPR defence system // Microbiology. – 2009. – V. 155. – P. 733–740.

97. Fonfara I., Richter H., Bratovič M., Le Rhun A., Charpentier E. The CRISPR-associated DNA-cleaving enzyme Cpf1 also processes precursor CRISPR RNA // Nature. – 2016. – V. 532. – P. 517–521.

98. 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. Cpf1 is a single RNA-guided endonuclease of a Class 2 CRISPR-Cas system // Cell. – 2015. – V. 163. – P. 759–771.

99. Ma H., Marti-Gutierrez N., Park S.W., Wu J., Lee Y., Suzuki K., Koski A., Ji D., Hayama T., Ahmed R., Darby H., Van Dyken C., Li Y., Kang E., Park A.R., Kim D., Kim S.T., Gong J., Gu Y., Xu X., Battaglia D., Krieg S.A., Lee D.M., Wu D.H., Wolf D.P., Heitner S.B., Belmonte J.C.I., Amato P., Kim J.S., Kaul S., Mitalipov S. Correction of a pathogenic gene mutation in human embryos // Nature. – 2017. – V. 548. – P. 413-419.

100. Marraffini L.A. CRISPR-Cas immunity in prokaryotes // Nature. – 2015. – V. 526. – P. 55-61.

101. Mohanraju P., Makarova K.S., Zetsche B., Zhang F., Koonin E.V., van der Oost J. Diverse evolu-tionary roots and mechanistic variations of the CRISPR-Cas systems // Science. – 2016. – V. 353. – P. aad5147.

102. Haurwitz R.E., Jinek M., Wiedenheft B., Zhou K., Doudna J.A. Sequence- and structure-specific RNA processing by a CRISPR endonuclease // Science. – 2010. – V. 329. – P. 1355–1358.

103. Wei J., Shao Y., Liang Y., Bu X., Zhou W., Liang Y., Li Y. A type III-associated Cas6 functions as a negative regulator of type I-B CRISPR-Cas system in Thermus thermophilus // Commun Biol. – 2025, - V. 8. – P. 793.

104. Chen J., Lin X., Xiang W., Chen Y., Zhao Y., Huang L., Liu L. DNA target binding-induced pre-crRNA processing in type II and V CRISPR-Cas systems // Nucleic Acids Res. – 2025. – V. 53. – P. gkae1241.

105. Yang Z.X., Fu Y.W., Zhao J.J., Zhang F., Li S.A., Zhao M., Wen W., Zhang L., Cheng T., Zhang J.P., Zhang X.B. Superior Fidelity and Distinct Editing Outcomes of SaCas9 Compared with SpCas9 in Genome Editing // Genomics Proteomics Bioinformatics. – 2023. – V. 21. – P. 1206-1220.

106. Sakovina L., Vokhtantsev I., Akhmetova E., Vorobyeva M., Vorobjev P., Zharkov D.O., Novopashina D. Photocleavable Guide crRNAs for a Light-Controllable CRISPR/Cas9 System // Int. J. Mol. Sci. – 2024. – V. 25. – P. 12392.

107. Gleditzsch D., Pausch P., Müller-Esparza H., Özcan A., Guo X., Bange G., Randau L. PAM identification by CRISPR-Cas effector complexes: diversified mechanisms and structures // RNA Biol. – 2019. – V. 16. – P. 504-517.

108. Zhang Y., Ge X., Yang F., Zhang L., Zheng J., Tan X., Jin Z.B., Qu J., Gu F. Comparison of non-canonical PAMs for CRISPR/Cas9-mediated DNA cleavage in human cells // Sci Rep. – 2014. – V. 4. – P. 5405.

109. Yamano T., Zetsche B., Ishitani R., Zhang F., Nishimasu H., Nureki O. Structural Basis for the Canonical and Non-canonical PAM Recognition by CRISPR-Cpf1 // Mol Cell. – 2017. – V. 67. – P. 633-645.e3.

110. Lin Y., Li J., Li C., Tu Z., Li S., Li X.J., Yan S. Application of CRISPR/Cas9 System in Establishing Large Animal Models // Front. Cell Dev. Biol. – 2022. – V. 10. – P. 919155.

111. Tuncel A., Pan C., Clem J.S., Liu D., Qi Y. CRISPR–Cas applications in agriculture and plant research // Nat Rev Mol Cell Biol. – 2025. – V. 26. – P. 419–441.

112. Aliev T.I., Yudkin. D.V. AAV-based vectors for human diseases modeling in laboratory animals // Front. Med. – 2025. – V. 11. – P. 1499605.

113. Weber J.A., Lang J.F., Carrell E.M., Alameh M.G., Davidson B.L. Temporal restriction of Cas9 expression improves CRISPR-mediated deletion efficacy and fidelity // Mol Ther Nucleic Acids. – 2024. – V. 35. – P. 102172.

114. Cai Y., Chen L., Sun S., Wu C., Yao W., Jiang B., Han T., Hou W. CRISPR/Cas9-Mediated Deletion of Large Genomic Fragments in Soybean // Int. J. Mol. Sci. – 2018. – V. 19. – P. 3835.

115. Saito A.C., Higashi T., Chiba H. Protocol for establishing knockout cell clones by deletion of a large gene fragment using CRISPR-Cas9 with multiple guide RNAs // STAR Protoc. – 2024. – V. 5. – P. 103179.

116. Yuan G., Deng S., Czajka J.J., Dai Z., Hofstad B.A., Kim J., Pomraning K.R. CRISPR-Cas9/Cas12a systems for efficient genome editing and large genomic fragment deletions in Aspergillus niger // Front. Bioeng. Biotechnol. – 2024. – V. 12. – P. 1452496.

117. Chang H.H.Y., Pannunzio N.R., Adachi N., Lieber M.R. Non-homologous DNA end joining and alternative pathways to double-strand break repair // Nat Rev Mol Cell Biol. – 2017. – V. 18. – P. 495–506.

118. Wei X., Lai Y., Li B., Qin L., Xu Y., Lin S., Wang S., Wu Q., Liang Q., Huang G., Deng Q., Liu P., Wu D., Lai L., Yao Y., Li P. CRISPR/Cas9-Mediated Deletion of Foxn1 in NOD/SCID/IL2rg−/− Mice Results in Severe Immunodeficiency // Sci Rep. – 2017. – V. 7. – P. 7720.

119. Chen X., Du J., Yun S., Xue C., Yao Y., Rao S. Recent advances in CRISPR-Cas9-based genome insertion technologies // Mol Ther Nucleic Acids. – 2024. – V. 35. – P. 102138.

120. Wang X., Xu G., Johnson W.A., Qu Y., Yin D., Ramkissoon N., Xiang H., Cong L. Long sequence insertion via CRISPR/Cas gene-editing with transposase, recombinase, and integrase // Curr Opin Biomed Eng. – 2023. – V. 28. – P. 100491.

121. Shrestha D., Bag A., Wu R., Zhang Y., Tang X., Qi Q., Xing J., Cheng Y. Genomics and epigenetics guided identification of tissue-specific genomic safe harbors // Genome Biol. – 2022. – V. 23. – P. 199.

122. Leitão M.C., Cabral L.S., Piva L.C., Queiroz P.F.S., Gomes T.G., de Andrade R.V., Perez A.L.A., de Paiva K.L.R., Báo S.N., Reis V.C.B., Moraes L.M.P., Togawa R.C., Barros L.M.G., Torres F.A.G., Pappas Júnior G.J., Coelho C.M. SHIP identifies genomic safe harbors in eukaryotic organisms using genomic general feature annotation // Sci Rep. – 2025. – V. 15. – P. 7193.

123. Mali P., Yang L., Esvelt K.M., Aach J., Guell M., DiCarlo J.E., Norville J.E., Church G.M. RNA-guided human genome engineering via Cas9 // Science. – 2013. – V. 339. – P. 823-826.

124. Yang D., Song J., Zhang J., Xu J., Zhu T., Wang Z., Lai L., Chen Y.E. Identification and characterization of rabbit ROSA26 for gene knock-in and stable reporter gene expression // Sci Rep. – 2016. – V. 6. – P. 25161.

125. Ferreira P., Choupina A.B. CRISPR/Cas9 a simple, inexpensive and effective technique for gene editing // Mol Biol Rep. – 2022. – V. 49. – P. 7079–7086.

126. Guerra-Resendez R.S., Lydon S.L., Ma A.J., Bedford G.C., Reed D.R., Kim S., Terán E.R., Nishiguchi T., Escobar M., DiNardo A.R., Hilton I.B. Characterization of Rationally Designed CRISPR/Cas9-Based DNA Methyltransferases with Distinct Methyltransferase and Gene Silencing Activities in Human Cell Lines and Primary Human T Cells // ACS Synth Biol. – 2025. – V. 14. – P. 384-397.

127. Heidersbach A.J., Dorighi K.M., Gomez J.A., Jacobi A.M., Haley B. A versatile, high-efficiency platform for CRISPR-based gene activation // Nat Commun. 2023. – V. 14. – P. 902.

128. Bendixen L., Jensen T.I., Bak R.O. CRISPR-Cas-mediated transcriptional modulation: The therapeutic promises of CRISPRa and CRISPRi // Mol Ther. – 2023. – V. 31. – P. 1920-1937.

129. Jakhanwal S., Cress B.F., Maguin P., Lobba M.J., Marraffini L.A., Doudna J.A. A CRISPR-Cas9-integrase complex generates precise DNA fragments for genome integration // Nucleic Acids Res. – 2021. – V. 49. – P. 3546-3556.

130. Lahr W.S., Sipe C.J., Skeate J.G., Webber B.R., Moriarity B.S. CRISPR-Cas9 base editors and their current role in human therapeutics // Cytotherapy. – 2023. – V. 25. – P. 270-276.

131. Porto E.M., Komor A.C., Slaymaker I.M., Yeo G.W. Base editing: advances and therapeutic opportunities // Nat Rev Drug Discov. – 2020. – V. 19. – P. 839–859.

132. Liu D., Cao D., Han R. Recent advances in therapeutic gene-editing technologies // Mol Ther. – 2025. – V. 33. – P. 2619-2644.

133. Musunuru K., Grandinette S.A., Wang X., Hudson T.R., Briseno K., Berry A.M., Hacker J.L., Hsu A., Silverstein R.A., Hille L.T., Ogul A.N., Robinson-Garvin

N.A., Small J.C., McCague S., Burke S.M., Wright C.M., Bick S., Indurthi V., Sharma S., Jepperson M., Vakulskas C.A., Collingwood M., Keogh K., Jacobi A., Sturgeon M., Brommel C., Schmaljohn E., Kurgan G., Osborne T., Zhang H., Kinney K., Rettig G., Barbosa C.J., Semple S.C., Tam Y.K., Lutz C., George L.A., Kleinstiver B.P., Liu D.R., Ng K., Kassim S.H., Giannikopoulos P., Alameh M.G., Urnov F.D., Ahrens-Nicklas R.C. Patient-Specific In Vivo Gene Editing to Treat a Rare Genetic Disease // N Engl J Med. – 2025. – V. 392. – P. 2235-2243.

134. Klaus V., Vermeulen T., Minassian B., Israelian N., Engel K., Lund A.M., Roebrock K., Christensen E., Häberle J. Highly variable clinical phenotype of carbamylphosphate synthetase 1 deficiency in one family: an effect of allelic variation in gene expression? // Clin Genet. – 2009. – V. 76. – P. 263-269.

135. Shilova O.N., Tsyba D.L., Shilov E.S. Mutagenic Activity of AID/APOBEC Deaminases in Antiviral Defense and Carcinogenesis // Mol Biol. – 2022. – V. 56. – P. 46-58.

136. Komor A.C., Kim Y.B., Packer M.S., Zuris J.A., Liu D.R. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage // Nature. – 2016. – V. 533. – P. 420-424.

137. Schormann N., Ricciardi R., Chattopadhyay D. Uracil-DNA glycosylases-structural and functional perspectives on an essential family of DNA repair enzymes // Protein Sci. – 2014. – V. 23. – P. 1667-1685.

138. Kunz C., Saito Y., Schär P. DNA Repair in mammalian cells: Mismatched repair: variations on a theme // Cell Mol Life Sci. – 2009. – V. 66. – P. 1021-1038.

139. Mol C.D., Arvai A.S., Sanderson R.J., Slupphaug G., Kavli B., Krokan H.E., Mosbaugh D.W., Tainer J.A. (1995) Crystal structure of human uracil-DNA glycosylase in complex with a protein inhibitor: protein mimicry of DNA. Cell. 82, 701-8.

140. Carrington B., Weinstein R.N., Sood R. BE4max and AncBE4max Are Efficient in Germline Conversion of C:G to T:A Base Pairs in Zebrafish // Cells. – 2020. – V. 9. – P. 1690.

141. Tan J., Zhang F., Karcher D., Bock R. Engineering of high-precision base editors for site-specific single nucleotide replacement // Nat Commun. – 2019. – V. 10. – P. 439.

142. Keegan L.P., Hajji K., O'Connell M.A. Adenosine Deaminase Acting on RNA (ADAR) Enzymes: A Journey from Weird to Wondrous // Acc Chem Res. – 2023. – V.56. – P. 3165-3174.

143. Soma A., Kubota A., Tomoe D., Ikeuchi Y., Kawamura F., Arimoto H., Shiwa Y., Kanesaki Y., Nanamiya H., Yoshikawa H., Suzuki T., Sekine Y. yaaJ, the tRNA-Specific Adenosine Deaminase, Is Dispensable in Bacillus subtilis // Genes (Basel). – 2023. – V. 14. – P. 1515.

144. Ranzau B.L., Rallapalli K.L., Evanoff M., Paesani F., Komor A.C. The Wild-Type tRNA Adenosine Deaminase Enzyme TadA Is Capable of Sequence-Specific DNA Base Editing // Chembiochem. – 2023. – V. 24. – P. e202200788.

145. Kim J., Malashkevich V., Roday S., Lisbin M., Schramm V.L., Almo S.C. Structural and kinetic characterization of Escherichia coli TadA, the wobble-specific tRNA deaminase // Biochemistry. – 2006. – V. 45. – P. 6407-6416.

