БИОХИМИЯ, 2022, том 87, вып. 5, с. 617 – 626

УДК 577.21

РАЦИОНАЛЬНЫЙ ДИЗАЙН ОДНОЦЕПОЧЕЧНОГО ОЛИГОДЕЗОКСИРИБОНУКЛЕОТИДА, ИСПОЛЬЗУЕМОГО ДЛЯ КОРРЕКЦИИ МУТАЦИЙ ПРИ ГЕНОМНОМ РЕДАКТИРОВАНИИ

© 2022 О.В. Володина[1], А.А. Анучина[1], М.И. Зайнитдинова[1], Н.А. Евтушенко[2], А.В. Лавров[1], С.А. Смирнихина[1]*

[1] *Медико-генетический научный центр имени академика Н.П. Бочкова, 115522 Москва, Россия; электронная почта: smirnikhinas@gmail.com*

[2] *Центр высокоточного редактирования и генетических технологий для биомедицины Российского национального исследовательского медицинского университета имени Н.И. Пирогова, 117997 Москва, Россия*

Поступила в редакцию 03.12.2021
После доработки 19.04.2022
Принята к публикации 19.04.2022

Геномное редактирование позволяет целенаправленно вносить разнообразные изменения в геном, что потенциально может быть использовано для лечения наследственных заболеваний человека. Несмотря на многочисленные исследования в этой области, эффективность методов коррекции мутаций все еще остается невысокой, что не позволяет использовать данные методы в рутинной практике. Показано, что рациональный дизайн компонентов геномного редактирования может существенно повысить эффективность исправления мутаций. В данной работе мы предлагаем дизайн одноцепочечных олигодезоксирибонуклеотидов (оцОДН) для более эффективного редактирования генов. С использованием модельной системы с восстановлением нокаутированного *EGFP*, интегрированного в геном клеточной культуры HEK293T, показано, что лишь небольшая часть оцОДН (около 20 нуклеотидов – с 14-го нуклеотида с 5′-конца от двухцепочечного разрыва до 4-го нуклеотида в сторону 3′-конца), используемой в качестве донорной молекулы для репарации двухцепочечного разрыва ДНК, интегрирует в место разрыва. На основе полученных данных можно рационально подходить к дизайну оцОДН для исправления мутаций с помощью метода CRISPR-Cas9 для разработки генной терапии наследственных болезней человека.

КЛЮЧЕВЫЕ СЛОВА: CRISPR-Cas9, репарация с помощью однонитевой матрицы, одноцепочечные олигодезоксирибонуклеотиды, *EGFP*, проточная цитофлуориметрия.

DOI: 10.31857/S0320972522050037, **EDN:** ASTYLE

ВВЕДЕНИЕ

Одним из инструментов генной терапии является геномное редактирование, которое позволяет целенаправленно вносить изменения в геном: корректировать мутации, нокаутировать аллели, влиять на экспрессию генов, что потенциально может быть использовано для лечения наследственных заболеваний человека [1]. Эффективность коррекции мутаций при моногенных заболеваниях, несмотря на все прилагаемые усилия, зачастую невысокая, что препятствует быстрому внедрению новых методов лечения на основе геномного редактирования в клиническую практику [2, 3]. Рациональный дизайн компонентов геномного редактирования позволяет существенно повысить эффективность коррекции мутаций [4]. В данной работе мы предлагаем рациональный дизайн коротких одноцепочечных олигодезоксирибонуклеотидов для более эффективного редактирования генов.

При лечении наследственных заболеваний с помощью геномного редактирования можно корректировать мутации, то есть вносить целенаправленные изменения в последовательность ДНК. При этом в молекуле ДНК сначала создается двухцепочечный разрыв (ДЦР) с помощью современных программируемых нуклеаз [5] –

Принятые сокращения: ДЦР – двухцепочечный разрыв; НГСК – негомологичное соединение концов; НГР – направленная гомологичная репарация; оцОДН – одноцепочечный олигодезоксирибонуклеотид; РОНМ – репарация с помощью однонитевой матрицы; crРНК – спейсер гидовой РНК, Cas – CRISPR-ассоциированный белок; CRISPR – система, основанная на коротких палиндромных повторах, регулярно расположенных группами; PAM – последовательность прилежащего к протоспейсеру мотива (protospacer adjacent motif); TALENs – эффекторные нуклеазы, подобные активаторам транскрипции; ZNFs – нуклеазы цинковых пальцев.

* Адресат для корреспонденции.

нуклеаз цинковых пальцев (ZNFs) [6], эффекторных нуклеаз, подобных активаторам транскрипции (TALENs) [7], или системы, основанной на коротких палиндромных повторах, регулярно расположенных группами (CRISPR) и CRISPR-ассоциированных белках (Cas) [8] – а затем клетка с помощью собственных механизмов репарации исправляет ДЦР.

Есть несколько путей репарации ДНК в клетке после искусственного введения ДЦР. Их можно разделить на RAD51-зависимые и RAD51-независимые. К RAD51-зависимым путям репарации относятся генные конверсии – механизмы, основанные на рекомбинации с донорной молекулой. Они имеют низкую эффективность и зависят от фазы клеточного цикла, однако обладают высокой точностью. Использование донорной молекулы позволяет в конечном итоге изменить последовательность по желанию исследователя. Генные конверсии с использованием двухцепочечной матрицы проходят по каноничному пути, при котором RAD51 образует филаменты на 3′-концах нитей ДНК-мишени, чтобы обеспечить инвазию донорного дуплекса. При этом также необходимо наличие факторов, которые обеспечивают взаимодействие RAD51 с ДНК, включая BRCA2, PALB2 и DSS1. Эффективность генных конверсий с использованием двухцепочечной матрицы снижается при истощении или инактивации этих факторов [9].

К RAD51-независимым путям относятся негомологичное соединение концов (НГСК, NHEJ), отжиг одиночной цепи (SSA), репарация с помощью однонитевой матрицы (РОНМ, SSTR) и некоторые виды индуцируемой разрывом репликации (BIR), механизм которых еще изучается. НГСК сопряжено с резекцией и лигированием концов ДНК. Этот механизм репарации приводит к возникновению множества коротких инсерций и делеций, что ведет к сдвигу рамки считывания, в связи с этим этот путь репарации используют для нокаутирования генов. НГСК не зависит от наличия донорной молекулы и фазы клеточного цикла, поэтому является доминирующим в эукариотической клетке [10]. При SSA разрыв фланкируется гомологичными последовательностями ДНК, что приводит к делеции промежуточной последовательности.

