



Оригинальная статья

# Эффективность программируемых нуклеаз SpCas9 и AsCpf1 (Cas12a) в геномных локусах safe harbor клеток линии HEK293

Павлова С.В.<sup>1</sup> • Елисафенко Е.А.<sup>1</sup> • Шаяхметова Л.Ш.<sup>1,2</sup> • Медведев С.П.<sup>1</sup>

**Павлова Софья Викторовна** – канд. биол. наук, науч. сотр. лаборатории эпигенетики развития<sup>1</sup>; ORCID: <https://orcid.org/0000-0003-1095-3692>

✉ 630090, г. Новосибирск, ул. Мальцева, 1–16, Российская Федерация. Тел.: +7 (913) 903 13 62. E-mail: [sonpavlova@gmail.com](mailto:sonpavlova@gmail.com)

**Елисафенко Евгений Анатольевич** – канд. биол. наук, ст. науч. сотр. лаборатории эпигенетики развития<sup>1</sup>; ORCID: <https://orcid.org/0000-0002-3204-8178>. E-mail: [antares@bionet.nsc.ru](mailto:antares@bionet.nsc.ru)

**Шаяхметова Лилия Шагитовна** – лаборант научно-образовательного отдела<sup>1</sup>; бакалавр<sup>2</sup>

**Медведев Сергей Петрович** – канд. биол. наук, вед. науч. сотр. лаборатории эпигенетики развития<sup>1</sup>; ORCID: <https://orcid.org/0000-0002-1520-5549>

**Актуальность.** Создание инструментов редактирования генома эукариот на основе программируемых нуклеаз бактерий из систем CRISPR-Cas открывает обширные перспективы для разработки методов генной терапии, клеточных моделей заболеваний человека, а также изучения проявлений патологического фенотипа, наблюдения за клеточными процессами. От точности внесения двуцепочечных разрывов в целевые участки ДНК зависят безопасность и корректность экспериментов как на клеточном, так и организменном уровнях. Поиск новых вариантов более точных нуклеаз CRISPR-Cas и изучения их способности гидролизовать ДНК в составе нуклеосом *in vivo* представляется актуальной задачей развития технологий геномного редактирования.

**Цель** – провести анализ активности программируемой нуклеазы AsCpf1 (Cas12a), обладающей низким уровнем нецелевой активности, в локусах генома человека, безопасных для внесения трансгенных конструкций (safe harbor), и сравнить с эффективностью широко применяемой нуклеазы SpCas9 в клетках линии HEK293.

**Материал и методы.** Выполнен биоинформационный анализ ассоциации целевых районов с нуклеосомами и другими белками в safe harbor локусах AAVS1 и GSH-Ch1 и транскрипционно неактивном гене MYBPC3 кардиального миозинсвязывающего белка 3 в клетках линии HEK293FT на основе данных, полученных методом ATAC-seq базы NCBI SRA для хроматина клеток линии HEK293FT. Проведено создание и внесение плазмидных конструкций, кодирующих нуклеазы SpCas9 и AsCpf1 и направляющих РНК в клетки HEK293FT. Осуществлен анализ событий в целевых районах генома клеток линии HEK293FT методом изучения секвеннограмм с помощью алгоритма TIDE.

**Результаты.** Изучение данных экспериментов ATAC-seq для клеток HEK293FT показало, что локус AAVS1 можно отнести к открытому хроматину с низкой плотностью нуклеосом, а локус GSH-Ch1 – к закрытому хроматину. В клетках HEK293FT ген кардиального белка MyBPC3 имеет промежуточные характеристики хроматина. Проведенное исследование эффективности внесения разрывов в изучаемые локусы хроматина клеток HEK293FT нуклеазами выявило, что SpCas9 справляется с хроматином любой плотности нуклеосом, тогда как AsCpf1 эффективно вносит разрывы в ДНК только в локусах с открытым хроматином – AAVS1 и MYBPC3. В локусе GSH-Ch1 с высокой плотностью нуклеосом события редактирования происходят на очень низком уровне.

**Заключение.** Показана низкая эффективность нуклеазы AsCpf1 в геномном safe harbor локусе GSH-Ch1, который характеризуется высокой плотностью нуклеосом. При планировании эксперимента по геномному редактированию с помощью нуклеазы AsCpf1 следует учитывать эпигенетический ландшафт хроматина и плотность нуклеосом, а также использовать вещества, влияющие на структуру хроматина.

**Ключевые слова:** редактирование геномов, CRISPR-Cas нуклеазы SpCas9 и AsCpf1, направляющая РНК, нуклеосомы, открытый хроматин

**Для цитирования:** Павлова СВ, Елисафенко ЕА, Шаяхметова ЛШ, Медведев СП. Эффективность программируемых нуклеаз SpCas9 и AsCpf1 (Cas12a) в геномных локусах safe harbor клеток линии HEK293. Альманах клинической медицины. 2021;49(6):385–395. doi: 10.18786/2072-0505-2021-49-037.