146. Shi K., Carpenter M.A., Banerjee S., Shaban N.M., Kurahashi K., Salamango D.J., McCann J.L., Starrett G.J., Duffy J.V., Demir Ö., Amaro R.E., Harki D.A., Harris R.S., Aihara H. Structural basis for targeted DNA cytosine deamination and mutagenesis by APOBEC3A and APOBEC3B // Nat Struct Mol Biol. – 2016. – V. 24. – P. 131-139.

147. Gaudelli N.M., Komor A.C., Rees H.A., Packer M.S., Badran A.H., Bryson D.I., Liu D.R. Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage // Nature. – 2018. – V. 551. – P. 464-471.

148. Grünewald J., Zhou R., Iyer S., Lareau C.A., Garcia S.P., Aryee M.J., Joung J.K. CRISPR DNA base editors with reduced RNA off-target and self-editing activities // Nat Biotechnol. – 2019. – V. 37 – P. 1041-1048.

149. Rallapalli K.L., Komor A.C., Paesani F. Computer simulations explain mutation-induced effects on the DNA editing by adenine base editors // Sci Adv. – 2020. – V. 6. – P. eaaz2309.

150. Arantes P.R., Chen X., Sinha S., Saha A., Patel A.C., Sample M., Nierzwicki Ł., Lapinaite A., Palermo G. Dimerization of the deaminase domain and locking interactions with Cas9 boost base editing efficiency in ABE8e // Nucleic Acids Res. – 2024. – V. 52. – P. 13931-13944.

151. Evanoff M., Korpal S., Krill Z.D., Cowan Q.T., Komor A.C. Precise, minimally evolved adenine base editors generated through mutation reversion analysis // Nat Biotechnol. – 2026. Mar 18. doi: 10.1038/s41587-026-03045-z. Epub ahead of print. PMID: 41851306.

152. Zhang C., Chen Z., Cao J., Zhang Z., Li W.K., Zhang X., Chen J., Liu J., Yuan Z., Gao F., Shi Z., Zhao X.M., Chen J., Zhan C., Cheng T.L. RNA inosine sensor-guided TadA mutational scanning for toxicity minimization of adenine base editors // Mol Ther. – 2026. – V. 34. – P. 380-396.

153. 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. Dual base editor catalyzes both cytosine and adenine base conversions in human cells // Nature Biotechnology. – 2020. – V. 38. – P. 856–860.

154. Lunev E., Karan A., Egorova T., Bardina M. Adeno-Associated Viruses for Modeling Neurological Diseases in Animals: Achievements and Prospects // Biomedicines. – 2022. – V. 10. – P. 1140.

155. Wang J.H., Gessler D.J., Zhan W., Gallagher T.L., Gao G. Adeno-associated virus as a delivery vector for gene therapy of human diseases // Sig Transduct Target Ther. – 2024. – V. 9. – P. 78.

156. Anzalone A.V., Randolph P.B., Davis J.R., Sousa A.A., Koblan L.W., Levy J.M., Chen P.J., Wilson C., Newby G.A., Raguram A., Liu D.R. Search-and-replace genome editing without double-strand breaks or donor DNA // Nature. – 2019. – V. 576. – P. 149–157.

157. Shuto Y., Nakagawa R., Zhu S., Hoki M., Omura S.N., Hirano H., Itoh Y., Zhang F., Nureki O. Structural basis for pegRNA-guided reverse transcription by a prime editor // Nature. – 2024. – V. 631. – P. 224–231.

158. 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. Enhanced prime editing systems by manipulating cellular determinants of editing outcomes // Cell. – 2021. – V. 184. – P. 5635-5652.e29.

159. Doman J.L., Pandey S., Neugebauer M.E., An M., Davis J.R., Randolph P.B., McElroy A., Gao X.D., Raguram A., Richter M.F., Everette K.A., Banskota S., Tian K., Tao Y.A., Tolar J., Osborn M.J., Liu D.R. Phage-assisted evolution and protein engineering yield compact, efficient prime editors // Cell. – 2023. – V. 186. – P. 3983-4002.e26.

160. Yan J., Oyler-Castrillo P., Ravisankar P., Ward C.C., Levesque S., Jing Y., Simpson D., Zhao A., Li H., Yan W., Goudy L., Schmidt R., Solley S.C., Gilbert L. A., Chan M.M., Bauer D.E., Marson A., Parsons, L.R., Adamson B. Improving prime editing with an endogenous small RNA-binding protein // Nature. – 2024. – V. 628. – P. 639–647.

161. Li F., Hou H., Song M., Chen Z., Peng T., Du Y., Zhao Y., Li J., Miao C. Targeted insertion of large DNA fragments through template-jumping prime editing in rice // Plant Biotechnol J. – 2025. – V. 23. – P. 2645-2647.

162. Wang J., He Z., Wang G., Zhang R., Duan J., Gao P., Lei X., Qiu H., Zhang C., Zhang Y., Yin H. Efficient targeted insertion of large DNA fragments without DNA donors // Nat Methods. – 2022. – V. 19. – P. 331-340.

163. Banskota S., Raguram A., Suh S., Du S.W., Davis J.R., Choi E.H., Wang X., Nielsen S.C., Newby G.A., Randolph P.B., Osborn M.J., Musunuru K., Palczewski K., Liu D.R. Engineered virus-like particles for efficient in vivo delivery of therapeutic proteins // Cell. – 2022. – V. 185. – P. 250–265.e16.

164. Borovikova S.E., Shepelev M.V., Mazurov D.V., Kruglova N.A. Efficient Genome Editing Using 'NanoMEDIC' AsCas12a-VLPs Produced with Pol II-Transcribed crRNA // Int. J. Mol. Sci. – 2024. – V. 25. – P. 12768.

165. ter Brake O., Konstantinova P., Ceylan M., Berkhout B. Silencing of HIV-1 with RNA interference: A multiple shRNA approach // Mol. Ther. – 2006. – V. 14. – P. 883–892.

166. Kozak M. An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs // Nucleic Acids Res. – 1987. – V. 15. – P. 8125-8148.

167. Yin S., Zhang M., Liu Y., Sun X., Guan Y., Chen X., Yang L., Huo Y., Yang J., Zhang X., Han H., Zhang J., Xiao M.M., Liu M., Hu J., Wang L., Li D. Engineering of efficiency-enhanced Cas9 and base editors with improved gene therapy efficacies // Mol. Ther. – 2023. – V. 31. – P. 744–759.

168. Kempton H.R., Love K.S., Guo L.Y., Qi L.S. Scalable biological signal recording in mammalian cells using Cas12a base editors // Nat. Chem. Biol. – 2022. – V. 18. – P. 742–750.

169. Standage-Beier K., Tekel S.J., Brookhouser N., Schwarz G., Nguyen T., Wang X., Brafman D.A. A transient reporter for editing enrichment (TREE) in human cells // Nucleic Acids Res. – 2019. – V. 47. – P. e120.

170. Jiang L., Long J., Yang Y., Zhou L., Su J., Qin F., Tang W., Tao R., Chen Q., Yao S. Internally inlaid SaCas9 base editors enable window specific base editing // Theranostics. – 2022. – V. 12. – P. 4767–4778.

171. Nguyen Tran M.T., Mohd Khalid M.K.N., Wang Q., Walker J.K.R., Lidgerwood G.E., Dilworth K.L. Lisowski L., Pébay A., Hewitt A.W. Engineering domain-inlaid SaCas9 adenine base editors with reduced RNA off-targets and increased on-target DNA editing // Nat. Commun. – 2020. – V. 11. – P. 4871.

172. Zhang X., Tao C., Li M., Zhang S., Liang P., Huang Y., Liu H., Wang Y. Engineering of SauriCas9 with enhanced specificity // Mol. Ther. Nucleic Acids. – 2025. – V. 36. – P. 102455.

173. Li X., Wang Y., Liu Y., Yang B., Wang X., Wei J., Lu Z., Zhang Y., Wu J., Huang X., Yang L., Chen J. Base editing with a Cpf1–cytidine deaminase fusion // Nat. Biotechnol. – 2018. – V. 36. – P. 324–327.

174. Chen F., Lian M., Ma B., Gou S., Luo X., Yang K., Shi H., Xie J., Ge W., Ouyang Z., Lai C., Li N., Zhang Q., Jin Q., Liang Y., Chen T., Wang J., Zhao X., Li L., Yu M., Ye Y., Wang K., Wu H., Lai L. Multiplexed base editing through Cas12a variant-mediated cytosine and adenine base editors // Commun. Biol. – 2022. – V. 5. – P. 1163.

175. Yoon D.E., Kim N.R., Park S.J., Jeong T.Y., Eun B., Cho Y., Lim S.Y., Lee H., Seong J.K., Kim K. Precise base editing without unintended indels in human cells and mouse primary myoblasts // Exp. Mol. Med. – 2023. -V. 55. – P. 2586–2595.

176. Zeng D., Zheng Z., Liu Y., Liu T., Li T., Liu J., Luo Q., Xue Y., Li S., Chai N., Yu S., Xie X., Liu Y.G., Zhu Q. Exploring C-to-G and A-to-Y Base Editing in Rice by Using New Vector Tools // Int J Mol Sci. – 2022. – V. 23. – P. 7990.

177. An M., Raguram A., Du S.W., Banskota S., Davis J.R., Newby G.A., Chen P.Z., Palczewski K., Liu D.R. Engineered virus-like particles for transient delivery of

prime editor ribonucleoprotein complexes in vivo // Nat Biotechnol. – 2024. – V. 42. – P. 1526–1537.

178. Jia J., Hao Y., Zhang L., Cao X., An L., Wang H., Ma Q., Jin X., Ma X. Development and validation of optimized lentivirus-like particles for gene editing tool delivery with Gag-Only strategy // Eur J Med Res. – 2025. – V. 30. – P. 242.

179. Raguram A., An M., Chen P.Z., Liu DR. Directed evolution of engineered virus-like particles with improved production and transduction efficiencies // Nat Biotechnol. – 2025. – V. 43. – P. 1635–1647.

180. Moore-Kelly C., Reddem R., Alberts B.M., Wright J., Evans T., Kulkarni A., Clarkson N.G., Farley D.C., Mitrophanous K.A., Saraiva Raposo R.A. Enhancing titers of therapeutic lentiviral vectors using PKC agonists // Mol Ther Methods Clin Dev. – 2025. – V. 33. – P. 101484.

181. Shen C.F., Tremblay S., Sabourin-Poirier C., Burney E., Broussau S., Manceur A., Rodenbrock A., Voyer R., Loignon M., Ansorge S., Gilbert R. Culture media selection and feeding strategy for high titer production of a lentiviral vector by stable producer clones cultivated at high cell density // Bioprocess Biosyst Eng. – 2022. – V. 45. – P. 1267-1280.

182. Abramson J., Adler J., Dunger J., Evans R., Green T., Pritzel A., Ronneberger O., Willmore L., Ballard A.J., Bambrick J., Bodenstein S.W., Evans D.A., Hung C.C., O'Neill M., Reiman D., Tunyasuvunakool K., Wu Z., Žemgulytė A., Arvaniti E., Beattie C., Bertolli O., Bridgland A., Cherepanov A., Congreve M., Cowen-Rivers A.I., Cowie A., Figurnov M., Fuchs F.B., Gladman H., Jain R., Khan Y.A., Low C.M.R., Perlin K., Potapenko A., Savy P., Singh S., Stecula A., Thillaisundaram A., Tong C., Yakneen S., Zhong E.D., Zielinski M., Žídek A., Bapst V., Kohli P., Jaderberg M., Hassabis D., Jumper J.M. Accurate structure prediction of biomolecular interactions with AlphaFold 3 // Nature. – 2024. – V. 630. – P. 493-500.

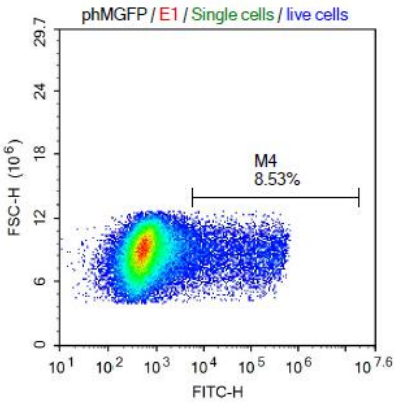
183. Kuratani M., Ishii R., Bessho Y., Fukunaga R., Sengoku T., Shirouzu M., Sekine S., Yokoyama S. Crystal structure of tRNA adenosine deaminase (TadA) from Aquifex aeolicus // J Biol Chem. – 2005. – V. 280. – P. 16002-16008.

184. Wu H., Sun Y., Wang Y., Luo L., Song Y. Advances in miniature CRISPR-Cas proteins and their applications in gene editing // Arch Microbiol. – 2024. – V. 206. – P. 231.

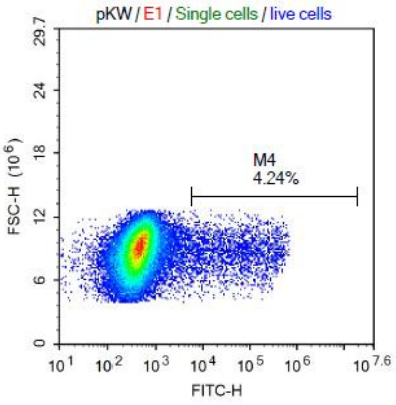
185. Nishimasu H., Ran F.A., Hsu P.D., Konermann S., Shehata S.I., Dohmae N., Ishitani R., Zhang F., Nureki O. Crystal structure of Cas9 in complex with guide RNA and target DNA // Cell. – 2014. – V. 156. – P. 935-949.

ПРИЛОЖЕНИЕ

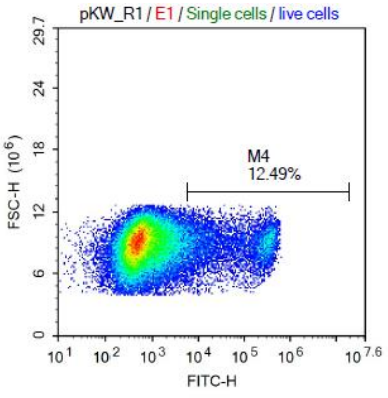
Приложение 1. Диаграммы рассеяния (дот-плоты), полученные при анализе методом проточной цитометрии клеток HEK293T, трансфицированных скрининговыми плазмидами phMGFP, pKW, pKW_R1, pKW_R3, pKW_R1m, pKW_R3m.