Еще один независимый от RAD51 путь репарации – РОНМ; он активируется при наличии одноцепочечной донорной молекулы ДНК, в том числе в процессе редактирования генома с использованием одноцепочечных олигодезоксирибонуклеотидов (оцОДН, ssODN). В ходе РОНМ сразу после внедрения двухцепочечного разрыва укорачиваются 5′-концы, а одноцепочечные 3′-концы участвуют в отжиге с экзогенным донором [11]. MRX связывается с молекулой ДНК и инициирует укорочение концов, белок RAD52 позволяет экзогенному донору отжигаться с доступными концами геномной ДНК. RAD59 и Srs2 предотвращают взаимодействие концов ДНК с RAD51 и останавливают отжиг, инициированный RAD52. Копирование последовательности одноцепочечного донора осуществляется предположительно ДНК-полимеразой δ, затем происходит отжиг на противоположной стороне ДЦР. Несовпадения на 5′-конце создают облигатный гетеродуплекс, который затем исправляется системой репарации неспаренных оснований. Бреши, оставшиеся после РОНМ, заполняются полимеразой ζ [12].

Для оптимизации метода коррекции мутаций и повышения эффективности этого процесса необходимо понимание фундаментальных основ репарации, которые в перспективе приведут к использованию наилучшей по своему строению донорной молекулы. На данный момент при геномном редактировании используют оцОДН, длинные одноцепочечные ДНК (lssDNA) [13, 14] и двухцепочечные донорные молекулы – плазмиды и продукты амплификации [15, 16].

ОцОДН – небольшие одноцепочечные молекулы длиной до 200 нуклеотидов (нт), которые состоят из последовательности для вставки и двух плеч гомологии. Длина плеча гомологии может варьировать от 30 до 60 нт [17]. ОцОДН в основном используются для исправления точечных мутаций [18]. Репарация двухцепочечного разрыва с использованием матриц оцОДН может проходить по пути РОНМ. В этом случае репарация не зависит от RAD51, но зависит от RAD52, RAD59, Srs2 и комплекса Mre11-Rad50-Xrs2 (MRX) [19].

Проведенные ранее исследования [19, 20] показывают, что используемый для репарации двухцепочечных разрывов ДНК оцОДН не полностью встраивается в место разрыва. Концевые нуклеотиды оцОДН участвуют только в создании гомологии с ДНК, тогда как центральная часть – встраивается. Точные длины фрагментов оцОДН, участвующих в создании гомологии и встраивании, были неизвестны. В результате данного исследования было точно определено, какая часть оцОДН участвует в создании гомологии с целевой последовательностью ДНК, а какая – встраивается. Полученные результаты позволят более эффективно подбирать оцОДН для дальнейшего развития подходов к лечению наследственных моногенных заболеваний.

МАТЕРИАЛЫ И МЕТОДЫ

Плазмидные конструкции и олигонуклеотиды. В работе были использованы следующие плазмиды: pLVT_turboRFP635-eGFPmut, pCMV-dR8.91, pMD2.G и Cas9(1.1)-sgGFP#2. Плазмида pLVT_turboRFP635-eGFPmut была получена путем амплификации гена *EGFP* с мутацией c.337delG, приводящей к сдвигу рамки считывания с образованием стоп-кодона, с плазмиды pEGFP-mut и клонирования в плазмиду pLVT_turboRFP635. Последовательность гена *EGFP* доступна по ссылке: https://www.addgene.org/browse/sequence_vdb/2487/. Мутация c.337delG обозначает делецию нуклеотида G в 337-м положении от начала открытой рамки считывания. Плазмида pLVT_turboRFP635 была получена на основе плазмиды pRRL.SIN.EF1.WPRE (в лаборатории Didier Trono), в которую были внесены изменения в последовательность полилинкера с введением сайтов для эндонуклеаз рестрикции AgeI и SalI [21]. Карта плазмиды доступна в статье Серебровской с соавт. [21]. В дальнейшем в плазмиду были клонированы фрагменты TurboRFP-T2A (по сайтам рестрикции BamHI и NheI) и EGFPmut (по сайтам рестрикции Nhe и SalI). Таким образом, была получена плазмида pLVT_turboRFP635-eGFPmut, верификацию которой проводили секвенированием по Сэнгеру. Упаковочные плазмиды pCMV-dR8.91 и pMD2.G были любезно предоставлены Didier Trono (https://www.epfl.ch/labs/tronolab/). Карты плазмид доступны по ссылкам: https://www.addgene.org/vector-database/2221/ (pCMV-dR8.91) и https://www.addgene.org/12259/ (pMD2.G).

Подбор спейсера гидовой РНК (сгРНК) проводили с помощью программного обеспечения с открытым доступом Benchling (https://www.benchling.com/) на ген *EGFP* с мутацией c.337delG. Последовательность сгРНК (TGTCGCCCTCGAACTTCACT) была синтезирована в виде двух частично комплементарных олигонуклеотидов с липкими концами для клонирования в плазмиду: sgGFP_2.f (5′-caccTGTCGCCCTCGAACTTCACT-3′) и sgGFP_2.r (5′-aaacAGTGAAGTTCGAGGGCGACA-3′). Спейсер гидовой РНК клонировали в исходную плазмиду eSpCas9(1.1) (любезно предоставлена Feng Zhang; Addgene plasmid #71814; http://n2t.net/addgene:71814; RRID:Addgene_71814), которая содержит мутантную форму SpCas9, по сайтам рестрикции для BbsI. Верификацию полученной плазмиды Cas9(1.1)-sgGFP#2 проводили путем секвенирования по Сэнгеру и рестрикционным анализом.

С помощью той же программы Benchling были подобраны четыре варианта оцОДН для восстановления правильной последовательности *EGFP*, содержащих синонимичные замены в разных положениях. Спейсер гидовой РНК был подобран таким образом, что имел гомологию (отжигался) со смысловой цепью ДНК, оцОДН – с антисмысловой.