Поступила 29.06.2021; принята к публикации 28.07.2021; опубликована онлайн 15.09.2021

<sup>1</sup>ФГБНУ «Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук»; 630090, г. Новосибирск, пр-т Академика Лаврентьева, 10, Российская Федерация

<sup>2</sup>ФГАУ ВО «Новосибирский национальный исследовательский государственный университет»; 630090, г. Новосибирск, ул. Пирогова, 1, Российская Федерация

С появлением эффективных методов редактирования нуклеотидных последовательностей такое направление биомедицины, как исследование функций генов и однонуклеотидных полиморфизмов при нормальной жизнедеятельности клетки и при патологиях, получило активное развитие [1, 2]. В 2007 г. у бактерий и архей была открыта система адаптивной защиты от вирусов и чужеродных молекул ДНК [3]. Геномы прокариот содержат элементы, состоящие из коротких прямых повторов ДНК, между которыми закодированы фрагменты бактериофагов или плазмид. Данные элементы были определены как короткие палиндромные повторы, регулярно расположенные группами (англ. clustered regularly interspaced short palindromic repeats, CRISPR). Повторы CRISPR входят в функциональную единицу бактериального генома и содержат

гены Cas (CRISPR associated), кодирующие белки, необходимые для разрезания ДНК [4, 5]. При попадании «знакомого» вируса в бактерию происходит синтез РНК (англ. crRNA, CRISPR RNA) по матрице ДНК из «коллекции» вирусных фрагментов, закодированных в районе CRISPR, и crRNA адресно (по принципу комплементарности между crRNA и участком ДНК в геноме вируса) направляет нуклеазы Cas для уничтожения вирусной ДНК. Наличие элементов CRISPR в геномах различных бактерий говорит об эволюционно-консервативном принципе РНК-опосредованной деградации инвазивных ДНК, однако механизмы функционирования иммунитета могут иметь различия. Например, если для деградации инвазивных ДНК необходимо несколько белков Cas, система CRISPR-Cas относится к 1-му классу (подразделяется на I, III и IV типы), а если эту функцию выполняет один белок – ко 2-му (II, V и VI типы) [6]. Имеются также различия в механизмах формирования направляющих crRNA для нуклеаз Cas у разных бактерий. В системе CRISPR-Cas9 бактерий *Streptococcus pyogenes* (класс 2, тип II) специфическая направляющая crRNA взаимодействует с tracrPНК (trans-activating crRNA), а в инвазивном геноме, в месте узнавания (протоспейсер, комплементарный crRNA), должен присутствовать смежный мотив ДНК PAM (Protospacer Adjacent Motif) с консенсусом 5'-NGG-3' [7, 8] (рис. 1А). Именно система CRISPR-Cas9 бактерий *S. pyogenes* послужила основой для создания в 2012 г. инструментов редактирования геномов эукариот – были добавлены сигналы ядерной локализации в белок SpCas9 (*S. pyogenes* Cas9), проведена оптимизация кодонов гена *SpCas9* для экспрессии в эукариотических клетках, а tracrRNA и crRNA для удобства были объединены в одну молекулу РНК – single guide RNA (sgRNA) [1, 2].

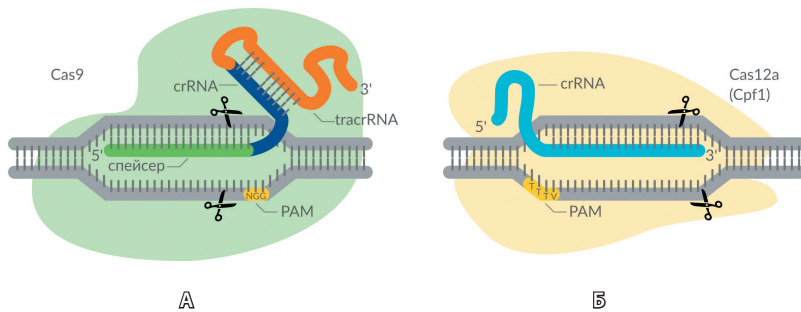
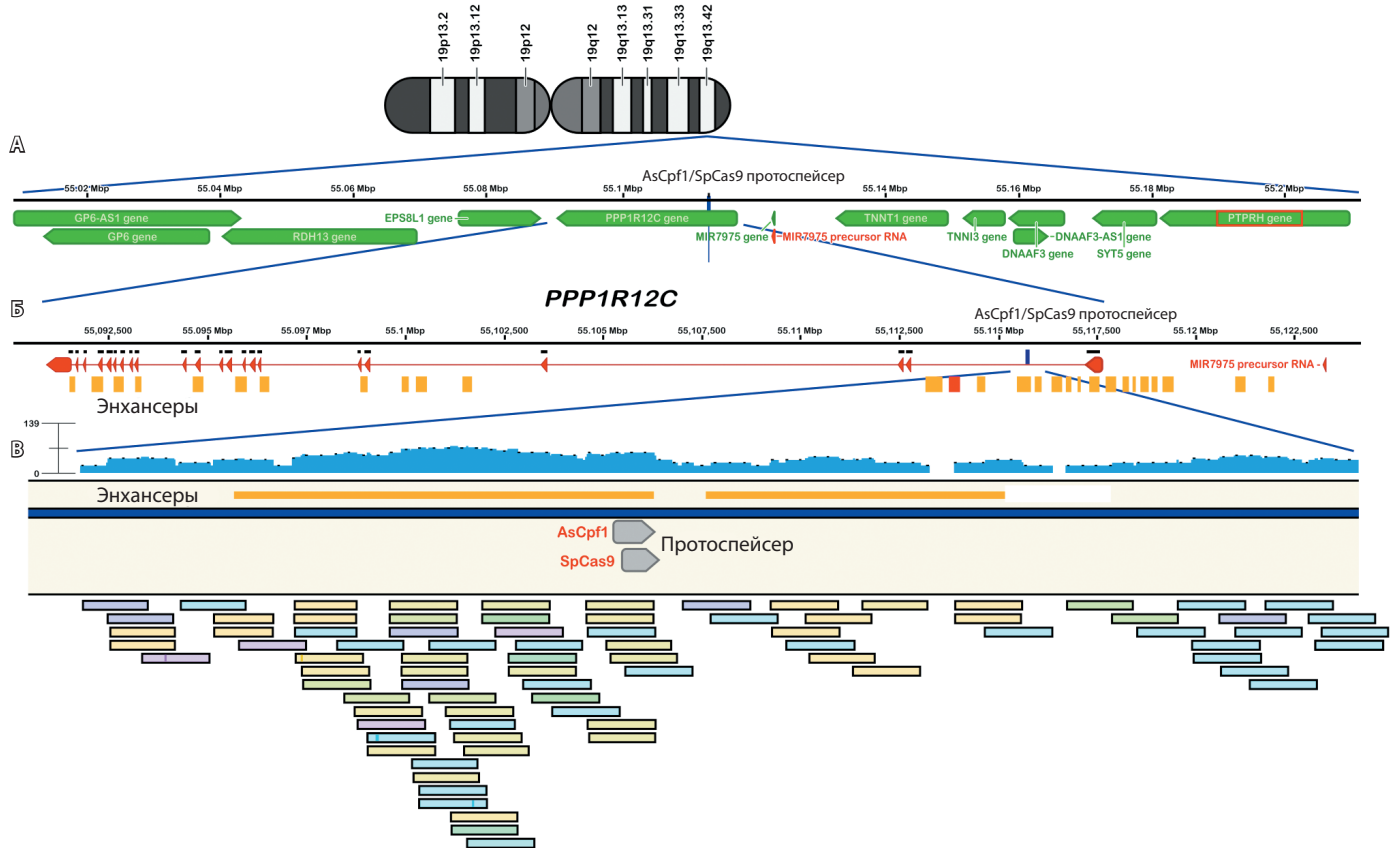