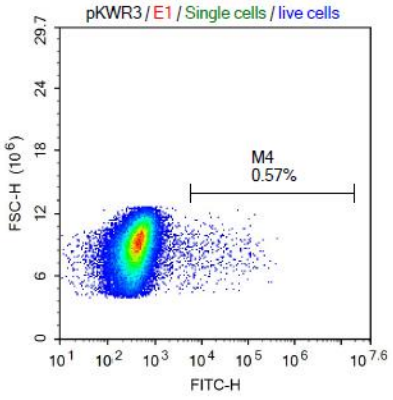
Gate	Count	% live cells	Median X
live cells	46,965	100.00%	578
M4	4,004	8.53%	49,961



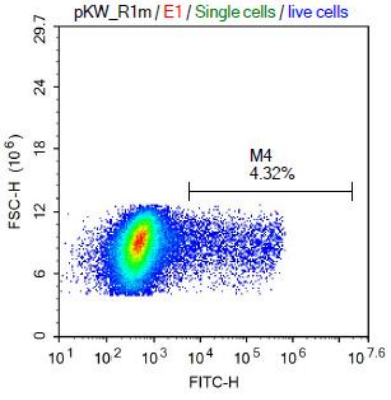
Gate	Count	% live cells	Median X
live cells	49,107	100.00%	435
M4	2,080	4.24%	51,821



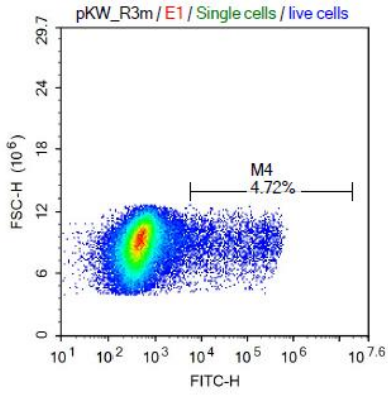
Gate	Count	% live cells	Median X
live cells	48,806	100.00%	637
M4	6,098	12.49%	64,388



Gate	Count	% live cells	Median X
live cells	49,880	100.00%	398
M4	286	0.57%	20,977



Gate	Count	% live cells	Median X
live cells	48,591	100.00%	476
M4	2,101	4.32%	50,152



Gate	Count	% live cells	Median X
live cells	48,401	100.00%	465
M4	2,284	4.72%	48,188

Приложение 2. Список олигонуклоетидов, использованных для сборки гена MmCas12m.

1aC12m	AGCCGGAAGCATAAGATGGACGAGCTGTACAAGG
2aC12m	CGCCGCTTCCGGATCCCTTGTACAGCTCGTCCAT
3aC12m	GATCCGGAAGCGGCGCCACCAATTTCAGCCTGCT
4aC12m	CGTCGCCAGCCTGTTTCAGCAGGCTGAAATTGGTGG
5aC12m	GAAACAGGCTGGCGACGTGGAAGAAAACCCTGGGC
6aC12m	CCACTTTCCGCTTTTTCTTGGGCCCAGGGTTTTCTTCCA
7aC12m	CCAAGAAAAAGCGGAAAGTGGGGCCCAGCGGCACC
8aC12m	CCATTGTATGCACTGTCATGGTGGTGCCGCTGGGCC
9aC12m	ACCATGACAGTGCATACAATGGGCGTGCACTACAAGTGGC
10aC12m	CTCAGAACCTCGGGGATCTGCCACTTGTAGTGCACGC
11aC12m	AGATCCCCGAGGTTCTGAGACAGCAGCTGTGGCTG
12aC12m	CTTCTCTCAGGTTGTGGGCCAGCCACAGCTGCTGT
13aC12m	GCCCACAACCTGAGAGAAGATCTGGTGTCTCTGCAGCT
14aC12m	GGTCGTCGTCGTAGGCCAGCTGCAGAGACACCAGAT
15aC12m	GGCCTACGACGACGACCTGAAGGCCATCTGGTCTAGC
16aC12m	CTGGGCCACATCAGGGTAGCTAGACCAGATGGCCTTCA
17aC12m	TACCCTGATGTGGCCCAGGCCGAGGATACAATGGCTG
18aC12m	GGCATCAGCTTCAGCGGCAGCCATTGTATCCTCGGC
19aC12m	CCGCTGAAGCTGATGCCGTGGCTCTGTCTGAGAGAGT
20aC12m	CTCGATTCTGGCCTGCTTCACTCTCTCAGACAGAGCCAC
21aC12m	GAAGCAGGCCAGAATCGAGGCCAGAAGCAAGAAGATCAG
22aC12m	CTGGGTCAGCTCGGTGCTGATCTTCTTGCTTCTGGC
23aC12m	CACCGAGCTGACCCAGCAGCTGAGGGACGCCAA
24aC12m	GGCGTCCTTCAGTCTCTTCTTGGCGTCCCTCAGCTG
25aC12m	GAAGAGACTGAAGGACGCCAGACAGGCCAGACGGG
26aC12m	GACCACGGCAATGGCATCCCGTCTGGCCTGTCT
27aC12m	ATGCCATTGCCGTGGTCAAAGATGATGCCGCCGAG
28aC12m	CAGATCTGGCCTTTCTCCGCTCGGCGGCATCATCTTT
29aC12m	CGGAGAAAGGCCAGATCTGATCAACTGGCCGCCGA
30aC12m	CCGTACAGGGCTTTCTGGTCGGCGGCCAGTTGAT
31aC12m	CCAGAAAGCCCTGTACGGCCAGTACTGCAGAGATGGC
32aC12m	AGCTGGCCCAGTACAGATCGCCATCTCTGCAGTACTGG
33aC12m	GATCTGTACTGGGCCAGCTTCAACACCGTGCTGGACC
34aC12m	ACGGCGGTCTTGTGGTGGTCCAGCACGGTGTTGA
35aC12m	ACCACAAGACCGCCGTGAAGAGAATCGCCGCTCAG
36aC12m	GGCTTGCCAGAGGCTCTCTGAGCGGCGATTCTCTTC
37aC12m	AGAGCCTCTGGCAAGCCTGCCACACTGAGACACCA
38aC12m	CCGCTGCCGTCAAATCTGTGGTGTCTCAGTGTGGCA
39aC12m	CAGATTTGACGGCAGCGGCACAATCGCCGTGCAACT
40aC12m	GCGCCAGCTTGTCTTTGGAGTTGCACGGCGATTGTG
41aC12m	CCAAAGACAAGCTGGCGCCCCTCCTCGGACACCTATG
42aC12m	CGGCCTCATCAGCAAGAACCATAGGTGTCCGAGGAGGG
43aC12m	GTTCTTGCTGATGAGGCCGGCAAGTACCGGAACGTGC
44aC12m	TCCAGCCAGGAATGTGCAGCACGTTCCGGTACTTGC
45aC12m	TGCACATTCCTGGCTGGACCGATCCTGATGTGTGGGA

46aC12m TGGGACCGTGTCATCTGCTCCCACACATCAGGATCGG
47aC12m GCAGATGACACGGTCCCAGTGCAGACAGTCTGGCAG
48aC12m TCTCATCCGCACTGTCACTCTGCCAGACTGTCTGCAC
49aC12m AGTGACAGTGCGGATGAGATGCGGCAGCACAGATG
50aC12m CGATCCACTGGGGCTGTCCATCTGTGCTGCCGCA
51aC12m GACAGCCCCAGTGGATCGATCTGCCTGTGCAGGTTC
52aC12m GCAGGCAGCCATCTGTGAACCTGCACAGGCAGAT
53aC12m ACAGATGGCTGCCTGCCGACGCCGATATTACAGGC
54aC12m TGTGACCACCAGCTCAGCGCCTGTAATATCGGCGTCG
55aC12m GCTGAGCTGGTGGTCACAAGAGTGGCCGGAATCTACA
56aC12m CTCTGTAGATTCCGGCCACTCT
1bC12m AGAGTGGCCGGAATCTACAGAGCCAAGCTG
2bC12m CCTGGCGGTCACGCACAGCTTGGCTCTGTAGATTCC
3bC12m TGCGTGACCGCCAGGATCGGAGATACCGAGCCTG
4bC12m ACTGTAGGGCCAGATGTCACAGGCTCGGTATCTCCGAT
5bC12m TGACATCTGGCCCTACAGTGGCCCTGCATCTTGGC
6bC12m CCTCTTCGGTAGATCGCCAGCCAAGATGCAGGGCC
7bC12m TGGCGATCTACCGAAGAGGGAACAGCCGTGGCCAC
8bC12m AGGGGCATCAGATCTCCAAGTGGCCACGGCTGTTC
9bC12m TTGGAGATCTGATGCCCCTCTGGACATCCCCTTCGGC
10bC12m CCACTCTCATCACTGTTCTCAGGCCGAAGGGGATGTCCAG
11bC12m CTGAGAACAGTGATGAGAGTGGATGCTGCCGGCACCT
12bC12m GGCACCACAATGATGCCAGAGGTGCCGGCAGCAT
13bC12m CTGGCATCATTGTGGTGCCTGCCACCATCGAGAGAAGG
14bC12m GATATTCTCTGTCCGGGTCAGCCTTCTCTCGATGGTGGCA
15bC12m CTGACCCGGACAGAGAATATCGCCAGCAGCAGATCTCTG
16bC12m CAGGGCGTCCAGGGCCAGAGATCTGCTGCTGGC
17bC12m GCCCTGGACGCCCTGAGAGATAAGGTTGTCGGCTG
18bC12m CGTCGTTGTCGCTCAGCCAGCCGACAACCTTATCTCT
19bC12m GCTGAGCGACAACGACGCCCCTACCTATAGAGATGCTCC
20bC12m GGTGGCGGCTTCCAGTGGAGCATCTCTATAGGTAGGGG
21bC12m ACTGGAAGCCGCCACCGTGAAGCAGTGGAAGTCCC
22bC12m CAGAGAGGCGAATCTCTGAGGGGACTTCCACTGCTTCAC
23bC12m CTCAGAGATTCGCCTCTCTGGCTCACGCCTGGAAGGA
24bC12m GCTGATCTCGGTGCCATTGTCCTTCCAGGCGTGAGC
25bC12m CAATGGCACCGAGATCAGCGATATCCTGTGGGCCTGGT
26bC12m TGCTTTCTGTCCAGGCTGAACCAGGCCCACAGGATATC
27bC12m TCAGCCTGGACAGAAAGCAGTGGGCTCAGCAAGAGAAC
28bC12m CAGGGCCTTTCTTCTGCCGTTCTCTTGCTGAGCCCAC
29bC12m GGCAGAAGAAAGGCCCTGGGACACAGAGATGACCTGTATC
30bC12m CGGCGGCGATCTGCCGATACAGGTCATCTCTGTGTCC
31bC12m GGCAGATCGCCGCCGTGATCTCTGATCAGGCTGGA
32bC12m TCGTCCACCAGCACGTGTCCAGCCTGATCAGAGATCA
33bC12m CACGTGCTGGTGGACGATACAAGTGTGGCCGAGCT
34bC12m TCCATAGCTCTGGCGCTCAGCTCGGCCACACTTGTA
35bC12m GAGCGCCAGAGCTATGGAAAGAACCGAGCTGCCTACC
36bC12m CGATCTTCTGCTGCACCTCGGTAGGCAGCTCGGTTCTT
37bC12m GAGGTGCAGCAGAAGATCGACAGAAGAAGGGACCACGC

38bC12m CAGTCCACCGGGAGCAGCGTGGTCCCTTCTTCTGT
39bC12m TGCTCCCGGTGGACTGAGAGCTTCTGTGGTGGCC
40bC12m CCCCATCTCTTGTCATGGCGGCCACCACAGAAGCTCT
41bC12m GCCATGACAAGAGATGGGGTGCCAGTGACAATCGTGG
42bC12m GGTGAAATCGGCTGCGGCCACGATTGTCACTGGCA
43bC12m CCGCAGCCGATTTCACCCGGACACACAGCAGATGC
44bC12m CGGGGTTCACGTGTCCGCATCTGCTGTGTGTCCG
45bC12m GGACACGTGAACCCCGCCGACGATAGATACCTGAGC
46bC12m TCGCATCTGACGGGGTTGCTCAGGTATCTATCGTCGG
47bC12m AACCCCGTCAGATGCGACGGCTGTGGCGCTATG
48bC12m AAGGATCTGTCCTGGTCGTACATAGCGCCACAGCCG
49bC12m TACGACCAGGACAGATCCTTCGTGACCCTGATGCTGAG
50bC12m GGAGCTGTGGCGGCTCTCAGCATCAGGGTCACG
51bC12m AGCCGCCACAGCTCCTAGCAATCCTGGCAAAAGAC

Приложение 3. Список олигонуклоетидов, использованных для сборки рекомбинантного гена TadA-8e – TadDE.