Клеточная культура. Клеточная культура HEK293T-GFPmut была получена в Центре высокоточного редактирования и генетических технологий для биомедицины Российского национального исследовательского медицинского университета имени Н.И. Пирогова путем лентивирусной трансдукции. Для получения лентивирусных частиц использовалась стандартная методика липофекции упаковочной линии клеток HEK293T. Клетки накануне высаживали в количестве, достигающем 70% монослоя (60 мм чашка). Для котрансфекции использовали смесь из трех плазмид pCMV-dR8.91 (2 мкг), pMD2.G (0,6 мкг) и pLVT_turboRFP635-eGFPmut (2 мкг), инкубировали в среде Opti-MEM («Thermo Fisher Scientific», США) с Lipofectamine 2000 («Thermo Fisher Scientific»). Через 6 ч после трансфекции указанных выше плазмид проводили замену среды на DMEM и накапливали вирусные частицы в течение 48 ч. Ростовая среда с вирусными частицами была пропущена через 0,45 мкм фильтр и использована для заражения клеток HEK293T. Для определения титра вируса следовали протоколу «Евроген» (Россия, https://evrogen.ru/kit-user-manuals/Lenti_LP002.pdf). Использовали 10 мкл вирусных частиц с титром $0,5 \times 10^5$ трансдуцирующих единиц/мл на 10^4 клеток. После заражения клетки подращивали в среде DMEM до состояния конфлуэнтности и сортировали их по наличию флуоресценции turboRFP635. С целью коррекции мутации c.337delG в гене *EGFP* проводили котрансфекцию 7,4 мкг плазмиды Cas9(1.1)-sgGFP#2 и 100 пмоль оцОДН на 200 тысяч клеток HEK293T-GFPmut с помощью Lipofectamine 2000 по протоколу производителя.

Сортировка клеток. Для получения клеточной культуры HEK293T-GFPmut после лентивирусной трансдукции при достижении клетками состояния конфлуэнтности проводили сортировку клеток по наличию флуоресценции turboRFP635 с помощью клеточного сортера S3 Cell Sorter («BioRad», США) с фильтром эмиссии 615 ± 25 нм. Через 72 ч после трансфекции Cas9(1.1)-sgGFP#2 и оцОДН клетки культуры HEK293T-GFPmut снимали с планшета. Количество GFP-положительных клеток (с успешным редактированием) оценивали с помощью того же клеточного сортера во время сортировки GFP-положительных клеток (длина волны 488 нм, фильтр эмиссии 525 ± 30 нм) для обога-

Таблица 1. Протокол амплификации

Первичная денатурация	95 °C	5 минут
Амплификация (37 циклов)	95 °C 61 °C 72 °C	30 секунд 30 секунд 50 секунд
Финальная элонгация	72 °C	5 минут

щения популяции клеток с успешным восстановлением правильной последовательности *EGFP*. Оценку проводили в двух технических повторностях.

Оценка эффективности редактирования. ДНК для амплификации и секвенирования фрагмента *EGFP* выделяли с помощью набора Quick-gDNA Miniprep Kit («Zymo Research», США). Амплификацию проводили по протоколу, представленному в табл. 1. Последовательность праймеров: eGFP65_F – 5′-ACGTAAACGGCCACAAGTTCA-3′, eGFP527_R – 5′-CTGCCGTCCTCGATGTTGT-3′.

После амплификации интересующего фрагмента *EGFP* образцы были секвенированы с обоих праймеров на приборе ABI Prism 3130XL Genetic Analyzer («Applied Biosystems», США) по протоколу производителя. Анализ хроматограмм проводили с помощью приложения Chromas и интернет-ресурса Benchling для определения нуклеотидов, встроившихся во время репарации ДНК. Процентное соотношение аллелей, восстановленных методом НГСК и РОНМ, оценивали с помощью программы TIDER (http://shinyapps.datacurators.nl/tider/). Оценку проводили в трех или четырех технических повторностях. Наличие аллелей дополнительно было валидировано с помощью программы DECODR (https://decodr.org).

Статистическая обработка. Для статистической обработки данных использовали критерий Манна–Уитни. Статистическую обработку проводили в программе STATISTICA, версия 10.0 («StatSoft», США). Различие считали значимым при $p < 0,05$.

РЕЗУЛЬТАТЫ

Эффективность восстановления EGFP в линии клеток. На подготовительном этапе работы была разработана модельная система для изучения параметров гомологичной репарации ДНК по матрице оцОДН. С помощью лентивирусной трансдукции была создана клеточная линия HEK293T-GFPmut с интегрированным геном *EGFPmut*. Ген зеленого флуоресцентного белка, *EGFPmut*, характеризуется однонуклеотидной делецией c.337delG, приводящей к сдвигу рамки считывания с появлением преждевременного стоп-кодона (рис. 1). При редактировании с помощью CRISPR-Cas9 и оцОДН происходит вставка нуклеотида G в положение c.337 и в клетках появляется флуоресценция eGFP.

Для оценки границ фрагментов оцОДН, участвующих в создании плеч гомологии и непосредственно репарации, были подобраны и синтезированы оцОДН с синонимичными точковыми заменами через разные промежутки (с заменами каждые 9 и 15 нт, а также с заменами на концах) и без замен (табл. 2). Все оцОДН содержали нуклеотид G, позволяющий восстанавливать корректную последовательность гена *EGFP*.