Рис. 1. Схема строения комплекса SpCas9:sgRNA (А) и AsCpf1:crRNA (Б) в целевом районе ДНК, содержащем протоспейсер и PAM (мотив, смежный с протоспейсером)

Таблица 1. Позиции протоспейсеров целевых районов и векторов для экспрессии нуклеаз sgRNA и CRISPR-Cas

Локус; нуклеаза	Позиция протоспейсера, сборка генома GRCh38/hg38	Вектор
AAVS1; Cas9	5'GTCACCAATCCTGTCCCTAGTGG-TGG-3' PAM chr19:55115745-55115764;	PX458 Addgene #48138
AAVS1; Cpf1	PAM5'TTTC-GTCACCAATCCTGTCCCTAG3' chr19:55115744-55115764;	pAsCpf1(TTTV)-2NLS *Addgene 69982/#89352
GSH-Ch1; Cas9	5'TATTATAATAGAAGTGCCTG-AGG-3' PAM chr1:189847623-189847642;	PX458 Addgene #48138
GSH-Ch1; Cpf1	PAM5'TTTA-TACAGTTGCCTTACAAGACA3' chr1:189847581-189847601; -	pAsCpf1(TTTV)-2NLS *Addgene 69982/#89352
MYBPC3; Cas9	5'ACACCACCTGATCATCAACG-AGG-3' PAM chr11:47342656-47342675	PX458 Addgene #48138
MYBPC3; Cpf1	PAM 5'TTCA-AGAAGGACGGGAGAGACACC chr11:47342671-47342691;	pAsCpf1(TYCV)(BB) Addgene #89352

\*Addgene 69982 – прототип плазмиды pAsCpf1(TTTV)-2NLS, куда был добавлен мотив 2NLS из плазмиды Addgene #89352