1TdA	ACAATTTCATTAAAGAGGAGAAATAATAAAT
2TdA	GGCTGAATTCCACTTCGCTCATTTATTATTTCTCCTCTTTAATGAAATTGT
3T7a	ATCAATATCGCCAAGAACGACTTCAGCGATATCGAGCTGGC
3TdA	GAGCGAAGTGGAATTCAGCCATGAATATTGGATGCGTCATGC
4T7a	GGTGTTGAAAGGAATGGCGGCCAGCTCGATATCGCTGA
4TdA	CGTTTTGCCAGGGTCAGTGCATGACGCATCCAATATTCAT
5T7a	CGCCATTCCTTTCAACACCCTGGCCGATCACTACGGC
5TdA	ACTGACCCTGGCAAAACGTGCCCGTGATGAAGGTGA
6T7a	CTGTTCTCTAGCCAGTCTCTCGCCGTAGTGATCGGCCAG
6TdA	TGCACCAACCGGTGCTTCACCTTCATCACGGGCA
7T7a	GAGAGACTGGCTAGAGAACAGCTGGCCCTGGAACACGA
7TdA	AGCACCGGTTGGTGCAGTTCTGGTTCTGAATAATCGTGTTA
8T7a	CTTCGCCCATCTCGTAGTACTCGTGTTCCAGGGCCAG
8TdA	GACGATTCCAACCTTCACCAATAACACGATTATTCAGAACCAGAAC
9T7a	GTACTACGAGATGGGCGAAGCCCGGTTCCGGAAGATGT
9TdA	TTGGTGAAGGTTGGAATCGTCGTATTGGTCTGCATGATCCGA
10TdA	CCATAATTTCTGCATGTGCGGTCGGATCATGCAGACCAATAC
11TdA	CCGCACATGCAGAAATTATGGCACTGCGTCAAGGTGGTC
12TdA	GACGGCTATTCTGCATAACCAGACCACCTTGACGCAGTG
13TdA	TGGTTATGCAGAATAGCCGTCTGATTGATGCAACCCTGTATGT
14TdA	CATAACACACGGTTCAAAGGTAACATACAGGGTTGCATCAATCA
15TdA	TACCTTTGAACCGTGTGTTATGTGTGCCGGTGCAATGATTA
16TdA	CACGACCAATGCGGCTATTAATCATTGCACCGGCACA
17TdA	ATAGCCGCATTGGTCGTGTTGTTTTTGGTGTTCGTAATAGCA
18TdA	ACCTGCTGCACCACGTTTGCTATTACGAACACCAAAAACAA
19TdA	AACGTGGTGCAGCAGGTAGCCTGATGAATGTTCTGAATTATC
20TdA	AACACGGTGATTCATACCCGGATAATTCAGAACATTCATCAGGCT
21TdA	CGGGTATGAATCACCGTGTTGAAATTACCGAAGGTATTCTGGC
22TdA	GCAGTGCTGCACATTCATCTGCCAGAATACCTTCGGTAATTTC
23TdA	AGATGAATGTGCAGCACTGCTGTGTGATTTTTATCGTATGCCTC
24TdA	TCTGTGCGTTAAAAACCTGACGAGGCATACGATAAAAATCACACA
25TdA	GTCAGGTTTTTAACGCACAGAAAAAGGCACAGAGCAGCATTA

Приложение 4. Ампликоны, использованные для создания pMmCas12m-TadDE-sgRNA, psgRNA-MmCas12m (А), и pdSpCas9-TadDE-sgRNA, psgRNA-dSpCas9 (Б). Дополнительно представлена электрофореграмма продуктов амплификации: М – 100+bp ДНК-маркер. Целевые ампликоны выделены рамкой.

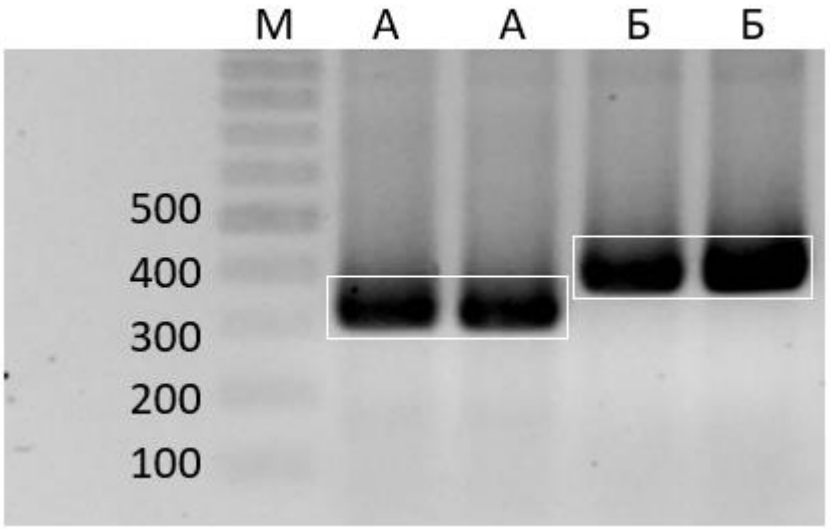
А

GAGGGCCTATTTCCCATGATTCCTTCATATTTGCATATACGATACAAGGCT
GTTAGAGAGATAATTGGAATTAATTTGACTGTAAACACAAAGATATTAGT
ACAAAATACGTGACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTT
TAAAATTATGTTTTAAAATGGACTATCATATGCTTACCGTAACTTGAAAGT
ATTTCGATTTCTTGGCTTTATATATCTTGTGGAAAGGACGAAACACCGTGT
CATAGCCCAGCTTGGCGGGCGAAGGCCAAGACGGAGACCGACACAAGCT
TCAGTGGTCTCTTTTTTTTT

Б

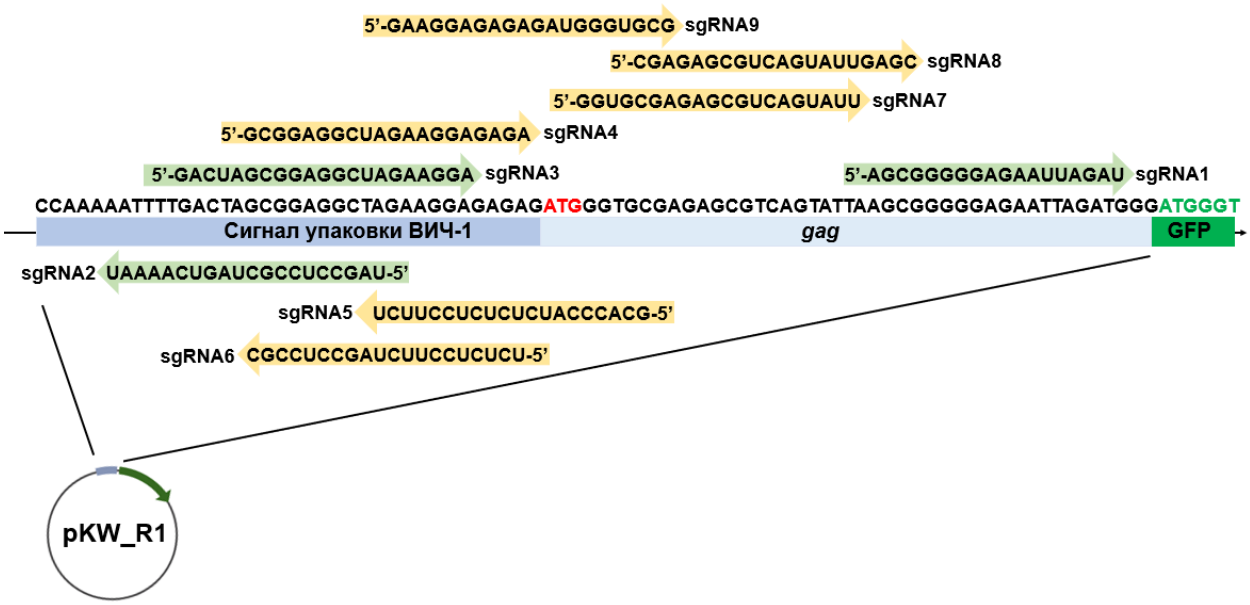
GAGGGCCTATTTCCCATGATTCCTTCATATTTGCATATACGATACAAGGCT
GTTAGAGAGATAATTGGAATTAATTTGACTGTAAACACAAAGATATTAGT
ACAAAATACGTGACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTT
TAAAATTATGTTTTAAAATGGACTATCATATGCTTACCGTAACTTGAAAGT
ATTTCGATTTCTTGGCTTTATATATCTTGTGGAAAGGACGAAACACCGGA
GACCGACACAAGCTTCAGTGGTCTCTGTTTAAGAGCTATGCTGGAAACAG
CATAGCAAGTTTAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCA
CCGAGTCGGTGC

Черный = U6-промотор, Фиолетовый = tracrRNA для MmCas12m, Голубой = tracrRNA для dSpCas9, Красный = участок с сайтами эндонуклеазы рестрикции *BsaI* для клонирования crRNA.

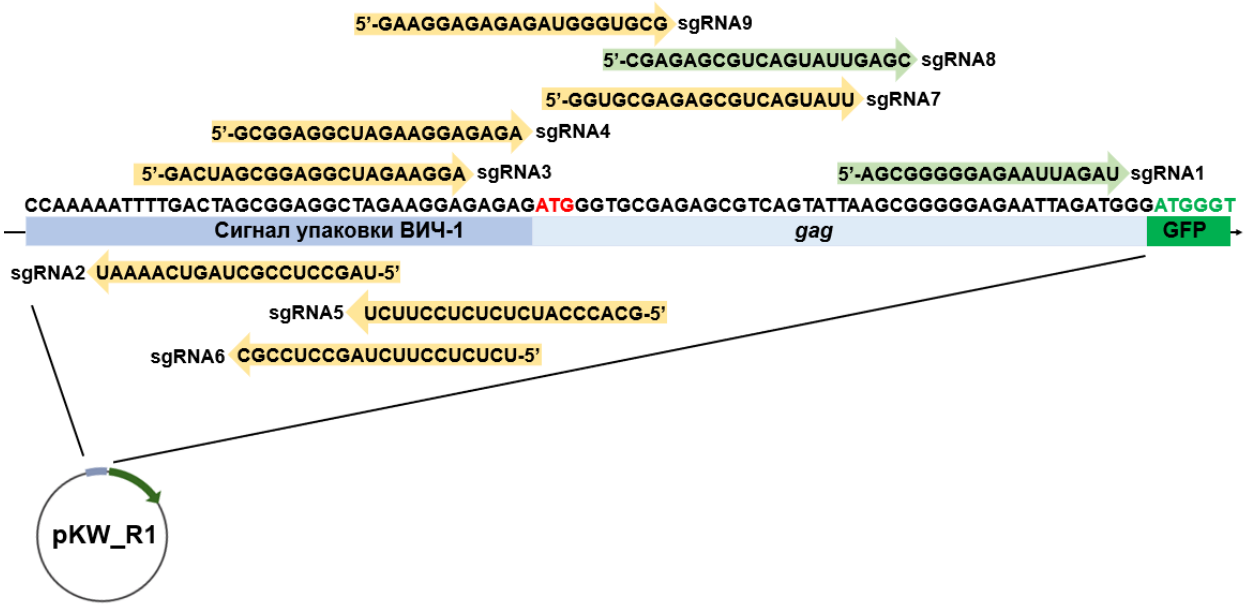


Приложение 5. Выравнивание sgRNA на область старт-кодона гена *gag*. sgRNA для MmCas12m (А) и dSpCas9 (Б), примыкающие к оптимальному PAM, выделены зеленым, к неоптимальному PAM – желтым. Старт-кодон гена *gag* выделен красным.

А



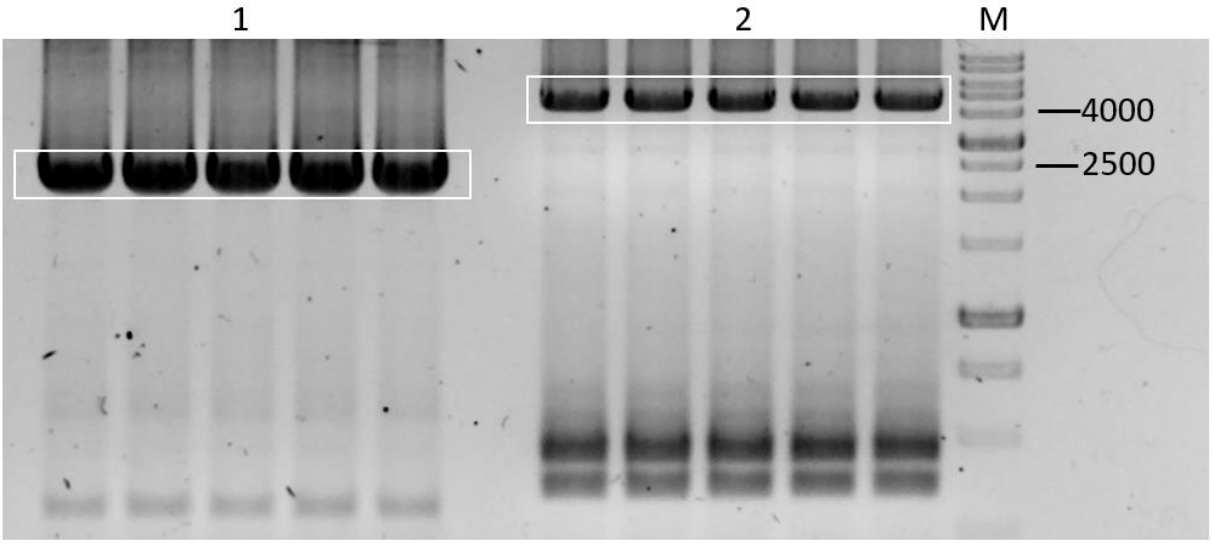
Б



Приложение 6. Применение системы скрининга для тестирования sgRNA1-9 при таргетировании MmCas12m и dSpCas9 на область старт-кодона гена *gag* провируса ВИЧ-1. Сравнение доли GFP-позитивных клеток при котрансфекции pKW_R1 и pMmCas12m-sgRNA*N* или pdSpCas9-sgRNA*N*. K- – группа клеток, трансфицированная pKW_R1 и pMmCas12m-sgRNA или pdSpCas9-sgRNA без sgRNA. Представлено среднее значение с указанием доверительного интервала. Размер выборки n = 3 (биологические повторы). *$p < 0,05$. sgRNA, примыкающие к оптимальному PAM, выделены зеленым, к неоптимальному PAM – желтым.

Исходное значение (K-) ≈49,15% (±0.37)	MmCas12m	dSpCas9
sgRNA1	≈39,21% (±0,82)	≈37,03% (±0,35)
sgRNA2	≈41,1% (±0,29)	≈40,07% (±0,32)
sgRNA3	≈40,15% (±0,46)	≈39,59% (±0,84)
sgRNA4	≈39,43% (±0,52)	≈49,84% (±0,36)
sgRNA5	≈37,4% (±0,42)	≈42,54% (±0,37)
sgRNA6	≈49,21% (±0,37)	≈49,2% (±0,76)
sgRNA7	≈43,1% (±0,35)	≈40,5% (±0,62)
sgRNA8	≈43,32% (±0,28)	≈24,42% (±1,65)
sgRNA9	≈43,01% (±0,29)	≈49,12% (±0,33)

Приложение 7. Электрофореграмма продуктов амплификации, кодирующих MmCas12m-TadDE (1) и dSpCas9-TadDE (2); M – 1kb ДНК-маркер. Целевые ампликоны выделены рамкой.



Приложение 8. Нуклеотидная последовательность, кодирующая двойной редактор оснований MmCas12m-TadDE.