Таблица 2. Последовательности оцОДН, используемые в работе

Название	Последовательность, 5′ → 3′
ssODN_sgGFP_9nt	TTCTTCAA**A**GACGACGG**T**AACTACAA**A**ACCCGCGC**T**GAG**G**TGAAGTTCGAGGG**T**GACACCCT-**C**GTGAACCG**T**ATCGAGCT**C**AAGGGCATCGACTTCAA**A**GAGGACGG**T**AACATCCTGGGG
ssODN_sgGFP_15nt	TTCTTCAAGGACGA**T**GGCAACTACAAGAC**G**CGCGCCGAG**G**TGAA**A**TTCGAGGGCGA-CAC**G**CTGGTGAACCGCAT**A**GAGCTGAAGGGCAT**A**GACTTCAAGGAGGA**T**GGCAACATC-CTGGGG
ssODN_sgGFP_es	TTCTT**T**AAGGACGACGGCAACTACAAGACCCGCGCCGAG**G**TGAAGTTCGAGGGCGACACC-CTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACAT**T**CTGGGG
ssODN_sgGFP_wt	TTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAG**G**TGAAGTTCGAGGGCGACACC-CTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGG

Примечание. Нижним подчеркиванием полужирным выделен нуклеотид, восстанавливающий правильную последовательность гена *EGFP*. Полужирным черным отмечены синонимичные замены в последовательности оцОДН (ssODN).

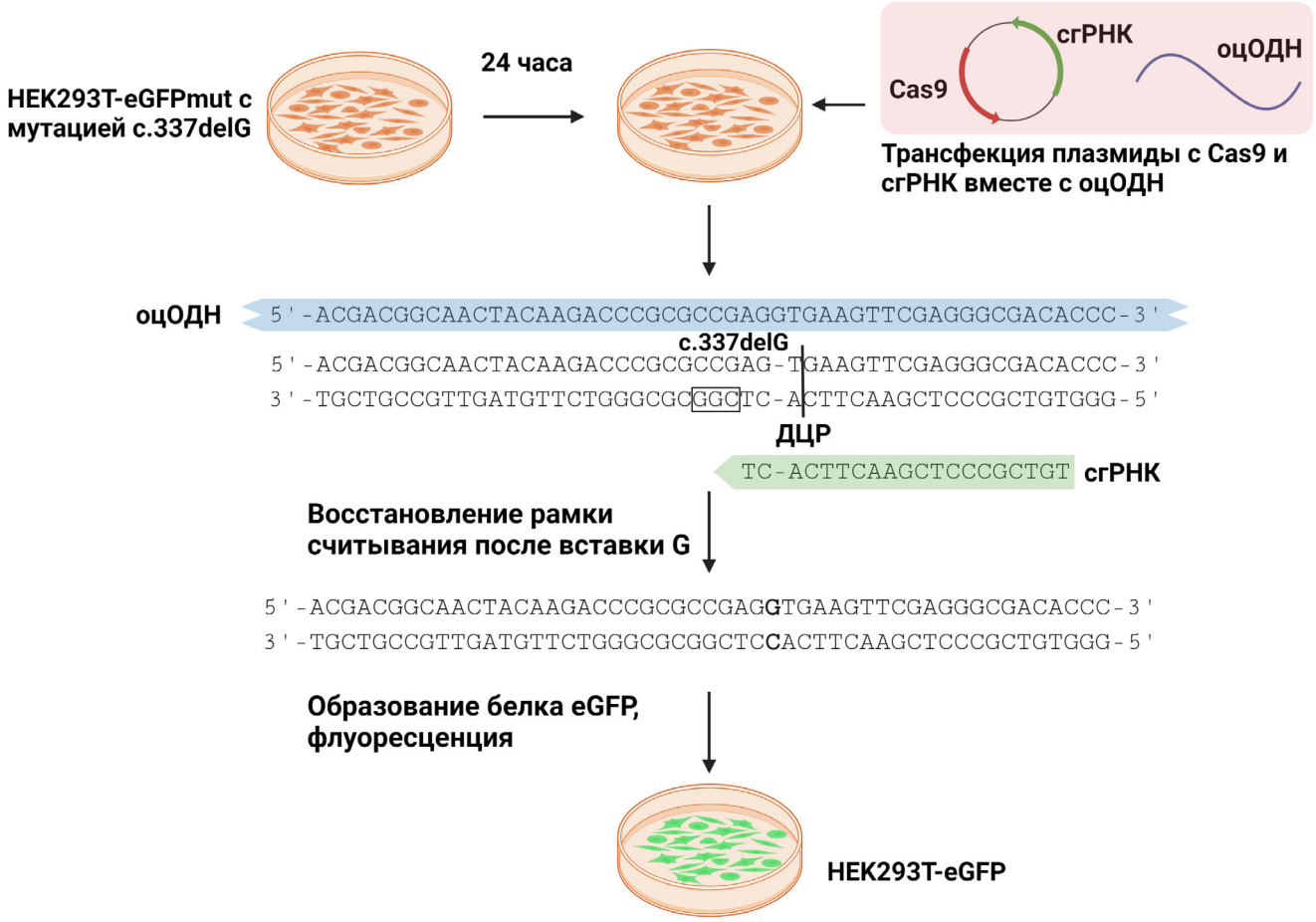

Рис. 1. Схема исправления мутации в гене *EGFP*. Мутация c.337delG приводит к сдвигу рамки считывания; после успешного редактирования наблюдается вставка нуклеотида G и восстановление рамки считывания, что приводит к синтезу нормального белка eGFP и появлению зеленой флуоресценции. Оценить результат редактирования можно с помощью проточной цитофлуориметрии и на хроматограммах. CrPHK – спейсер гидовой РНК, оцОДН – одноцепочечный олигодезоксирибонуклеотид, ДЦР – двухцепочечный разрыв

Далее клетки трансфицировали плазмидой Cas9(1.1)-sgGFP#2 с одним из подобранных оцОДН. Восстановление нормальной последовательности *EGFP* было зарегистрировано при использовании всех четырех оцОДН. Эффективность восстановления *EGFP* оценивали путем проточной цитофлуориметрии (оценивали долю GFP-положительных клеток). ОцОДН с синонимичными заменами каждые 9 или 15 нт обладают меньшей гомологией, поэтому ожидаемо продемонстрировали меньшую эффективность, чем оцОДН с двумя заменами на концах или вовсе без замен: 1,8–2,1% против ~6,7% (рис. 2, *а*).