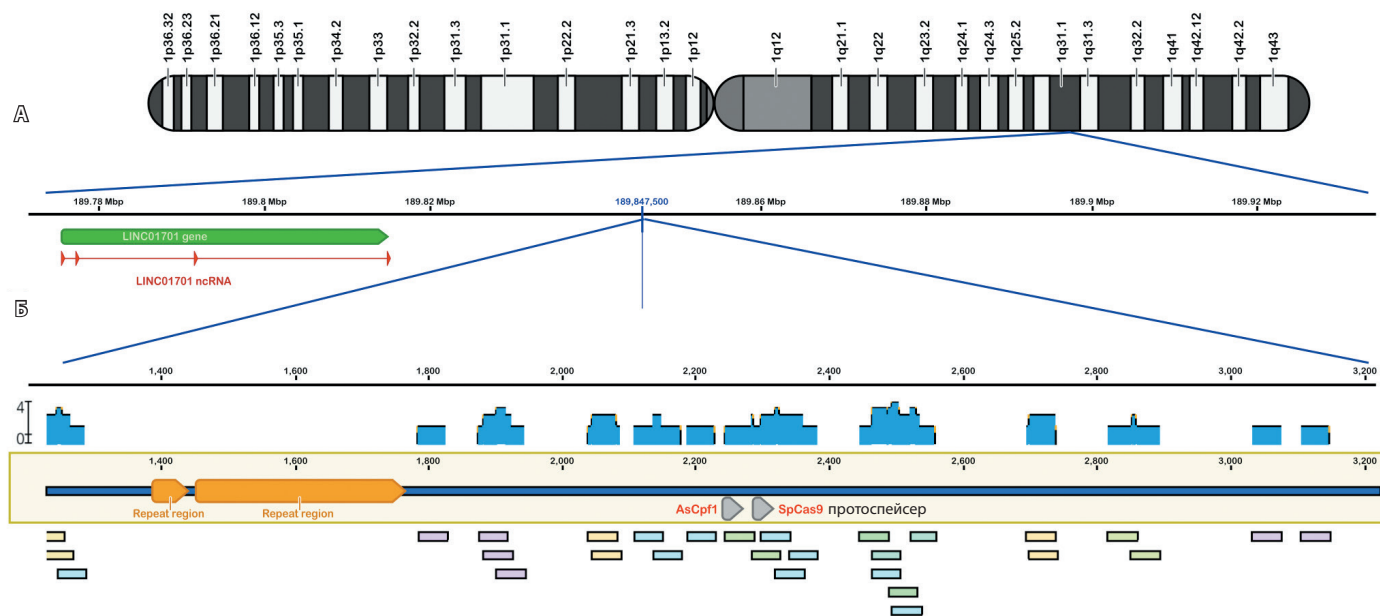


**Рис. 2.** Схема локуса AAVS1 человека, локализованного на 19-й хромосоме (позиция 19q13.42, идеограмма 19-й хромосомы). Ниже расположена схема района (200 т.п.о.), содержащего сайт интеграции аденоассоциированных вирусов (позиция протоспейсера для нуклеаз AsCpf1 и SpCas9). Гены изображены зелеными стрелками, отражающими направление транскрипции (**А**). Экзон-интронная структура гена *PPP1R12C* представлена на **Б**. Позиция протоспейсера в интроне 1 отмечена синим столбиком, экзоны – красные треугольники. Энхансеры обозначены желтыми прямоугольниками, потенциальный промотор – красный прямоугольник. Пример картирования прочтений фрагментов ДНК эксперимента ATAC-seq для клеток HEK293FT (SRR9602331, см. раздел «Материал и методы») показан на **В**. Голубым цветом обозначена плотность прочтений (диапазон представлен числом прочтений, картированных на референсную последовательность локуса AAVS1, – от 1 до 139). Серые стрелки – позиции протоспейсера AsCpf1 и SpCas9, прямоугольники – последовательности ДНК данного района, свободные от белков и выявляемые методом ATAC-seq

более точная и имеет более низкий уровень нецелевой активности по сравнению с Cas9 [10–12]. Нуклеаза AsCpf1 использует одну короткую направляющую РНК (около 40 нуклеотидов длиной), а также Т-богатый PAM с консенсусом 5'-TTTV-3' (V = A, G или C). Поскольку в геномах большинства организмов данное сочетание нуклеотидов встречается с низкой частотой, были созданы мутантные формы AsCpf1, узнающие консенсусы PAM 5'-TUCV-3' и 5'-TATV-3'. Для трех форм AsCpf1 частота встречаемости возможных мишеней в кодирующей части генома человека составляет 1 сайт на 11 пар нуклеотидов [13]. При гидролизе нуклеазой AsCpf1 образуются выступающие «липкие» концы ДНК, а сам разрез располагается в отдалении от PAM (рис. 1Б). В совокупности это делает систему CRISPR-Cpf1 крайне привлекательным инструментом редактирования геномов,

в частности для создания изогенных клеточных моделей наследственных заболеваний.

Для разработки методов генной терапии, создания клеточных моделей заболеваний человека и изучения проявлений патологического фенотипа на молекулярном уровне часто возникает необходимость проведения модификации генома вставками функциональных трансгенов или генетических элементов. Генную терапию на основе ретро- и аденовирусных векторов ранее применяли для лечения некоторых болезней, связанных с метаболическими или гематологическими нарушениями [14–16]. Однако при использовании ретро- и лентивирусных векторов происходит неконтролируемая интеграция вирусов в случайные места генома, что может приводить к инсерционному мутагенезу. Открытие систем редактирования CRISPR-Cas позволяет использовать определенные



**Рис. 3.** Схема локуса GSH-Ch1 человека, локализованного на 1-й хромосоме (позиция 1q31.1, идеограмма 1-й хромосомы). Ниже расположена схема района (200 т.п.о.), позиция протоспейсеров для нуклеаз AsCpf1 и SpCas9 отмечена синим столбиком. Зеленой стрелкой обозначен ген длинной некодирующей РНК *LINC01701* (А). Пример картирования прочтений фрагментов ДНК эксперимента ATAC-seq для клеток HEK293FT (SRR9602331, см. раздел «Материал и методы») дан на Б. Голубым цветом обозначена плотность прочтений (число прочтений, картированных на референсную последовательность локуса GSH-Ch1, – от 1 до 4). Серые стрелки – позиции протоспейсеров AsCpf1 и SpCas9, прямоугольники – последовательности ДНК данного района, свободные от белков и выявляемые методом ATAC-seq. Повторы ДНК представлены желтыми прямоугольниками

**Таблица 2.** Праймеры для амплификации целевых районов ДНК с целью анализа нуклеазной активности SpCas9 и AsCpf1

Название праймера	Структура праймера	Температура отжига праймера; размер ПЦР продукта	Целевой район амплификации ДНК
AAVS1_WT-F AAVS1_WT-R	5'CTCTGGCTCCATCGTAAGCAA3' 5'CCCAAAGTACCCCGTCTCCC3'	58С 555 п.о.	Локус AAVS1
1Chr-SHL-T-F 1Chr-SHL-T-R	5'GGTAGTGACAGAGCTAACATG3' 5'GTCTAACTATGACAGTGCTGCC3'	58С 490 п.о.	Локус SH-Chr1
MYBPC3-LD-F MYBPC3-R	5'CAAATGGTGAGTCCAGAAGC3' 5'GAGATGAGAAGGATGAGGTTTAGG3'	58С 405 п.о.	Ген <i>MYBPC3</i>

ПЦР – полимеразная цепная реакция

места генома, где взаимное влияние внесенных генетических конструкций и хозяйского генома минимально, – так называемые genome safe harbor (GSH) локусы. Наиболее популярное место для интеграции трансгенов в геном человека – локус AAVS1 (adeno-associated virus integration site 1) на 19-й хромосоме (19q13.42), в первом интроне гена *PPP1R12C* (protein phosphatase 1 regulatory subunit 12C), который кодирует белок, регулирующий активность фосфатазы 1δ [17] (рис. 2А).