GGGGCCCAAGAAAAAGCGGAAAGTGGGGCCCAGCGGCACCACCATGACAGTGCATAC
AATGGGCGTGCACTACAAGTGGCAGATCCCCGAGGTTCTGAGACAGCAGCTGTGGCTG
GCCCACAACCTGAGAGAAGATCTGGTGTCTCTGCAGCTGGCCTACGACGACGACCTGA
AGGCCATCTGGTCTAGCTACCCTGATGTGGCCCAGGCCGAGGATACAATGGCTGCCGC
TGAAGCTGATGCCGTGGCTCTGTCTGAGAGAGTGAAGCAGGCCAGAATCGAGGCCAGA
AGCAAGAAGATCAGCACCGAGCTGACCCAGCAGCTGAGGGACGCCAAGAAGAGACTG
AAGGACGCCAGACAGGCCAGACGGGATGCCATTGCCGTGGTCAAAGATGATGCCGCC
GAGCGGAGAAAGGCCAGATCTGATCAACTGGCCGCCGACCAGAAAGCCCTGTACGGC
CAGTACTGCAGAGATGGCGATCTGTACTGGGCCAGCTTCAACACCGTGCTGGACCACC
ACAAGACCGCCGTGAAGAGAATCGCCGCTCAGAGAGCCTCTGGCAAGCCTGCCACACT
GAGACACCACAGATTTGACGGCAGCGGCACAATCGCCGTGCAACTCCAAAGACAAGCT
GGCGCCCCTCCTCGGACACCTATGGTTCTTGCTGATGAGGCCGGCAAGTACCGGAACG
TGCTGCACATTCCTGGCTGGACCGATCCTGATGTGTGGGAGCAGATGACACGGTCCCA
GTGCAGACAGTCTGGCAGAGTGACAGTGCGGATGAGATGCGGCAGCACAGATGGACA
GCCCCAGTGGATCGATCTGCCTGTGCAGGTTCACAGATGGCTGCCTGCCGACGCCGAT
ATTACAGGCGCTGAGCTGGTGGTCACAAGAGTGGCCGGAATCTACAGAGCCAAGCTGT
GCGTGACCGCCAGGATCGGAGATACCGAGCCTGTGACATCTGGCCCTACAGTGGCCCT
GCATCTTGGCTGGCGATCTACCGAAGAGGGAACAGCCGTGGCCACTTGGAGATCTGAT
GCCCCTCTGGACATCCCCTTCGGCCTGAGAACAGTGATGAGAGTGGATGCTGCCGGCA
CCTCTGGCATCATTGTGGTGCCTGCCACCATCGAGAGAAGGCTGACCCGGACAGAGAA
TATCGCCAGCAGCAGATCTCTGGCCCTGGACGCCCTGAGAGATAAGGTTGTCGGCTGG
CTGAGCGACAACGACGCCCCTACCTATAGAGATGCTCCACTGGAAGCCGCCACCGTGA
AGCAGTGGAAGTCCCCTCAGAGATTCGCCTCTCTGGCTCACGCCTGGAAGGACAATGG
CACCGAGATCAGCGATATCCTGTGGGCCTGGTTCAGCCTGGACAGAAAGCAGTGGGCT
CAGCAAGAGAACGGCAGAAGAAAGGCCCTGGGACACAGAGATGACCTGTATCGGCAG
ATCGCCGCCGTGATCTCTGATCAGGCTGGACACGTGCTGGTGGACGATACAAGTGTGG
CCGAGCTGAGCGCCAGAGCTATGGAAAGAACCGAGCTGCCTACCGAGGTGCAGCAGA
AGATCGACAGAAGAAGGGACCACGCTGCTCCCGGTGGACTGAGAGCTTCTGTGGTGGC
CGCCATGACAAGAGATGGGGTGCCAGTGACAATCGTGGCCGCAGCCGATTTCACCCGG
ACACACAGCAGATGCGGACACGTGAACCCCGCCGACGATAGATACCTGAGCAACCCC
GTCAGATGCGACGGCTGTGGCGCTATGTACGACCAGGACAGATCCTTCGTGACCCTGA
TGCTGAGAGCCGCCACAGCTCCTAGCAATCCTGGCAAAAGACCTGCCGCCACAAAGAA
GGCCGGACAGGCCAAAAAGAAAAAGGGTCTCAGCGGTGGTAGCAGCGGCGGGCCTAC
CAGCGAAGTGGAATTCAGCCATGAATATTGGATGCGTCATGCACTGACCCTGGCAAAA
CGTGCCCGTGATGAAGGTGAAGCACCGGTTGGTGCAGTTCTGGTTCTGAATAATCGTGT
TATTGGTGAAGGTTGGAATCGTCGTATTGGTCTGCATGATCCGACCGCACATGCAGAA
ATTATGGCACTGCGTCAAGGTGGTCTGGTTATGCAGAATAGCCGTCTGATTGATGCAAC
CCTGTATGTTACCTTTGAACCGTGTGTTATGTGTGCCGGTGCAATGATTAATAGCCGCA
TTGGTCGTGTTGTTTTTGGTGTTCGTAATAGCAAACGTGGTGCAGCAGGTAGCCTGATG
AATGTTCTGAATTATCCGGGTATGAATCACCGTGTTGAAATTACCGAAGGTATTCTGGC
AGATGAATGTGCAGCACTGCTGTGTGATTTTTATCGTATGCCTCGTCAGGTTTTTAACG
CACAGAAAAAGGCACAGAGCAGCATTAAT

Красный = сигнал ядерной локализации большого Т-антигена вируса SV40; Голубой = ген *Mycolicibacterium mucogenicum* Cas12m; Оранжевый = сигнал ядерной локализации нуклеоплазмина; Фиолетовый = глицин-сериновый линкер; Зеленый = ген рекомбинантной дезаминазы TadA-8e (TadDE).

Приложение 9. Нуклеотидная последовательность, кодирующая двойной редактор оснований dSpCas9-TadDE.

GGGGCCCAAGAAAAAGCGGAAAGTGGACAAGAAGTACAGCATCGGCCTGGCCATCGG
CACCAACTCTGTGGGCTGGGCCGTGATCACCGACGAGTACAAGGTGCCCAGCAAGAAA
TTCAAGGTGCTGGGCAACACCGACCGGCACAGCATCAAGAAGAACCTGATCGGAGCCC
TGCTGTTCGACAGCGGCGAAACAGCCGAGGCCACCCGGCTGAAGAGAACCGCCAGAA
GAAGATACACCAGACGGAAGAACCGGATCTGCTATCTGCAAGAGATCTTCAGCAACGA
GATGGCCAAGGTGGACGACAGCTTCTTCCACAGACTGGAAGAGTCCTTCCTGGTGGAA
GAGGATAAGAAGCACGAGCGGCACCCCATCTTCGGCAACATCGTGGACGAGGTGGCCT
ACCACGAGAAGTACCCCACCATCTACCACCTGAGAAAGAAACTGGTGGACAGCACCG
ACAAGGCCGACCTGCGGCTGATCTATCTGGCCCTGGCCCACATGATCAAGTTCCGGGG
CCACTTCCTGATCGAGGGCGACCTGAACCCCGACAACAGCGACGTGGACAAGCTGTTC
ATCCAGCTGGTGCAGACCTACAACCAGCTGTTCGAGGAAAACCCCATCAACGCCAGCG
GCGTGGACGCCAAGGCCATCCTGTCTGCCAGACTGAGCAAGAGCAGACGGCTGGAAA
ATCTGATCGCCCAGCTGCCCGGCGAGAAGAAGAATGGCCTGTTCGGAAACCTGATTGC
CCTGAGCCTGGGCCTGACCCCCAACTTCAAGAGCAACTTCGACCTGGCCGAGGATGCC
AAACTGCAGCTGAGCAAGGACACCTACGACGACGACCTGGACAACCTGCTGGCCCAG
ATCGGCGACCAGTACGCCGACCTGTTTCTGGCCGCCAAGAACCTGTCCGACGCCATCCT
GCTGAGCGACATCCTGAGAGTGAACACCGAGATCACCAAGGCCCCCCTGAGCGCCTCT
ATGATCAAGAGATACGACGAGCACCACCAGGACCTGACCCTGCTGAAAGCTCTCGTGC
GGCAGCAGCTGCCTGAGAAGTACAAAGAGATTTTCTTCGACCAGAGCAAGAACGGCTA
CGCCGGCTACATTGACGGCGGAGCCAGCCAGGAAGAGTTCTACAAGTTCATCAAGCCC
ATCCTGGAAAAGATGGACGGCACCGAGGAACTGCTCGTGAAGCTGAACAGAGAGGAC
CTGCTGCGGAAGCAGCGGACCTTCGACAACGGCAGCATCCCCCACCAGATCCACCTGG
GAGAGCTGCACGCCATTCTGCGGCGGCAGGAAGATTTTTACCCATTCCTGAAGGACAA
CCGGGAAAAGATCGAGAAGATCCTGACCTTCCGCATCCCCTACTACGTGGGCCCTCTG
GCCAGGGGAAACAGCAGATTCGCCTGGATGACCAGAAAGAGCGAGGAAACCATCACC
CCCTGGAACTTCGAGGAAGTGGTGGACAAGGGCGCTTCCGCCCAGAGCTTCATCGAGC
GGATGACCAACTTCGATAAGAACCTGCCCAACGAGAAGGTGCTGCCCAAGCACAGCCT
GCTGTACGAGTACTTCACCGTGTATAACGAGCTGACCAAAGTGAAATACGTGACCGAG
GGAATGAGAAAGCCCGCCTTCCTGAGCGGCGAGCAGAAAAAGGCCATCGTGGACCTG
CTGTTCAAGACCAACCGGAAAGTGACCGTGAAGCAGCTGAAAGAGGACTACTTCAAG
AAAATCGAGTGCTTCGACTCCGTGGAAATCTCCGGCGTGGAAGATCGGTTCAACGCCT
CCCTGGGCACATACCACGATCTGCTGAAAATTATCAAGGACAAGGACTTCCTGGACAA
TGAGGAAAACGAGGACATTCTGGAAGATATCGTGCTGACCCTGACACTGTTTGAGGAC
AGAGAGATGATCGAGGAACGGCTGAAAACCTATGCCCACCTGTTCGACGACAAAGTG
ATGAAGCAGCTGAAGCGGCGGAGATACACCGGCTGGGGCAGGCTGAGCCGGAAGCTG
ATCAACGGCATCCGGGACAAGCAGTCCGGCAAGACAATCCTGGATTTCCTGAAGTCCG
ACGGCTTCGCCAACAGAAACTTCATGCAGCTGATCCACGACGACAGCCTGACCTTTAA
AGAGGACATCCAGAAAGCCCAGGTGTCCGGCCAGGGCGATAGCCTGCACGAGCACAT
TGCCAATCTGGCCGGCAGCCCCGCCATTAAGAAGGGCATCCTGCAGACAGTGAAGGTG
GTGGACGAGCTCGTGAAAGTGATGGGCCGGCACAAGCCCGAGAACATCGTGATCGAA
ATGGCCAGAGAGAACCAGACCACCCAGAAGGGACAGAAGAACAGCCGCGAGAGAAT
GAAGCGGATCGAAGAGGGCATCAAAGAGCTGGGCAGCCAGATCCTGAAAGAACACCC
CGTGGAAAACACCCAGCTGCAGAACGAGAAGCTGTACCTGTACTACCTGCAGAATGGG
CGGGATATGTACGTGGACCAGGAACTGGACATCAACCGGCTGTCCGACTACGATGTGG
ACGCCATCGTGCCTCAGAGCTTTCTGAAGGACGACTCCATCGACAACAAGGTGCTGAC
CAGAAGCGACAAGAACCGGGGCAAGAGCGACAACGTGCCCTCCGAAGAGGTCGTGAA
GAAGATGAAGAACTACTGGCGGCAGCTGCTGAACGCCAAGCTGATTACCCAGAGAAA
GTTCGACAATCTGACCAAGGCCGAGAGAGGCGGCCTGAGCGAACTGGATAAGGCCGG

CTTCATCAAGAGACAGCTGGTGGAAACCCGGCAGATCACAAAGCACGTGGCACAGATC
CTGGACTCCCGGATGAACACTAAGTACGACGAGAATGACAAGCTGATCCGGGAAGTG
AAAGTGATCACCCTGAAGTCCAAGCTGGTGTCCGATTTCCGGAAGGATTTCCAGTTTTA
CAAAGTGCGCGAGATCAACAACTACCACCACGCCCACGACGCCTACCTGAACGCCGTC
GTGGGAACCGCCCTGATCAAAAAGTACCCTAAGCTGGAAAGCGAGTTCGTGTACGGCG
ACTACAAGGTGTACGACGTGCGGAAGATGATCGCCAAGAGCGAGCAGGAAATCGGCA
AGGCTACCGCCAAGTACTTCTTCTACAGCAACATCATGAACTTTTTCAAGACCGAGATT
ACCCTGGCCAACGGCGAGATCCGGAAGCGGCCTCTGATCGAGACAAACGGCGAAACC
GGGGAGATCGTGTGGGATAAGGGCCGGGATTTTGCCACCGTGCGGAAAGTGCTGAGCA
TGCCCCAAGTGAATATCGTGAAAAAGACCGAGGTGCAGACAGGCGGCTTCAGCAAAG
AGTCTATCCTGCCCAAGAGGAACAGCGATAAGCTGATCGCCAGAAAGAAGGACTGGG
ACCCTAAGAAGTACGGCGGCTTCGACAGCCCCACCGTGGCCTATTCTGTGCTGGTGGT
GGCCAAAGTGGAAAAGGGCAAGTCCAAGAAACTGAAGAGTGTGAAAGAGCTGCTGGG
GATCACCATCATGGAAAGAAGCAGCTTCGAGAAGAATCCCATCGACTTTCTGGAAGCC
AAGGGCTACAAAGAAGTGAAAAAGGACCTGATCATCAAGCTGCCTGAGTACTCCCTGT
TCGAGCTGGAAAACGGCCGGAAGAGAATGCTGGCCTCTGCCGGCGAACTGCAGAAGG
GAAACGAACTGGCCCTGCCCTCCAAATATGTGAACTTCCTGTACCTGGCCAGCCACTAT
GAGAAGCTGAAGGGCTCCCCCGAGGATAATGAGCAGAAACAGCTGTTTGTGGAACAG
CACAAGCACTACCTGGACGAGATCATCGAGCAGATCAGCGAGTTCTCCAAGAGAGTGA
TCCTGGCCGACGCTAATCTGGACAAAGTGCTGTCCGCCTACAACAAGCACCGGGATAA
GCCCATCAGAGAGCAGGCCGAGAATATCATCCACCTGTTTACCCTGACCAATCTGGGA
GCCCCTGCCGCCTTCAAGTACTTTGACACCACCATCGACCGGAAGAGGTACACCAGCA
CCAAAGAGGTGCTGGACGCCACCCTGATCCACCAGAGCATCACCGGCCTGTACGAGAC
ACGGATCGACCTGTCTCAGCTGGGAGGCGACAAAAGACCTGCCGCCACAAAGAAGGC
CGGACAGGCCAAAAAGAAAAAGGGTCTCAGCGGTGGTAGCAGCGGCGGGCCTACCAG
CGAAGTGGAATTCAGCCATGAATATTGGATGCGTCATGCACTGACCCTGGCAAAACGT
GCCCGTGATGAAGGTGAAGCACCGGTTGGTGCAGTTCTGGTTCTGAATAATCGTGTTAT
TGGTGAAGGTTGGAATCGTCGTATTGGTCTGCATGATCCGACCGCACATGCAGAAATT
ATGGCACTGCGTCAAGGTGGTCTGGTTATGCAGAATAGCCGTCTGATTGATGCAACCCT
GTATGTTACCTTTGAACCGTGTGTTATGTGTGCCGGTGCAATGATTAATAGCCGCATTG
GTCGTGTTGTTTTTGGTGTTCGTAATAGCAAACGTGGTGCAGCAGGTAGCCTGATGAAT
GTTCTGAATTATCCGGGTATGAATCACCGTGTTGAAATTACCGAAGGTATTCTGGCAGA
TGAATGTGCAGCACTGCTGTGTGATTTTTATCGTATGCCTCGTCAGGTTTTTAACGCAC
AGAAAAAGGCACAGAGCAGCATTAAT