GFP-положительные клетки были сортированы для обогащения клеточной популяции с успешно отредактированной мутацией. Эффективность редактирования целевого локуса оценили с помощью секвенирования по Сэнгеру с последующей обработкой хроматограмм по алгоритму TIDER. Частота инделов (коротких вставок и делеций нуклеотидов) в целевом локусе колебалась от 10,5% до 14,6% и составила в среднем 11,6%. Этот показатель отражает точность редактирования целевого локуса в общем и обычно обусловлен случайными потерями или вставками нуклеотидов в области ДЦР. Частота вставки нуклеотида G в положение c.337 (восстановление правильной последовательности гена *EGFP*) в GFP-положительных клетках варьировала от 55,6% до 65,7% в зависимости от использованного оцОДН (рис. 2, *б*). Таким образом, несмотря на различия в эффективности репарации разными оцОДН, были получены отредактированные клетки с восстановленным eGFP с применением всех четырех типов оцОДН, что позволило перейти к основному этапу оценки границ оцОДН, участвующих в формировании плеч гомологии, и участка, используемого в качестве матрицы для репарации.

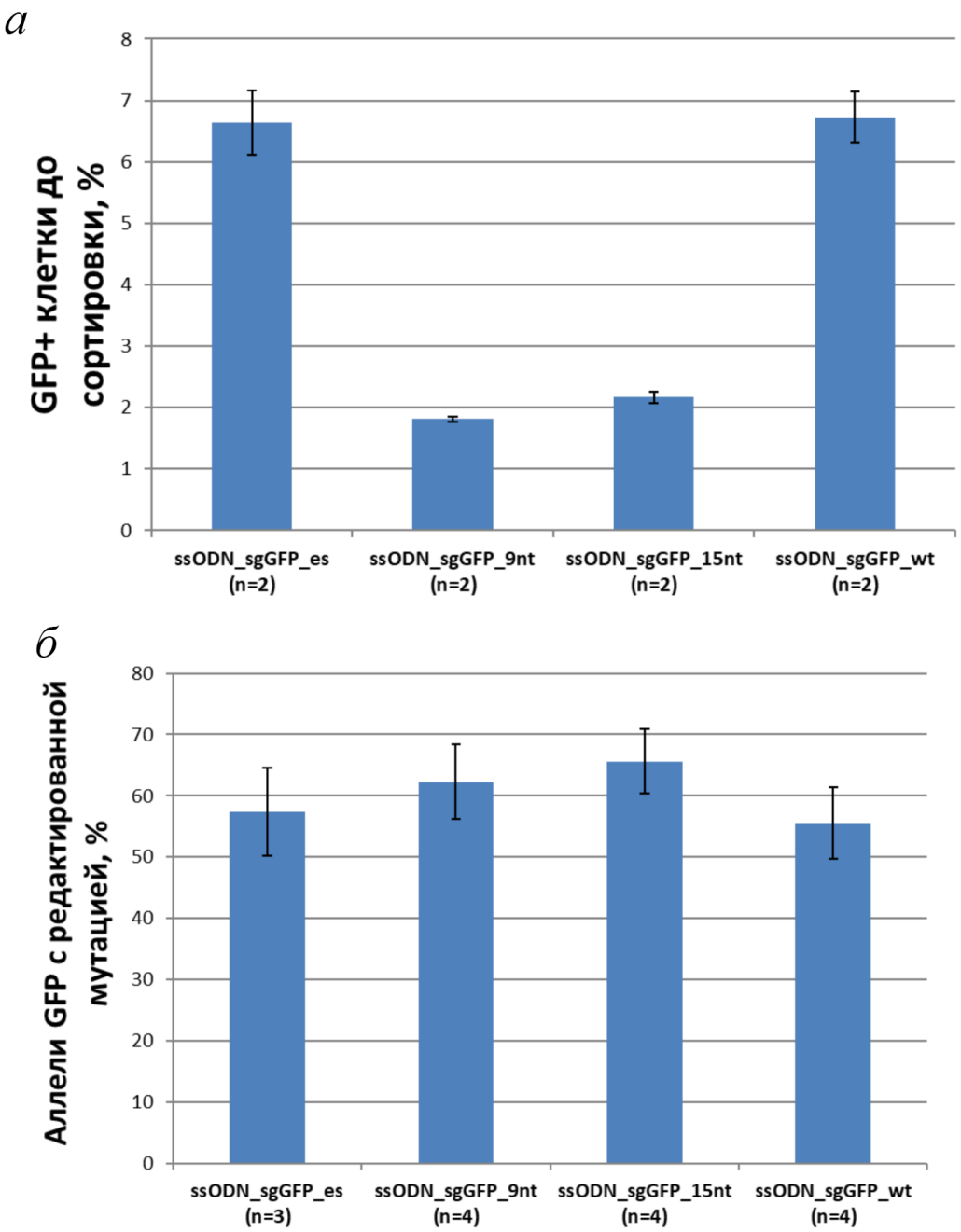

Рис. 2. Коррекция мутации c.337delG в гене *EGFP*. *а* – Процент GFP-положительных клеток в образцах, отредактированных с добавлением различных матриц оцОДН. *б* – Доля отредактированных (восстановленных) различными оцОДН аллелей в GFP-положительных клетках. ssODN_sgGFP_es – образцы с концевыми заменами, ssODN_sgGFP_9nt – образцы с заменами каждые 9 нт, ssODN_sgGFP_15nt – образцы с заменами каждые 15 нт, ssODN_sgGFP_wt – образцы без замен. Планки погрешностей по оси Y указывают 95%-ный доверительный интервал. Число n отражает количество технических повторностей

Длины фрагментов оцОДН, участвующих в создании плеч гомологии и встраивании. Целью работы была оценка границ частей оцОДН, участвующих в формировании плеч гомологии и используемых как матрица для репарации в экспериментах CRISPR-Cas9. Для анализа границ донорной последовательности, участвующей во вставке, анализировали только клетки, прошедшие сортировку по зеленой флуоресценции. В образцах с добавлением оцОДН с заменами на концах соответствующие замены в геноме определены не были, и хроматограммы не отлича-

лись от контрольных образцов, отредактированных оцОДН_sgGFP_wt без замен (данные не представлены).