Локус AAVS1 был выявлен экспериментально при анализе сайтов интеграции аденоассоциированных вирусов в ДНК клеток человека *in vitro*. Примечательно, что нарушение одного или двух

аллелей этого гена не приводит к отклонениям ни на уровне культуры клеток, ни на уровне целого организма, хотя *PPP1R12C* экспрессируется во всех тканях [17]. Тем не менее следует учитывать, что данный район 19-й хромосомы не может быть безопасным локусом safe harbor, так как обогащен кодирующими генами, в том числе там находится ген *PTPRH* (protein tyrosine phosphatase receptor type H), мутации в котором ассоциированы с раком печени (см. рис. 2А), и внесение чужеродных генетических элементов в данный локус потенциально может приводить к непредсказуемым событиям. Были сформулированы критерии для «идеальных» GSH-локусов: расстояние >50 т.п.о.





**Таблица 3.** Покрытие прочтений ДНК в целевых районах генома клеток HEK293, выявляемых в экспериментах ATAC-seq, в процентном соотношении от данных референсного гена *ACTB*

ATAC-seq	SRR6418076 HEK293	SRR8938025 HEK293	SRR9602331 HEK293	SRR10982335 Кардиомиоциты из ИПСК человека	SRR12503731 Кардиомиоциты из ИПСК человека	Исследуемый район (позиции в геноме, сборка hg38)
Покрытие <i>ACTB</i> , %	100	100	100	100	100	chr7 5,527,748-5,529,657 (1910 п.о.)
Покрытие GSH-Ch1, %	8,2	3,9	3,55	56,5	59,6	chr1 189847081-189848101 (1000 п.о.)
Покрытие <i>MYBPC3</i> , %	27,1	6,6	11,9	11,9	50,0	chr11 47342156-47343175 (1000 п.о.)
Покрытие <i>AAVS1</i> , %	68,9	23,3	31,0	33,8	33,3	chr19 55115245-55116264 (1000 п.о.)

ИПСК – индуцированные плюрипотентные стволовые клетки

от 5'-конца любого гена, более 300 т.п.о. от генов, ассоциированных с раком и любой микроРНК, удаленность от ультраконсервативных регионов генома. Таким идеальным кандидатом, отвечающим всем критериям, считается локус GSH-Ch1 (genomic safe harbor chromosome 1) (позиции 189847581–189847601, район 1q31.1 1-й хромосомы человека) [17, 18] (рис. 3А).

Системы CRISPR-Cas представляют собой «изобретение» бактерий для борьбы с вирусной ДНК, которая не образует регулярные комплексы с белками, тогда как в клетках эукариот ДНК присутствует в составе нуклеосом и других структурных белков хроматина. В экспериментах *in vitro* показано, что нуклеосомы действительно препятствуют гидролизу ДНК нуклеазами SpCas9 [19] и AsCpf1 (Cas12a) [20]. При этом *in vivo* SpCas9 справляется с хроматином любой плотности и доступности [21]. Для нуклеазы AsCpf1 этот аспект недостаточно изучен. В данном исследовании мы провели биоинформационный анализ уровня открытости хроматина в safe harbor локусах *AAVS1* и GSH-Ch1 и транскрипционно неактивном гене *MYBPC3* кардиального миозинсвязывающего белка 3 в клетках линии HEK293FT, а также оценили нуклеазную активность белков SpCas9 и AsCpf1 *in vivo*.

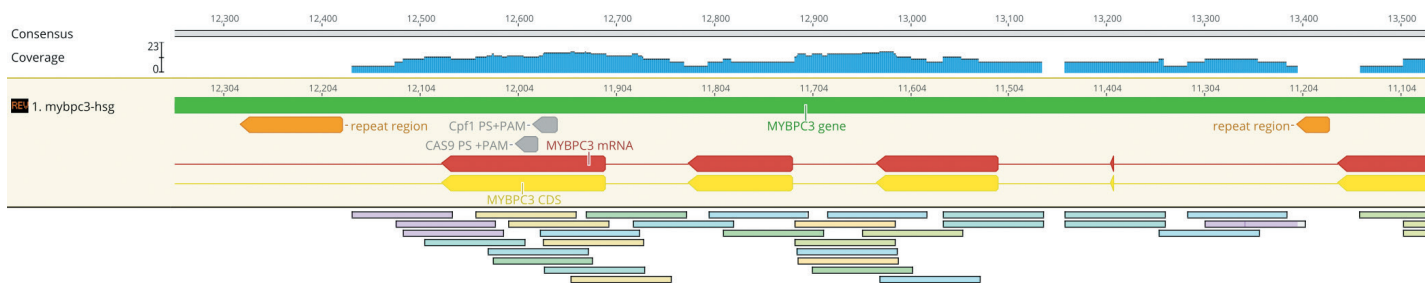
## Материал и методы

Получение генетических конструкций, экспрессирующих компоненты системы CRISPR-Cas. Дизайн протоспейсеров для синтеза направляющих РНК нуклеаз SpCas9 и AsCpf1 выполняли с помощью ресурсов Benchling (<https://www.benchling.com>). Олигонуклеотиды были синтезированы компанией «Биоссет» (Россия) и клонированы в соответствующие плазмиды, кодирующие

sgRNA и нуклеазы SpCas9 (pX458) и AsCpf1 (pAsCpf1(TYCV)(BB), pAsCpf1-2NLS) в соответствии с протоколом [22] (табл. 1). Плазида pAsCpf1(TTTV)-2NLS была сконструирована в лаборатории на основе AsCpf1 (TTTV) путем добавления второго сайта ядерной локализации NLS.