Красный = сигнал ядерной локализации большого Т-антигена вируса SV40; Голубой = ген каталитически неактивной формы *Streptococcus pyogenes* Cas9; Оранжевый = сигнал ядерной локализации нуклеоплазмина; Фиолетовый = глицин-сериновый линкер; Зеленый = ген рекомбинантной дезаминазы TadA-8e (TadDE).

Приложение 10. Нуклеотидная последовательность, кодирующая полипептид-прекурсор Gag-MmCas12m-TadDE для продукции VLP-MmCas12m-TadDE.

ATGGGTGCCCGCGCCAGCGTGCTCAGCGGGGGAGAATTAGATCGATGGGAAAAAATTC
GGTTAAGGCCAGGGGGAAAGAAAAAATATAAATTAAAACATATAGTATGGGCAAGCA
GGGAGCTAGAACGATTCGCAGTTAATCCTGGCCTGTTAGAAACATCAGAAGGCTGTAG
ACAAATACTGGGACAGCTACAACCATCCCTTCAGACAGGATCAGAAGAACTTAGATCA
TTATATAATACAGTAGCAACCCTCTATTGTGTGCATCAAAGGATAGAGATAAAAGACA
CCAAGGAAGCTTTAGACAAGATAGAGGAAGAGCAAAACAAAAGTAAGAAAAAAGCA
CAGCAAGCAGCAGCTGACACAGGACACAGCAATCAGGTCAGCCAAAATTACCCTATA
GTGCAGAACATCCAGGGGCAAATGGTACATCAGGCCATATCACCTAGAACTTTAAATG
CATGGGTAAAAGTAGTAGAAGAGAAGGCTTTCAGCCCAGAAGTGATACCCATGTTTTC
AGCATTATCAGAAGGAGCCACCCCACAAGATTTAAACACCATGCTAAACACAGTGGGG
GGACATCAAGCAGCCATGCAAATGTTAAAAGAGACCATCAATGAGGAAGCTGCAGAA
TGGGATAGAGTGCATCCAGTGCATGCAGGGCCTATTGCACCAGGCCAGATGAGAGAAC
CAAGGGGAAGTGACATAGCAGGAACTACTAGTACCCTTCAGGAACAAATAGGATGGA
TGACACATAATCCACCTATCCCAGTAGGAGAAATCTATAAAAGATGGATAATCCTGGG
ATTAAATAAAATAGTAAGAATGTATAGCCCTACCAGCATTCTGGACATAAGACAAGGA
CCAAAGGAACCCTTTAGAGACTATGTAGACCGATTCTATAAAACTCTAAGAGCCGAGC
AAGCTTCACAAGAGGTAAAAAATTGGATGACAGAAACCTTGTTGGTCCAAAATGCGAA
CCCAGATTGTAAGACTATTTTAAAAGCATTGGGACCAGGAGCGACACTAGAAGAAATG
ATGACAGCATGTCAGGGAGTGGGGGGACCCGGCCATAAAGCAAGAGTTTTGGCTGAA
GCAATGAGCCAAGTAACAAATCCAGCTACCATAATGATACAGAAAGGCAATTTTAGGA
ACCAAAGAAAGACTGTTAAGTGTTTCAATTGTGGCAAAGAAGGGCACATAGCCAAAA
ATTGCAGGGCCCCTAGGAAAAAGGGCTGTTGGAAATGTGGAAAGGAAGGACACCAAA
TGAAAGATTGTACTGAGAGACAGGCTAATTTTTTAGGGAAGATCTGGCCTTCCCACAA
GGGAAGGCCAGGGAATTTTCTTCAGAGCAGACCAGAGCCAACAGCCCCACCAGAAGA
GAGCTTCAGGTTTGGGGAAGAGACAACAACTCCCTCTCAGAAGCAGGAGCCGATAGAC
AAGGAACTGTATCCTTTAGCTTCCCTCAGATCACTCTTTGGCAGCGACCCCTCGTCACA
ATAAAGATAGGGGGGCAATTAAAGGAAGCTCTATTAGATACAGGAGCAGATGATACA
GTATTAGAAGAAATGAATTTGCCAGGAAGATGGAAACCAAAAATGATAGGGGGAATT
GGAGGTTTTATCAAAGTAAGACAGTATGATCAGATACTCATAGAAATCTGCGGACATA
AAGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAATCT
GTTGACTCAGATTGGCTGCACTTTAAATTTTCCCATTAGTCCTGGGGGGGCCCAAGAAAA
AGCGGAAAGTGGGGCCCAGCGGCACCACCATGACAGTGCATACAATGGGCGTGCACT
ACAAGTGGCAGATCCCCGAGGTTCTGAGACAGCAGCTGTGGCTGGCCCACAACCTGAG
AGAAGATCTGGTGTCTCTGCAGCTGGCCTACGACGACGACCTGAAGGCCATCTGGTCT
AGCTACCCTGATGTGGCCCAGGCCGAGGATACAATGGCTGCCGCTGAAGCTGATGCCG
TGGCTCTGTCTGAGAGAGTGAAGCAGGCCAGAATCGAGGCCAGAAGCAAGAAGATCA
GCACCGAGCTGACCCAGCAGCTGAGGGACGCCAAGAAGAGACTGAAGGACGCCAGAC
AGGCCAGACGGGATGCCATTGCCGTGGTCAAAGATGATGCCGCCGAGCGGAGAAAGG
CCAGATCTGATCAACTGGCCGCCGACCAGAAAGCCCTGTACGGCCAGTACTGCAGAGA
TGGCGATCTGTACTGGGCCAGCTTCAACACCGTGCTGGACCACCACAAGACCGCCGTG
AAGAGAATCGCCGCTCAGAGAGCCTCTGGCAAGCCTGCCACACTGAGACACCACAGAT
TTGACGGCAGCGGCACAATCGCCGTGCAACTCCAAAGACAAGCTGGCGCCCCTCCTCG
GACACCTATGGTTCTTGCTGATGAGGCCGGCAAGTACCGGAACGTGCTGCACATTCCT
GGCTGGACCGATCCTGATGTGTGGGAGCAGATGACACGGTCCCAGTGCAGACAGTCTG
GCAGAGTGACAGTGCGGATGAGATGCGGCAGCACAGATGGACAGCCCCAGTGGATCG
ATCTGCCTGTGCAGGTTCACAGATGGCTGCCTGCCGACGCCGATATTACAGGCGCTGA
GCTGGTGGTCACAAGAGTGGCCGGAATCTACAGAGCCAAGCTGTGCGTGACCGCCAGG

ATCGGAGATACCGAGCCTGTGACATCTGGCCCTACAGTGGCCCTGCATCTTGGCTGGC
GATCTACCGAAGAGGGAACAGCCGTGGCCACTTGGAGATCTGATGCCCCTCTGGACAT
CCCCTTCGGCCTGAGAACAGTGATGAGAGTGGATGCTGCCGGCACCTCTGGCATCATT
GTGGTGCCTGCCACCATCGAGAGAAGGCTGACCCGGACAGAGAATATCGCCAGCAGC
AGATCTCTGGCCCTGGACGCCCTGAGAGATAAGGTTGTCGGCTGGCTGAGCGACAACG
ACGCCCCTACCTATAGAGATGCTCCACTGGAAGCCGCCACCGTGAAGCAGTGGAAGTC
CCCTCAGAGATTCGCCTCTCTGGCTCACGCCTGGAAGGACAATGGCACCGAGATCAGC
GATATCCTGTGGGCCTGGTTCAGCCTGGACAGAAAGCAGTGGGCTCAGCAAGAGAACG
GCAGAAGAAAGGCCCTGGGACACAGAGATGACCTGTATCGGCAGATCGCCGCCGTGA
TCTCTGATCAGGCTGGACACGTGCTGGTGGACGATACAAGTGTGGCCGAGCTGAGCGC
CAGAGCTATGGAAAGAACCGAGCTGCCTACCGAGGTGCAGCAGAAGATCGACAGAAG
AAGGGACCACGCTGCTCCCGGTGGACTGAGAGCTTCTGTGGTGGCCGCCATGACAAGA
GATGGGGTGCCAGTGACAATCGTGGCCGCAGCCGATTTCACCCGGACACACAGCAGAT
GCGGACACGTGAACCCCGCCGACGATAGATACCTGAGCAACCCCGTCAGATGCGACGG
CTGTGGCGCTATGTACGACCAGGACAGATCCTTCGTGACCCTGATGCTGAGAGCCGCC
ACAGCTCCTAGCAATCCTGGCAAAAGACCTGCCGCCACAAAGAAGGCCGGACAGGCC
AAAAAGAAAAAGGGTCTCAGCGGTGGTAGCAGCGGCGGGCCTACCAGCGAAGTGGAA
TTCAGCCATGAATATTGGATGCGTCATGCACTGACCCTGGCAAAACGTGCCCGTGATG
AAGGTGAAGCACCGGTTGGTGCAGTTCTGGTTCTGAATAATCGTGTTATTGGTGAAGGT
TGGAATCGTCGTATTGGTCTGCATGATCCGACCGCACATGCAGAAATTATGGCACTGC
GTCAAGGTGGTCTGGTTATGCAGAATAGCCGTCTGATTGATGCAACCCTGTATGTTACC
TTTGAACCGTGTGTTATGTGTGCCGGTGCAATGATTAATAGCCGCATTGGTCGTGTTGT
TTTTGGTGTTCGTAATAGCAAACGTGGTGCAGCAGGTAGCCTGATGAATGTTCTGAATT
ATCCGGGTATGAATCACCGTGTTGAAATTACCGAAGGTATTCTGGCAGATGAATGTGC
AGCACTGCTGTGTGATTTTTATCGTATGCCTCGTCAGGTTTTTAACGCACAGAAAAAGG
CACAGAGCAGCATTAATCATCACCATCACCATCAC

Черный = Белок Gag ВИЧ-1; Красный = сигнал ядерной локализации большого Т-антигена вируса SV40; Голубой = ген *Mycolicibacterium mucogenicum* Cas12m; Оранжевый = сигнал ядерной локализации нуклеоплазмина; Фиолетовый = глицин-сериновый линкер; Зеленый = ген рекомбинантной дезаминазы TadA-8e (TadDE); Красный = метка 6×His-tag.

Приложение 11. Нуклеотидная последовательность, кодирующая полипептид-прекурсор Gag-dSpCas9-TadDE для продукции VLP-dSpCas9-TadDE.