При редактировании с помощью оцОДН с заменами каждые 9 нт обнаружены замены нуклеотидов в 5-м и 14-м положениях от ДЦР в сторону 5′-конца оцОДН-матрицы (рис. 3, *а*). Подчеркнём, что оцОДН были подобраны комплементарными к сгРНК; здесь и далее направления указаны относительно матрицы оцОДН, если не указано иначе. После редактирования клеток с использованием оцОДН с заменами каждые 15 нт замена нуклеотидов произошла в 4-м положении в сторону 3′-конца (рис. 3, *б*). При этом на хроматограммах наблюдали наложение пиков нуклеотидов, начиная со встроенного нуклеотида G, что свидетельствует о том, что в образце присутствовало минимум два вида аллелей – редактированный (содержащий дополнительный G) и исходный (с мутацией). Полные хроматограммы, фрагменты которых изображены на рис. 3, представлены в Приложении (рис. S1 и S2). Также в Приложении приведены таблицы с частотами аллелей *EGFP* после редактирования разными оцОДН (рис. S3).

Таким образом, результаты проведенной работы свидетельствуют о том, что со стороны 5′-конца от ДЦР были заменены –14 и –5 нуклеотиды, а со стороны 3′-конца от ДЦР – нуклеотид в позиции +4 (рис. 4).

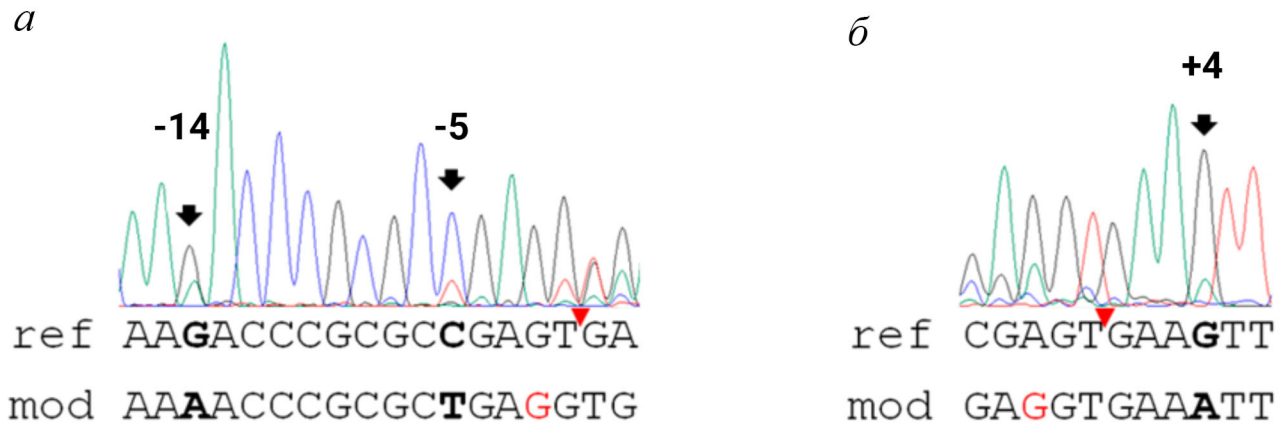

Рис. 3. Замены нуклеотидов в редактированном фрагменте *EGFP*. *а* – Хроматограмма образца после редактирования с помощью ssODN_sgGFP_9nt. Стрелками указаны замены нуклеотидов G на A и C на T. Красной стрелкой указано место введения ДЦР. *б* – Хроматограмма образца после редактирования с помощью ssODN_sgGFP_15nt. Стрелкой указана замена нуклеотида G на A. Красной стрелкой указано место введения ДЦР. Красным выделена вставка нуклеотида G в позиции c.337. Наложение последовательностей, прочитанных с прямого праймера вправо от вставки G (панель *а*), и последовательностей, прочитанных с обратного праймера влево от вставки G (панель *б*), обусловлено тем, что не все аллели *EGFP* в клеточной культуре подвергаются редактированию, что, однако, не затрудняет чтение синонимичных замен до вставки с соответствующего праймера. Ref – исходная последовательность, mod – измененная в результате геномного редактирования последовательность

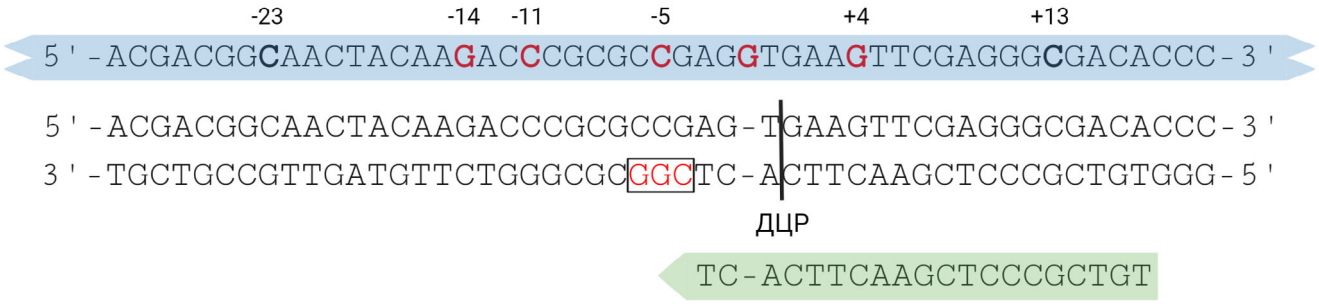

Рис. 4. Расположение заменяемых нуклеотидов при редактировании мутации c.337delG в гене *EGFP* с помощью оцОДН. Голубым указан фрагмент оцОДН, зеленым – последовательность сгРНК, рамкой выделен прилежащий к протоспейсеру мотив (PAM). Красным полужирным в последовательности оцОДН выделены нуклеотиды, встроенные в последовательность ДНК, полужирным черным – не встроенные. Цифрами указаны анализируемые нуклеотиды от ДЦР

ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

ОцОДН широко используют в качестве матриц для репарации при CRISPR-Cas9-опосредованном редактировании. Однако до сих пор нет единого мнения насчёт оптимальной длины олигонуклеотидов, их расположения относительно места разрыва и величины плеч гомологии и, соответственно, размеров внутренней части, служащей непосредственно матрицей для восстановления. В данном исследовании мы применили оцОДН длиной 120 нт с асимметричным расположением вокруг места разреза ДНК – 5′-конец длиной в 41 нт и 3′-конец – 79 нт. Для определения размеров и границ области, являющейся матрицей для гомологичной репарации, были использованы оцОДН с синонимичными заменами каждые 9 и каждые 15 нт. Такие оцОДН позволили определить, что для репарации используется небольшой фрагмент длиной около 20 нт (от положения –15 с 5′-конца от ДЦР до положения +4 в сторону 3′-конца от ДЦР). Таким образом, с помощью оцОДН можно скорректировать последовательность в 20 нт, асимметрично расположенных вокруг точки разрыва ДНК. Данное утверждение справедливо для оцОДН, подобранных комплементарными сгРНК, т.е. являющимися идентичными той цепи ДНК, на которую отжигается сгРНК (рис. 4), что важно учитывать при дизайне соответствующих экспериментов.