Анализ эффективности редактирования CRISPR-Cas на клетках HEK293FT

Доставку плазмидных конструкций в клетки линии HEK293FT (cat. R70007 Thermo Fisher Scientific) проводили с помощью электропорации на приборе Neon Transfection System (Thermo Fisher Scientific) согласно рекомендациям производителя. Клетки линии HEK293FT культивировали при 5% CO<sub>2</sub>, 37 °C, в ростовой среде DMEM/F12 (Gibco), 10% фетальной бычьей сыворотке (Life Technologies), 2 mM GlutaMax (Thermo Fisher Scientific), 100 ед/мл раствора пенициллина/стрептомицина (Life Technologies), 0,1 mM раствора аминокислот NEAA (Invitrogen). Перед экспериментом клетки диссоциировали реагентом TrypLE, центрифугировали в фосфатном буфере PBS в течение 5 минут (300 g) и ресуспендировали в буфере R (Neon Transfection System 10ul kit, Thermo Fisher Scientific) из расчета конечной концентрации 5 × 10<sup>7</sup>/мл. Реакцию электропорации 0,5 мкг плазмиды, кодирующей компоненты CRISPR-Cas, проводили в 10 мкл суспензии клеток в буфере R. После электропорации клетки перенесли в подготовленные культуральные планшеты с теплой средой без антибиотика. Через 48 часов после трансфекции из клеток выделяли геномную ДНК и амплифицировали целевой участок, содержащий сайт CRISPR-индуцируемого разрыва, со специфических праймеров (табл. 2). Полученный продукт полимеразной цепной реакции (ПЦР)



**Рис. 4.** Пример картирования прочтений фрагментов ДНК эксперимента ATAC-seq для клеток HEK293FT (SRR9602331, см. раздел «Материал и методы») на ген *MYBP3*. Голубым цветом обозначена плотность прочтений (диапазон представлен числом прочтений, картированных на референсную последовательность гена *MYBP3*, – от 1 до 23). Серые стрелки – позиции протоспейсеров AsCpf1 и SpCas9, прямоугольники – последовательности ДНК данного района, свободные от белков и выявляемые методом ATAC-seq. Красные стрелки – экзоны гена, желтые – экзоны, кодирующие белок

очищали методом элюции из агарозного геля с помощью набора Cleanup Standard (Евроген). Реакции секвенирования по методу Сэнгера проводили с применением набора реагентов BigDye Terminator v3.1 Cycle Sequencing Kit (Thermo Fisher Scientific). Продукты реакции секвенирования анализировали в Центре коллективного пользования «Геномика» СО РАН.

Общую эффективность CRISPR-Cas9 оценивали по доле модифицированных последовательностей ДНК, образовавшихся в результате негомологичного сшивания концов, выявляемых при анализе секвенограмм с помощью программного алгоритма TIDE (Tracking of Indels by Decomposition) (<https://tide.nki.nl>).

Анализ открытости хроматина в целевых районах генома клеток HEK293FT

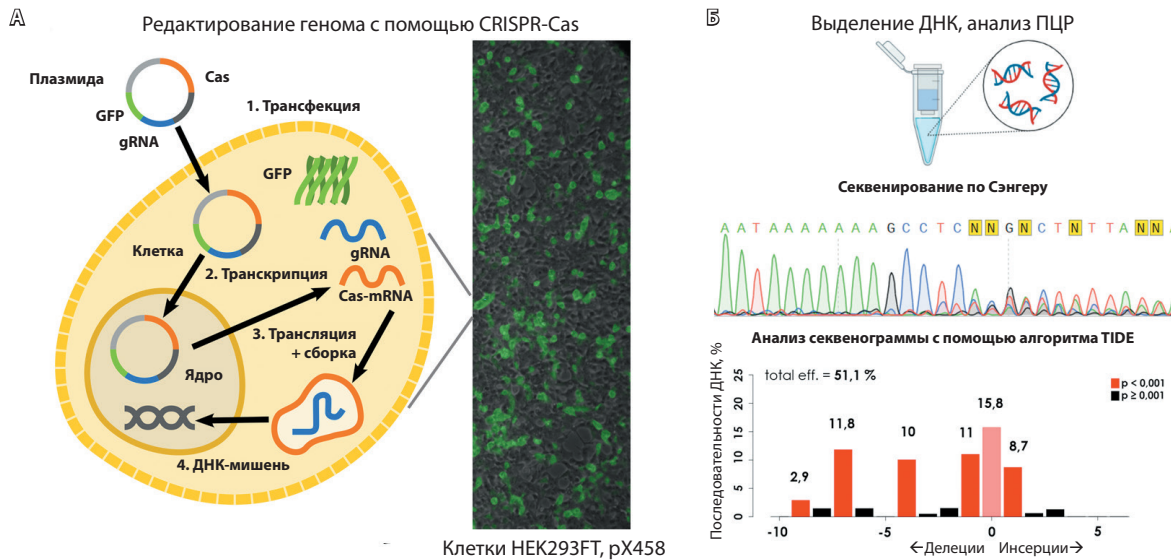
Последовательности локусов были экстрагированы из GenBank (<https://www.ncbi.nlm.nih.gov/nucleotide>): AP023479.1:55589039-55593668 Homo sapiens DNA, chromosome 19, AL591504.5:14265-18238 Human DNA sequence from clone RP11-445J9 on chromosome 1. Анализ открытости хроматина в клетках HEK293FT проводили на основе данных из базы NCBI SRA (Sequence Read Archive; <http://www.ncbi.nlm.nih.gov/sra>), полученных методом ATAC-seq (Assay for Transposase-Accessible Chromatin using sequencing) [23]. Использовали данные экспериментов ATAC-seq для клеток HEK293FT SRR6418076, SRR8938025, SRR9602331 и для кардиомиоцитов человека, полученных из индуцированных плюрипотентных стволовых клеток (ИПСК) SRR10982335, SRR12503731. Данные экспериментов ATAC-seq картировали относительно референсных последовательностей ДНК с использованием программы HISAT2 [24]. Результирующие файлы в формате SAM визуализировали и проводили первичный анализ

с использованием программного обеспечения Geneious 11.0.2 (<http://www.geneious.com>) [25]. Участки, содержащие повторы ДНК, исключали из анализа с помощью программы RepeatMasker [26].