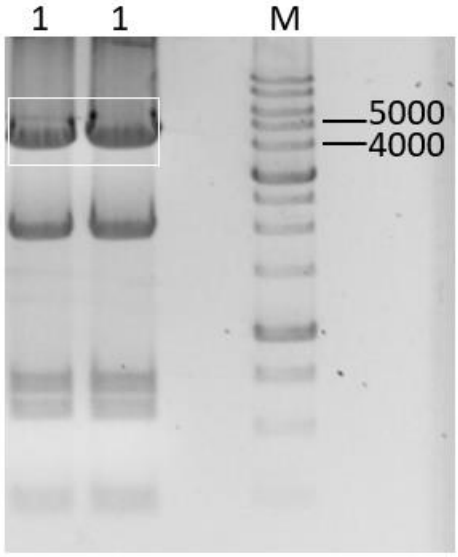
ATGGGTGCCCGCGCCAGCGTGCTCAGCGGGGGAGAATTAGATCGATGGGAAAAAATTC
GGTTAAGGCCAGGGGGAAAGAAAAAATATAAATTAAAACATATAGTATGGGCAAGCA
GGGAGCTAGAACGATTCGCAGTTAATCCTGGCCTGTTAGAAACATCAGAAGGCTGTAG
ACAAATACTGGGACAGCTACAACCATCCCTTCAGACAGGATCAGAAGAACTTAGATCA
TTATATAATACAGTAGCAACCCTCTATTGTGTGCATCAAAGGATAGAGATAAAAGACA
CCAAGGAAGCTTTAGACAAGATAGAGGAAGAGCAAAACAAAAGTAAGAAAAAAGCA
CAGCAAGCAGCAGCTGACACAGGACACAGCAATCAGGTCAGCCAAAATTACCCTATA
GTGCAGAACATCCAGGGGCAAATGGTACATCAGGCCATATCACCTAGAACTTTAAATG
CATGGGTAAAAGTAGTAGAAGAGAAGGCTTTCAGCCCAGAAGTGATACCCATGTTTTC
AGCATTATCAGAAGGAGCCACCCCACAAGATTTAAACACCATGCTAAACACAGTGGGG
GGACATCAAGCAGCCATGCAAATGTTAAAAGAGACCATCAATGAGGAAGCTGCAGAA
TGGGATAGAGTGCATCCAGTGCATGCAGGGCCTATTGCACCAGGCCAGATGAGAGAAC
CAAGGGGAAGTGACATAGCAGGAACTACTAGTACCCTTCAGGAACAAATAGGATGGA
TGACACATAATCCACCTATCCCAGTAGGAGAAATCTATAAAAGATGGATAATCCTGGG
ATTAAATAAAATAGTAAGAATGTATAGCCCTACCAGCATTCTGGACATAAGACAAGGA
CCAAAGGAACCCTTTAGAGACTATGTAGACCGATTCTATAAAACTCTAAGAGCCGAGC
AAGCTTCACAAGAGGTAAAAAATTGGATGACAGAAACCTTGTTGGTCCAAAATGCGAA
CCCAGATTGTAAGACTATTTTAAAAGCATTGGGACCAGGAGCGACACTAGAAGAAATG
ATGACAGCATGTCAGGGAGTGGGGGGACCCGGCCATAAAGCAAGAGTTTTGGCTGAA
GCAATGAGCCAAGTAACAAATCCAGCTACCATAATGATACAGAAAGGCAATTTTAGGA
ACCAAAGAAAGACTGTTAAGTGTTTCAATTGTGGCAAAGAAGGGCACATAGCCAAAA
ATTGCAGGGCCCCTAGGAAAAAGGGCTGTTGGAAATGTGGAAAGGAAGGACACCAAA
TGAAAGATTGTACTGAGAGACAGGCTAATTTTTTAGGGAAGATCTGGCCTTCCCACAA
GGGAAGGCCAGGGAATTTTCTTCAGAGCAGACCAGAGCCAACAGCCCCACCAGAAGA
GAGCTTCAGGTTTGGGGAAGAGACAACAACTCCCTCTCAGAAGCAGGAGCCGATAGAC
AAGGAACTGTATCCTTTAGCTTCCCTCAGATCACTCTTTGGCAGCGACCCCTCGTCACA
ATAAAGATAGGGGGGCAATTAAAGGAAGCTCTATTAGATACAGGAGCAGATGATACA
GTATTAGAAGAAATGAATTTGCCAGGAAGATGGAAACCAAAAATGATAGGGGGAATT
GGAGGTTTTATCAAAGTAAGACAGTATGATCAGATACTCATAGAAATCTGCGGACATA
AAGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAATCT
GTTGACTCAGATTGGCTGCACTTTAAATTTTCCCATTAGTCCTGGGGGGGCCCAAGAAAA
AGCGGAAAGTGGACAAGAAGTACAGCATCGGCCTGGCCATCGGCACCAACTCTGTGG
GCTGGGCCGTGATCACCGACGAGTACAAGGTGCCCAGCAAGAAATTCAAGGTGCTGGG
CAACACCGACCGGCACAGCATCAAGAAGAACCTGATCGGAGCCCTGCTGTTCGACAGC
GGCGAAACAGCCGAGGCCACCCGGCTGAAGAGAACCGCCAGAAGAAGATACACCAGA
CGGAAGAACCGGATCTGCTATCTGCAAGAGATCTTCAGCAACGAGATGGCCAAGGTGG
ACGACAGCTTCTTCCACAGACTGGAAGAGTCCTTCCTGGTGGAAGAGGATAAGAAGCA
CGAGCGGCACCCCATCTTCGGCAACATCGTGGACGAGGTGGCCTACCACGAGAAGTAC
CCCACCATCTACCACCTGAGAAAGAAACTGGTGGACAGCACCGACAAGGCCGACCTGC
GGCTGATCTATCTGGCCCTGGCCCACATGATCAAGTTCCGGGGCCACTTCCTGATCGAG
GGCGACCTGAACCCCGACAACAGCGACGTGGACAAGCTGTTCATCCAGCTGGTGCAGA
CCTACAACCAGCTGTTCGAGGAAAACCCCATCAACGCCAGCGGCGTGGACGCCAAGGC
CATCCTGTCTGCCAGACTGAGCAAGAGCAGACGGCTGGAAAATCTGATCGCCCAGCTG
CCCGGCGAGAAGAAGAATGGCCTGTTCGGAAACCTGATTGCCCTGAGCCTGGGCCTGA
CCCCCAACTTCAAGAGCAACTTCGACCTGGCCGAGGATGCCAAACTGCAGCTGAGCAA
GGACACCTACGACGACGACCTGGACAACCTGCTGGCCCAGATCGGCGACCAGTACGCC
GACCTGTTTCTGGCCGCCAAGAACCTGTCCGACGCCATCCTGCTGAGCGACATCCTGAG

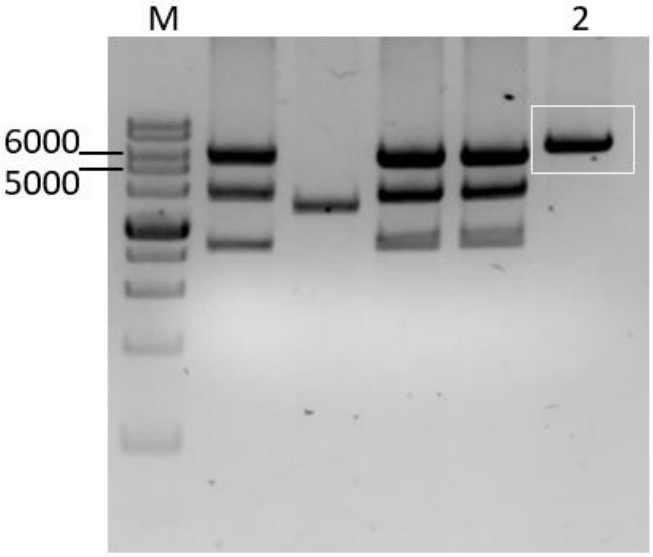
AGTGAACACCGAGATCACCAAGGCCCCCCTGAGCGCCTCTATGATCAAGAGATACGAC
GAGCACCACCAGGACCTGACCCTGCTGAAAGCTCTCGTGCGGCAGCAGCTGCCTGAGA
AGTACAAAGAGATTTTCTTCGACCAGAGCAAGAACGGCTACGCCGGCTACATTGACGG
CGGAGCCAGCCAGGAAGAGTTCTACAAGTTCATCAAGCCCATCCTGGAAAAGATGGAC
GGCACCGAGGAACTGCTCGTGAAGCTGAACAGAGAGGACCTGCTGCGGAAGCAGCGG
ACCTTCGACAACGGCAGCATCCCCCACCAGATCCACCTGGGAGAGCTGCACGCCATTC
TGCGGCGGCAGGAAGATTTTTACCCATTCCTGAAGGACAACCGGGAAAAGATCGAGAA
GATCCTGACCTTCCGCATCCCCTACTACGTGGGCCCTCTGGCCAGGGGAAACAGCAGA
TTCGCCTGGATGACCAGAAAGAGCGAGGAAACCATCACCCCCTGGAACTTCGAGGAAG
TGGTGGACAAGGGCGCTTCCGCCCAGAGCTTCATCGAGCGGATGACCAACTTCGATAA
GAACCTGCCCAACGAGAAGGTGCTGCCCAAGCACAGCCTGCTGTACGAGTACTTCACC
GTGTATAACGAGCTGACCAAAGTGAAATACGTGACCGAGGGAATGAGAAAGCCCGCC
TTCCTGAGCGGCGAGCAGAAAAAGGCCATCGTGGACCTGCTGTTCAAGACCAACCGGA
AAGTGACCGTGAAGCAGCTGAAAGAGGACTACTTCAAGAAAATCGAGTGCTTCGACTC
CGTGGAAATCTCCGGCGTGGAAGATCGGTTCAACGCCTCCCTGGGCACATACCACGAT
CTGCTGAAAATTATCAAGGACAAGGACTTCCTGGACAATGAGGAAAACGAGGACATTC
TGGAAGATATCGTGCTGACCCTGACACTGTTTGAGGACAGAGAGATGATCGAGGAACG
GCTGAAAACCTATGCCCACCTGTTCGACGACAAAGTGATGAAGCAGCTGAAGCGGCGG
AGATACACCGGCTGGGGCAGGCTGAGCCGGAAGCTGATCAACGGCATCCGGGACAAG
CAGTCCGGCAAGACAATCCTGGATTTCCTGAAGTCCGACGGCTTCGCCAACAGAAACT
TCATGCAGCTGATCCACGACGACAGCCTGACCTTTAAAGAGGACATCCAGAAAGCCCA
GGTGTCCGGCCAGGGCGATAGCCTGCACGAGCACATTGCCAATCTGGCCGGCAGCCCC
GCCATTAAGAAGGGCATCCTGCAGACAGTGAAGGTGGTGGACGAGCTCGTGAAAGTG
ATGGGCCGGCACAAGCCCGAGAACATCGTGATCGAAATGGCCAGAGAGAACCAGACC
ACCCAGAAGGGACAGAAGAACAGCCGCGAGAGAATGAAGCGGATCGAAGAGGGCAT
CAAAGAGCTGGGCAGCCAGATCCTGAAAGAACACCCCGTGGAAAACACCCAGCTGCA
GAACGAGAAGCTGTACCTGTACTACCTGCAGAATGGGCGGGATATGTACGTGGACCAG
GAACTGGACATCAACCGGCTGTCCGACTACGATGTGGACGCCATCGTGCCTCAGAGCT
TTCTGAAGGACGACTCCATCGACAACAAGGTGCTGACCAGAAGCGACAAGAACCGGG
GCAAGAGCGACAACGTGCCCTCCGAAGAGGTCGTGAAGAAGATGAAGAACTACTGGC
GGCAGCTGCTGAACGCCAAGCTGATTACCCAGAGAAAGTTCGACAATCTGACCAAGGC
CGAGAGAGGCGGCCTGAGCGAACTGGATAAGGCCGGCTTCATCAAGAGACAGCTGGT
GGAAACCCGGCAGATCACAAAGCACGTGGCACAGATCCTGGACTCCCGGATGAACACT
AAGTACGACGAGAATGACAAGCTGATCCGGGAAGTGAAAGTGATCACCCTGAAGTCC
AAGCTGGTGTCCGATTTCCGGAAGGATTTCCAGTTTTACAAAGTGCGCGAGATCAACA
ACTACCACCACGCCCACGACGCCTACCTGAACGCCGTCGTGGGAACCGCCCTGATCAA
AAAGTACCCTAAGCTGGAAAGCGAGTTCGTGTACGGCGACTACAAGGTGTACGACGTG
CGGAAGATGATCGCCAAGAGCGAGCAGGAAATCGGCAAGGCTACCGCCAAGTACTTC
TTCTACAGCAACATCATGAACTTTTTCAAGACCGAGATTACCCTGGCCAACGGCGAGA
TCCGGAAGCGGCCTCTGATCGAGACAAACGGCGAAACCGGGGAGATCGTGTGGGATA
AGGGCCGGGATTTTGCCACCGTGCGGAAAGTGCTGAGCATGCCCCAAGTGAATATCGT
GAAAAAGACCGAGGTGCAGACAGGCGGCTTCAGCAAAGAGTCTATCCTGCCCAAGAG
GAACAGCGATAAGCTGATCGCCAGAAAGAAGGACTGGGACCCTAAGAAGTACGGCGG
CTTCGACAGCCCCACCGTGGCCTATTCTGTGCTGGTGGTGGCCAAAGTGGAAAAGGGC
AAGTCCAAGAAACTGAAGAGTGTGAAAGAGCTGCTGGGGATCACCATCATGGAAAGA
AGCAGCTTCGAGAAGAATCCCATCGACTTTCTGGAAGCCAAGGGCTACAAAGAAGTGA
AAAAGGACCTGATCATCAAGCTGCCTGAGTACTCCCTGTTCGAGCTGGAAAACGGCCG
GAAGAGAATGCTGGCCTCTGCCGGCGAACTGCAGAAGGGAAACGAACTGGCCCTGCC
CTCCAAATATGTGAACTTCCTGTACCTGGCCAGCCACTATGAGAAGCTGAAGGGCTCC
CCCGAGGATAATGAGCAGAAACAGCTGTTTGTGGAACAGCACAAGCACTACCTGGACG
AGATCATCGAGCAGATCAGCGAGTTCTCCAAGAGAGTGATCCTGGCCGACGCTAATCT

GGACAAAGTGCTGTCCGCCTACAACAAGCACCGGGATAAGCCCATCAGAGAGCAGGC
CGAGAATATCATCCACCTGTTTACCCTGACCAATCTGGGAGCCCCTGCCGCCTTCAAGT
ACTTTGACACCACCATCGACCGGAAGAGGTACACCAGCACCAAAGAGGTGCTGGACGC
CACCCTGATCCACCAGAGCATCACCGGCCTGTACGAGACACGGATCGACCTGTCTCAG
CTGGGAGGCGACAAAAGACCTGCCGCCACAAAGAAGGCCGGACAGGCCAAAAAGAAA
AAGGGTCTCAGCGGTGGTAGCAGCGGCGGGCCTACCAGCGAAGTGGAATTCAGCCATG
AATATTGGATGCGTCATGCACTGACCCTGGCAAAACGTGCCCGTGATGAAGGTGAAGC
ACCGGTTGGTGCAGTTCTGGTTCTGAATAATCGTGTTATTGGTGAAGGTTGGAATCGTC
GTATTGGTCTGCATGATCCGACCGCACATGCAGAAATTATGGCACTGCGTCAAGGTGG
TCTGGTTATGCAGAATAGCCGTCTGATTGATGCAACCCTGTATGTTACCTTTGAACCGT
GTGTTATGTGTGCCGGTGCAATGATTAATAGCCGCATTGGTCGTGTTGTTTTTGGTGTTC
GTAATAGCAAACGTGGTGCAGCAGGTAGCCTGATGAATGTTCTGAATTATCCGGGTAT
GAATCACCGTGTTGAAATTACCGAAGGTATTCTGGCAGATGAATGTGCAGCACTGCTG
TGTGATTTTTATCGTATGCCTCGTCAGGTTTTTAACGCACAGAAAAAGGCACAGAGCAG
CATTAATCATCACCATCACCATCAC

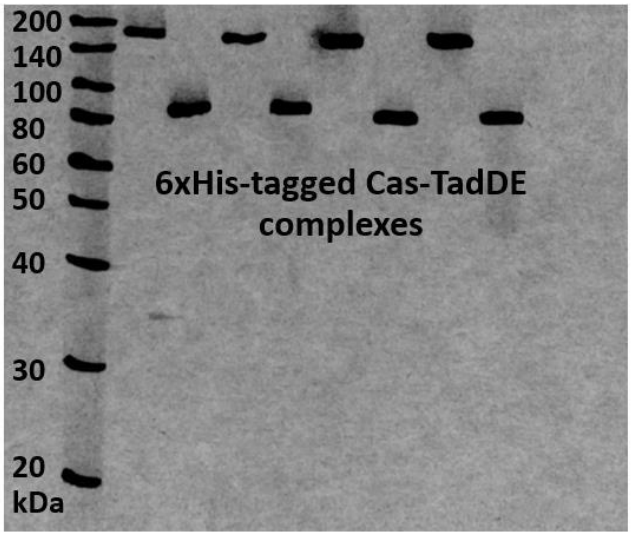
Черный = Белок Gag ВИЧ-1; Красный = сигнал ядерной локализации большого Т-антигена вируса SV40; Голубой = ген каталитически неактивной формы *Streptococcus pyogenes* Cas9; Оранжевый = сигнал ядерной локализации нуклеоплазмина; Фиолетовый = глицин-сериновый линкер; Зеленый = ген рекомбинантной дезаминазы TadA-8e (TadDE); Красный = метка 6×His-tag.