Попытки изучить детально, какая часть оцОДН интегрируется в место двухцепочечного разрыва предпринимались ранее [19, 20]. В работе Gallagher et al. [17] анализировали только область из восьми нуклеотидов вокруг разрыва (по четыре с каждой стороны), а в работе Harmsen et al. [19] изучали влияние факторов репарации неспаренных оснований на эффективность редактирования с помощью оцОДН. Таким образом, насколько нам известно, наша работа является единственной направленной именно на определение длин фрагментов оцОДН, интегрируемых в локусы ДЦР.

В целом, полученные нами данные о длинах фрагментов оцОДН, участвующих в создании гомологии и встраивании, хорошо коррелируют с процитированными выше работами [19, 20]. Со стороны 5′-плеча оцОДН встроилось больше нуклеотидов, чем со стороны 3′-плеча. По данным исследования системы репарации неспаренных оснований эффективность вставки на расстоянии более 9 нт от разрыва сильно зависит от того, на какой стороне была замена. Замена со стороны 5′-плеча происходила в 6–10% случаев, со стороны 3′-плеча – не происходила совсем [20]. Предположительно, причина в том, что 3′-плечо может отжигаться с выступающим концом на месте разрыва геномной ДНК, а 5′-плечо не способно на отжиг с выступающим концом. Также возможно влияние активности ДНК-полимеразы Polδ(3′-5′), которая может устранить несовпадения нуклеотидов близко к 3′-концу матрицы [19]. В нашей работе мы также можем отметить, что процент успешно редактированных клеток имел тенденцию к увеличению в тех образцах, где были использованы оцОДН с меньшим количеством замен (рис. 2, *а*). Вероятнее всего, это обусловлено тем, что плечи гомологии не являются строго комплементарными геномной ДНК, из-за чего эффективность геномного редактирования после инициации ДЦР могла заметно снизиться.

Перекрывающиеся пики на хроматограммах, полученных в данной работе, обусловлены секвенированием смеси аллелей, часть из которых редактировалась, а часть – нет. Современные методы анализа хроматограмм, один из которых – TIDER [22] – использованный нами, позволяют разложить сложную хроматограмму на составляющие. Этот метод широко используется в подобных работах [23, 24], т.к. он позволяет получать точные и воспроизводимые данные по результатам анализа сложных хроматограмм.

Таким образом, в результате работы были четко установлены длины фрагментов оцОДН, участвующие в создании гомологии и интеграции в локус ДЦР. На основе этих данных можно рационально подходить к дизайну оцОДН для коррекции мутаций с целью разработки генной терапии наследственных болезней человека на основе метода CRISPR-Cas9.

Финансирование. Работа выполнена в рамках государственного задания Министерства науки и высшего образования Российской Федерации для ФГБНУ «МГНЦ». Создание клеточной культуры HEK293T-GFPmut выполнено за счет средств гранта РФФИ № 19-29-04044.

Благодарности. Коллектив выражает благодарность к.б.н. Н.Г. Гурской из Центра высокоточного редактирования и генетических технологий для биомедицины Российского национального исследовательского медицинского университета имени Н.И. Пирогова (Москва) за помощь в получении клеточной культуры HEK293T-GFPmut.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

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

Дополнительные материалы. Приложение к статье опубликовано на сайте журнала «Биохимия» (https://biochemistrymoscow.com).