## Результаты

Для исследования программируемых нуклеаз SpCas9 и AsCpf1 (Cas12a) в safe harbor локусах генома клеток линии HEK293 были выбраны протоспейсеры для локуса AAVS1 в районе hg38 chr19:55115745-55115764 (интрон гена *PPP1R12C*) (см. рис. 1А) [27] и для локуса GSH-Ch1 в позиции hg38 chr1:189847581-189847601 (см. рис. 2А) [28]. Дополнительно для изучения активности нуклеаз в транскрипционно неактивном хроматине был выбран ген кардиального миозинсвязывающего белка сMyBP-C, который не экспрессируется в клетках HEK293FT (протоспейсер hg38 chr11:47342656-47342675) (рис. 4).

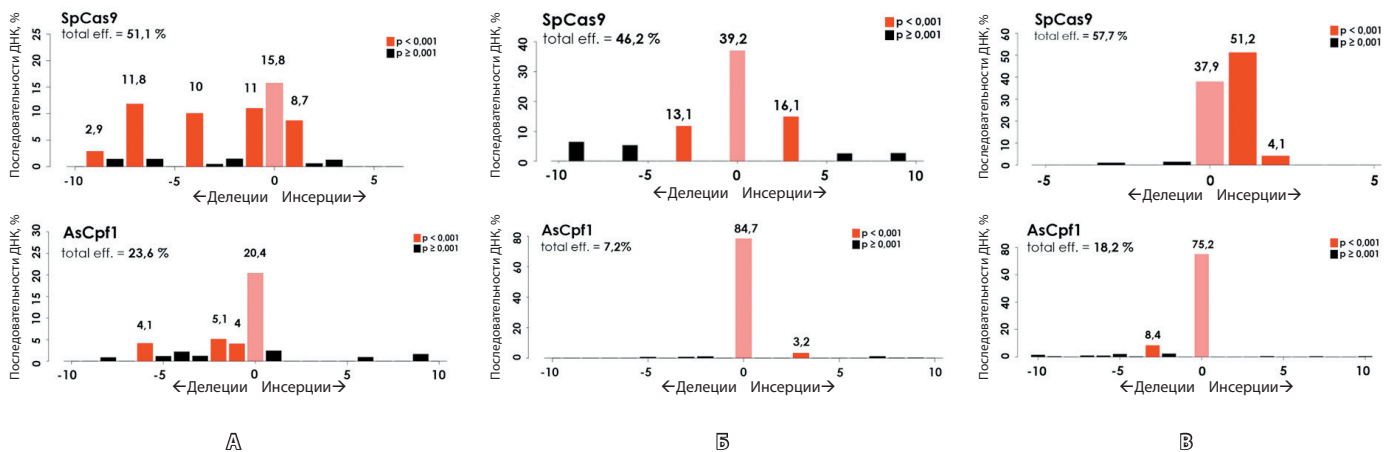
Предварительно проводили оценку степени ассоциации целевых районов с нуклеосомами и другими белками с помощью анализа данных, полученных методом ATAC-seq базы NCBI SRA для хроматина клеток линии HEK293FT и кардиомиоцитов. Метод ATAC-seq выявляет доступные участки ДНК в составе хроматина с помощью гиперактивной мутантной формы транспозазы Tn5, которая интегрирует линкеры ДНК в открытые участки генома. Фрагменты генома, содержащие линкеры, очищают, амплифицируют с помощью ПЦР, после чего определяют последовательность ДНК с помощью секвенирования нового поколения (Illumina). На основании анализа прочтений ДНК можно выявить открытые участки хроматина, сайты связывания транскрипционных факторов, а также позиции нуклеосом. Чем более открыт хроматин, тем больше прочтений ДНК приходится на соответствующий участок генома [29, 30].



**Рис. 5.** Схема эксперимента. **А** – плазмиды, кодирующие компоненты систем CRISPR-SpCas9 и CRISPR-AsCpf1, трансфицировали в клетки HEK293FT. *In vivo* в ядрах клеток происходила транскрипция ДНК плазмиды и синтез направляющей РНК (gRNA), мРНК нуклеаз Cas (Cas-mRNA) и флуоресцентного белка GFP. В цитоплазме происходила трансляция мРНК и сборка комплексов CRISPR-Cas:gRNA, которые перемещались в ядро и вносили двуцепочечные разрывы в ДНК целевых регионов (ДНК-мишень). По наличию сигнала белка GFP, закодированного в плазмиде рХ458, определяли уровень трансфекции плазмид в клетки линии HEK293FT. **Б** – выделяли ДНК клеток линии HEK293FT, амплифицировали целевой район методом полимеразной цепной реакции (ПЦР), определяли последовательность нуклеотидов ДНК по Сэнгеру и анализировали секвеннограммы при помощи алгоритма TIDE

В данном исследовании мы оценили количество прочтений ДНК (плотность покрытия) в локусах AAVS1, GSH-Ch1 и гене MYBPC3 в аннотируемых экспериментах ATAC-seq, проведенных на клетках HEK293 (3 варианта) и кардиомиоцитах, полученных из ИПСК (2 варианта). Полученные данные для каждого эксперимента нормировали на

плотность покрытия, выявляемую для хроматина кодирующей области гена домашнего хозяйства ACTB (beta-actin). Из данных табл. 3 видно, что в клетках HEK293 локус GSH-Ch1, вероятнее всего, относится к закрытому хроматину, поскольку плотность прочтений ДНК составляет в среднем 5% от плотности прочтений, картируемых в гене



**Рис. 6.** Результаты анализа секвеннограмм целевых регионов ДНК локусов AAVS1 (**А**), GSH-Ch1 (**Б**), MYBPC3 (**В**) клеток линии HEK293FT после внесения плазмид, кодирующих компоненты CRISPR-SpCas9 и CRISPR-AsCpf1, с помощью алгоритма TIDE. На оси абсцисс в нуле показано отсутствие событий редактирования, слева от нуля отложены делеции, справа от нуля – инсерции, цифрами указано, сколько нуклеотидов удалено или вставлено. По оси ординат отложен процент выявленных событий

домашнего хозяйства *ACTB*, хроматин которого мы считаем открытым. Разброс значений покрытия прочтений ДНК для локуса *AAVS1* составляет от 23,3 до 68,9% относительно гена *ACTB*.