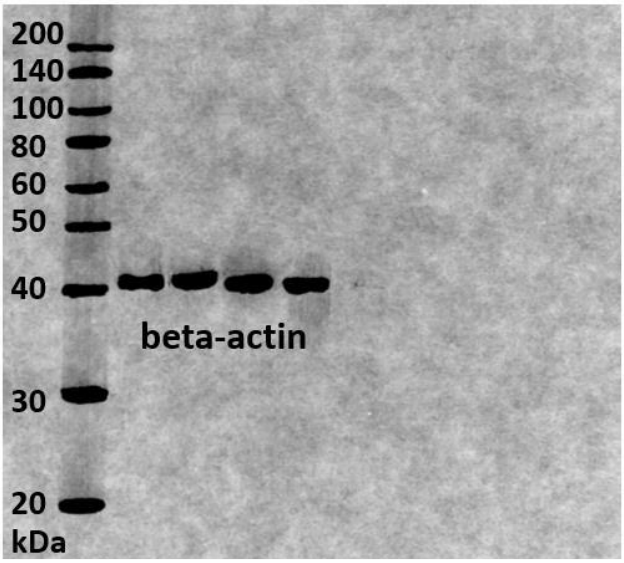
Приложение 12. Электрофореграмма продуктов амплификации, кодирующих полипептид-прекурсор Gag-MmCas12m-TadDE (1) и Gag-dSpCas9-TadDE (2); M – 1kb ДНК-маркер. Целевые ампликоны выделены рамкой.

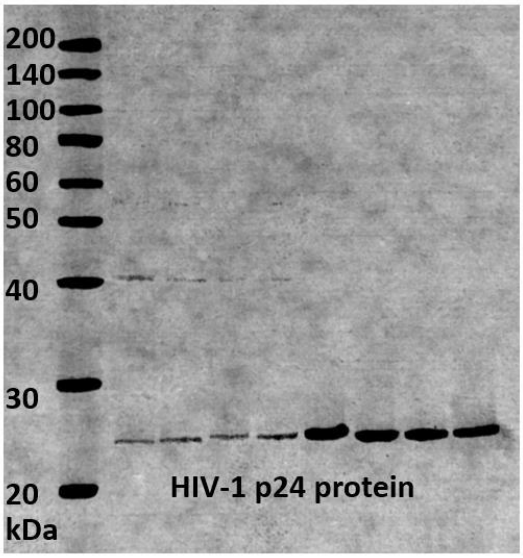




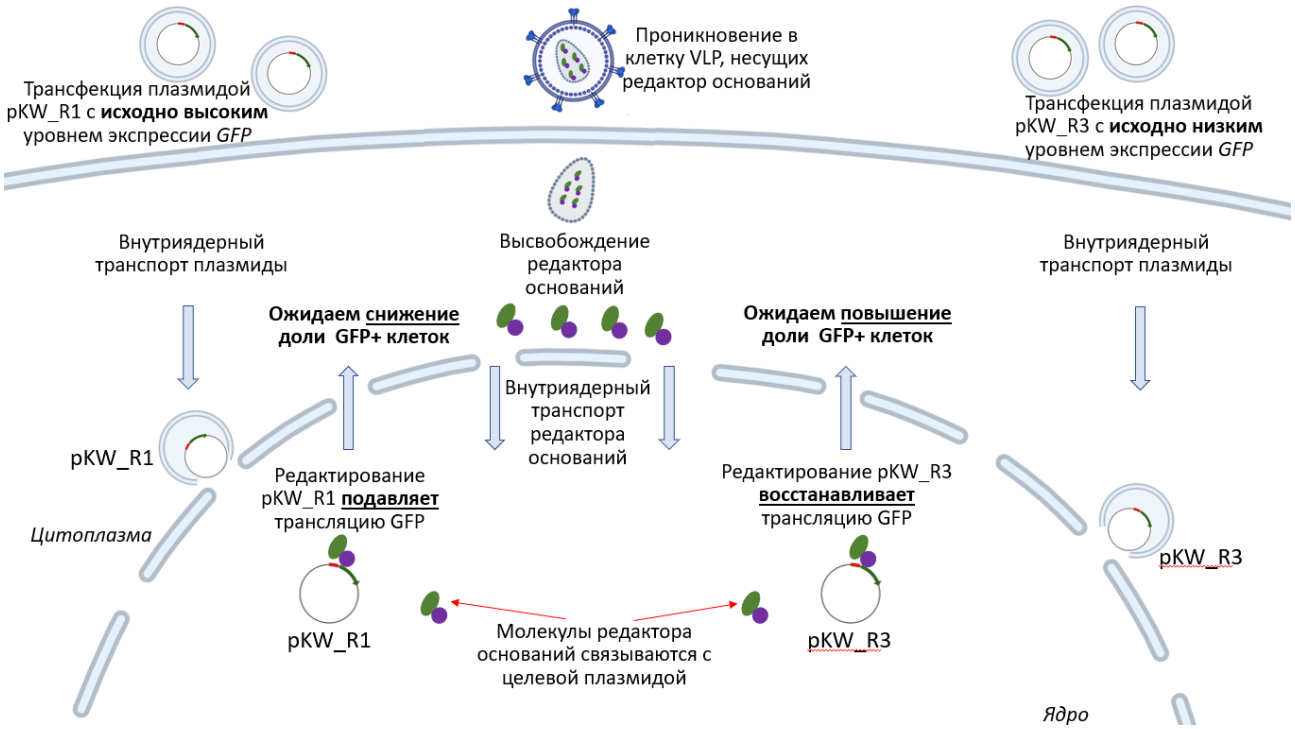
Приложение 13. Необработанные результаты вестерн-блота с наличием маркера молекулярного веса белков, 20-200 кДа.



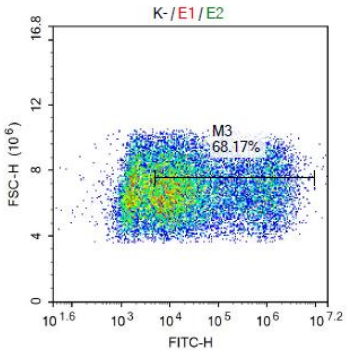




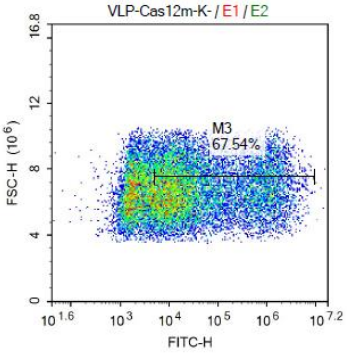
Приложение 14. Схема, отображающая принцип эксперимента для подтверждения функциональности VLP с использованием плазмид системы скрининга.



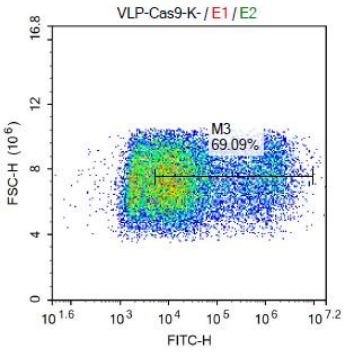
Приложение 15. Диаграммы рассеяния (дот-плоты), полученные при анализе методом проточной цитометрии клеток HEK293T, трансфицированных скрининговой плазмидой pKW_R1 и обработанных VLP. (K-) – только трансфицированные клетки, (VLP-Cas12-K-, VLP-Cas9-K-) – клетки, трансфицированные и обработанные VLP без sgRNA. (VLP-Cas12, VLP-Cas9) – клетки, трансфицированные и обработанные VLP, несущих редактор оснований в комплексе с sgRNA5.



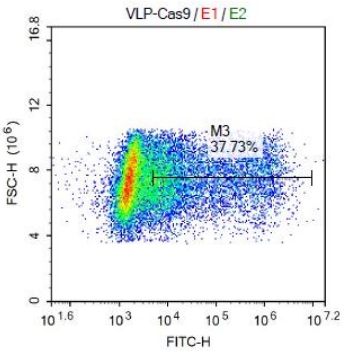
Gate	Count	% E2	Median X
E2	18,506	100.00%	11.916
M3	12,616	68.17%	30.418



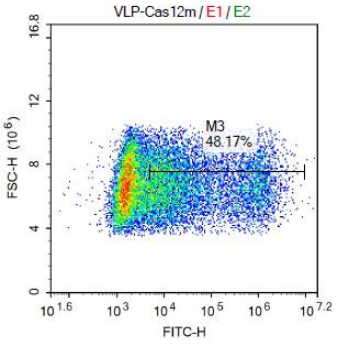
Gate	Count	% E2	Median X
E2	18,704	100.00%	11.739
M3	12,633	67.54%	29.784



Gate	Count	% E2	Median X
E2	19,493	100.00%	12.383
M3	13,468	69.09%	26.026

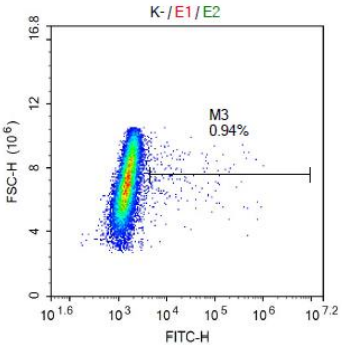


Gate	Count	% E2	Mean X
E2	19,767	100.00%	107.931
M3	7,458	37.73%	276.490

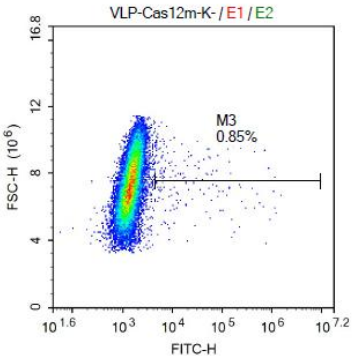


Gate	Count	% E2	Median X
E2	18,705	100.00%	4.558
M3	9,011	48.17%	29.410

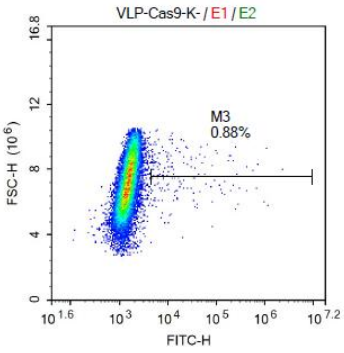
Приложение 16. Диаграммы рассеяния (дот-плоты), полученные при анализе методом проточной цитометрии клеток HEK293T, трансфицированных скрининговой плазмидой pKW_R3 и обработанных VLP. (K-) – только трансфицированные клетки, (VLP-Cas12-K-, VLP-Cas9-K-) – клетки, трансфицированные и обработанные VLP без sgRNA. (VLP-Cas12, VLP-Cas9) – клетки, трансфицированные и обработанные VLP, несущих редактор оснований в комплексе с sgRNA5.



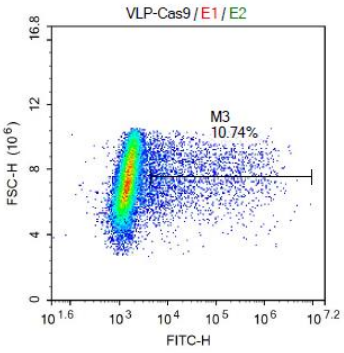
Gate	Count	% E2	Mean X
E2	20,000	100.00%	2,160
M3	187	0.94%	69,559



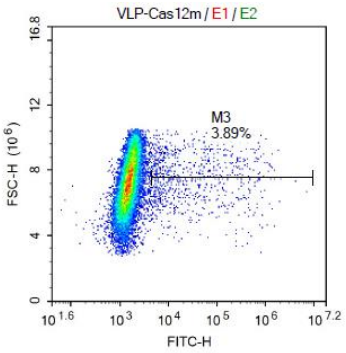
Gate	Count	% E2	Mean X
E2	20,385	100.00%	2,367
M3	173	0.85%	107,800



Gate	Count	% E2	Mean X
E2	20,000	100.00%	2,212
M3	175	0.88%	81,893



Gate	Count	% E2	Mean X
E2	20,000	100.00%	20,263
M3	2,148	10.74%	175,911



Gate	Count	% E2	Mean X
E2	20,000	100.00%	6,882
M3	778	3.89%	138,810

Приложение 17. Двумерные тепловые карты взаимодействия доменов редакторов оснований MmCas12m-TadDE (А) и dSpCas9-TadDE (Б) при редактировании старт-кодона *gag* провируса ВИЧ-1. Получено с помощью AlphaFold 3.