СПИСОК ЛИТЕРАТУРЫ

1. Fellmann, C., Gowen, B. G., and Lin, P. C. (2017) Cornerstones of CRISPR-Cas in drug discovery and therapy, *Nat Rev Drug Discov.*, **16**, 89-100, doi: 10.1038/nrd. 2016.238.
2. Peters-Hall, J. R., Coquelin, M. L., Torres, M. J., LaRanger, R., Alabi, B. R., et al. (2018) Long-term culture and cloning of primary human bronchial basal cells that maintain multipotent differentiation capacity and CFTR channel function, *Am. J. Physiol. Lung Cell Mol. Physiol.*, **315**, L313-L327, doi: 10.1152/ajplung.00355.2017.
3. Smirnikhina, S. A., Kondrateva, E. V., Adilgereeva, E. P., Anuchina, A. A., Zaynitdinova, M. I., et al. (2020) P.F508del editing in cells from cystic fibrosis patients, *PLoS One*, **15**, e0242094, doi: 10.1371/journal.pone.0242094.
4. Richardson, C. D., Ray, G. J., DeWitt, M. A., Curie, G. L., and Corn, J. E. (2016) Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA, *Nat. Biotechnol.*, **34**, 339-344, doi: 10.1038/nbt.3481.
5. Guha, T. K., Wai, A., and Hausner, G. (2017) Programmable genome editing tools and their regulation for efficient genome engineering, *Comput. Struct. Biotechnol. J.*, **15**, 146-160, doi: 10.1016/j.csbj.2016.12.006.
6. Durai, S., Mani, M., Kandavelou, K., Wu, J., Porteus, W. H., and Chandrasegaran, S. (2005) Zinc finger nucleases: Custom-designed molecular scissors for genome engineering of plant and mammalian cells, *Nucleic Acids Res.*, **33**, 5978-5990, doi: 10.1093/nar/gki912.
7. Cermak, T., Doyle, E. L., Christian, M., Wang, Li, Zhang, Y., et al. (2011) Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting, *Nucleic Acids Res.*, **39**, e82, doi: 10.1093/nar/gkr218.
8. Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., et al. (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity, *Science*, **337**, 816-821, doi: 10.1126/science.1225829.
9. Liu, M., Rehman, S., Tang, X., Gu, K., Fan, Q., et al. (2019) Methodologies for improving HDR Efficiency, *Front. Genet.*, **9**, doi: 10.3389/fgene.2018.00691.
10. Seol, J. H., Shim, E. Y., and Lee, S. E. (2018) Microhomology-mediated end joining: Good, bad and ugly, *Mutat. Res.*, **809**, 81-87, doi: 10.1016/j.mrfmmm.2017.07.002.
11. Davis, L., Zhang, Y., Maizels, N. (2018) Assaying repair at DNA nicks, *Methods Enzymol.*, **601**, 71-89, doi: 10.1016/bs.mie.2017.12.001.
12. Gallagher, D. N., and Haber, J. E. (2021) Single-strand template repair: Key insights to increase the efficiency of gene editing, *Curr. Genet.*, **67**, 747-753, doi: 10.1007/s00294-021-01186-z.
13. Hu, Z., Zhou, M., Wu, Y., Li, Z., Liu, X., et al. (2019) ssODN-mediated in-frame deletion with CRISPR/Cas9 restores FVIII function in hemophilia A-patient-derived iPSCs and ECs, *Mol. Ther. Nucleic Acids*, **17**, 198-209, doi: 10.1016/j.omtn.2019.05.019.
14. Bennett, H., Aguilar-Martinez, E., and Adamson, A. D. (2021) CRISPR-mediated knock-in in the mouse embryo using long single stranded DNA donors synthesised by biotinylated PCR, *Methods*, **191**, 3-14, doi: 10.1016/j.ymeth.2020.10.012.
15. Cristea, S., Freyvert, Y., and Santiago, Y. (2013) *In vivo* cleavage of transgene donors promotes nuclease-mediated targeted integration, *Biotechnol. Bioeng.*, **110**, 871-880, doi: 10.1002/bit.24733.
16. Fueller, J., Herbst, K., Meurer, M., Gubicza, K., Kurtulmus, B., et al. (2020) CRISPR-Cas12a-assisted PCR tagging of mammalian genes, *J. Cell Biol.*, **219**, e201910210, doi: 10.1083/jcb.201910210.
17. Bai, H., Liu, L., An, K., Lu, X., Harrison, M., et al. (2020) CRISPR/Cas9-mediated precise genome modification by a long ssDNA template in zebrafish, *BMC Genomics*, **21**, 67, doi: 10.1186/s12864-020-6493-4.
18. Lim, D., Sreekanth, V., Cox, K. J., Law, B. K., Wagner, B. K., et al. (2020) Engineering designer beta cells with a CRISPR-Cas9 conjugation platform, *Nat. Commun.*, **11**, 4043, doi: 10.1038/s41467-020-17725-0.
19. Harmsen, T., Klaasen, S., van de Vrugt, H., Riele, H. T. (2018) DNA mismatch repair and oligonucleotide end-protection promote base-pair substitution distal from a CRISPR/Cas9-induced DNA break, *Nucleic Acids Res.*, **46**, 2945-2955, doi: 10.1093/nar/gky076.
20. Galagher, D. N., Pham, N., Tsai, A. M., Janto, N. V., Choi, J., et al. (2020) A Rad51-independent pathway promotes single-strand template repair in gene editing, *PLoS Genet.*, **16**, e1008689, doi: 10.1371/journal.pgen.1008689.
21. Serebrovskaya, E. O., Podvalnaya, N. M., Dudenkova, V. V., Efremova, A. S., Gurskaya, N. G., et al. (2020) Genetically encoded fluorescent sensor for poly-ADP-ribose, *Int. J. Mol. Sci.*, **21**, 5004, doi: 10.3390/ijms21145004.
22. Brinkman, E. K., Kousholt, A. N., Harmsen, T., Leemans, C., Chen, T., et al. (2018) Easy quantification of template-directed CRISPR/Cas9 editing, *Nucleic Acids Res.*, **46**, e58, doi: 10.1093/nar/gky164.
23. Etard, C., Joshi, S., and Stegmaier, J., (2017) Tracking of indels by DEcomposition is a Simple and effective method to assess efficiency of guide RNAs in zebrafish, *Zebrafish*, **14**, 586-588, doi: 10.1089/zeb.2017.1454.
24. Sentmanat, M. F., Peters, S. T., Florian, C. P., Connelly, J. P., and Pruett-Miller, S. M. (2018) A survey of validation strategies for CRISPR-Cas9 editing, *Sci. Rep.*, **8**, 888, doi: 10.1038/s41598-018-19441-8.

RATIONAL DESIGN OF SSODN TO CORRECT MUTATIONS BY GENE EDITING

O. V. Volodina[1], A. A. Anuchina[1], M. I. Zainitdinova[1], N. A. Evtushenko[2], A. V. Lavrov[1], and S. A. Smirnikhina[1]*

[1] *Research Centre for Medical Genetics, 115522 Moscow, Russia; e-mail: smirnikhinas@gmail.com*

[2] *Center for Precision Genome Editing and Genetic Technologies for Biomedicine, Pirogov Russian National Research Medical University, 117997 Moscow, Russia*

Gene editing allows to make a variety of targeted changes in genome, which can potentially be used to treat hereditary human diseases. Despite numerous studies in this area, effectiveness of gene editing methods for correcting mutations is still low, so these methods are not allowed in routine practice. It has been shown that rational design of genome editing components can significantly increase efficiency of mutation correction. In this work, we propose design of single-stranded oligodeoxyribonucleotides (ssODNs) to increase efficiency of gene editing. Using a model system to repair knocked out *EGFP* that is integrated into genome of HEK293T cell culture, we have shown that only a small part of ssODN (about 20 nucleotides: from the 15th nucleotide at 3′-end to the 4th nucleotide at 5′-end), a donor molecule for repairing double-stranded DNA breaks, integrates into the site of the break. Based on obtained data, it is possible to rationally approach design of ssODNs to correct mutations using CRISPR-Cas9 method for development of gene therapy for hereditary human diseases.

Keywords: CRISPR-Cas9, single-strand template repair, single-stranded oligodeoxyribonucleotides, *EGFP*, flow cytometry