Наряду с этим в локусе *AAVS1* (интрон 1 гена *PPP1R12C*) с помощью программы BLAT (<http://genome.ucsc.edu>) были выявлены потенциальные сайты регуляторных элементов cCREs (candidate cis-Regulatory Elements), а именно потенциальные энхансеры (см. рис. 2Б). Элементы cCREs выявляются в районах открытого хроматина (свободного от нуклеосом), гиперчувствительных к ДНКазе 1 (см. рис. 2Б, В). Таким образом, можно считать, что район *AAVS1*, в котором расположен протоспейсер в клетках НЕК293, содержит открытый хроматин, доступный для связывания белков комплекса CRISPR-Cas, а район safe harbor на первой хромосоме GSH-Ch1 – закрытый хроматин, который может быть проблемным для активности AsCpf1. Ген *MYBPC3* неактивен в клетках НЕК293, однако значения покрытия прочтений ДНК в экспериментах ATAC-seq в этом случае варьируют от 6,6 до 27,1% относительно гена домашнего хозяйства *ACTB*. Мы оценили плотность прочтений в экспериментах ATAC-seq, проведенных на кардиомиоцитах, полученных из ИПСК, в гене *MYBPC3*, который экспрессируется в клетках. Анализ данных показывает, что транскрипционно активный ген *MYBPC3* действительно приобретает черты открытого хроматина. Интересно, что хроматин локуса GSH-Ch1 в кардиомиоцитах также становится открытым в отличие от НЕК293 (см. табл. 3).

Примеры картирования прочтений в исследуемых районах клеток НЕК293FT даны на рис. 2В (локус *AAVS1*), 3Б (локус GSH-Ch1), 4 (ген *MYBPC3*).

Протоспейсеры для нуклеаз SpCas9 и AsCpf1 были клонированы в соответствующие плазмиды (см. табл. 2), и с помощью электропорации проведена доставка кодирующих конструкций в клетки НЕК293FT. Уровень трансфекции оценивали по флуоресценции белка EGFP, который присутствует в плазмиде pX458, кодирующей белок SpCas9 и направляющей sgRNA (доставка конструкции происходила в 70% клеток). Уровень доставки плазмид, кодирующих компоненты CRISPR-AsCpf1, считали таким же. Далее геномную ДНК выделяли спустя 48 часов после трансфекции, целевые районы амплифицировали и секвенировали. Полученные данные секвенограмм были проанализированы с помощью алгоритма TIDE (рис. 5). Программа статистически оценивает вероятность присутствия делеций и инсерций, появившихся в районе протоспейсера после негомологичной репарации разрывов ДНК, и представляет результаты в виде

диаграмм, на которых можно оценить степень активности системы CRISPR-Cas в целевом районе.

Как и ожидалось, в локусе *AAVS1* нуклеазы SpCas9 и AsCpf1 эффективно взаимодействовали с хроматином (рис. 6А). По данным программы TIDE, общая эффективность редактирования составляла 51,1% для SpCas9 и 23,6% для AsCpf1. В локусе GSH-Ch1 была активна исключительно SpCas9 (46,2%), в то время как нуклеаза AsCpf1 вносила разрывы в хроматин с очень низкой эффективностью – 7,2% (рис. 6Б). В целевом районе гена *MYBPC3* эффективность редактирования для SpCas9 была равна 57,7%, для AsCpf1 – 18,2% (рис. 6В).

### Обсуждение и заключение

Открытие программированных нуклеаз CRISPR-Cas и создание на их основе инструментов редактирования генома эукариот позволяют эффективно изучать мутации и создавать модели для изучения наследственных заболеваний. Для безопасного и эффективного трансгенеза в геноме человека выявлены специальные GSH-локусы. С целью уменьшения вероятности ошибок редактирования подбирают более точные нуклеазы Cas, оценивают разные аспекты – частоту встречаемости PAM, эффективность нуклеазной активности в различных районах хроматина. Система редактирования CRISPR-Cas была создана на основе механизмов иммунитета бактерий, основанного на специфическом узнавании и расщеплении ДНК вирусом и не была адаптирована для расщепления ДНК в составе нуклеосом, которые служат структурной единицей хроматина эукариот. Участки ДНК, свободные от нуклеосом эукариот (так называемый открытый хроматин), встречаются в промоторах и энхансерах генов – сайтах связывания транскрипционных факторов или регуляторных белков. Открытый хроматин также присутствует в экзонах и интронах активно транскрибируемых генов [31]. Участки ДНК, свободные от нуклеосом, в активных генах образуются в результате динамического процесса взаимодействия ДНК и белков нуклеосом (октамера гистонов 2xH2A, 2xH2B, 2xH3 и 2xH4), так называемого дыхания нуклеосом за счет привлечения белков ремоделирования хроматина. Данные белки узнают активные модификации гистонов и способствуют уменьшению сродства между ДНК и гистонами, за счет чего ДНК чаще остается свободной от нуклеосом. Эпигенетические модификации гистонов активного хроматина (например, ацетилованный лизин гистона H3 (H3K9ac, H3K27ac) в положении 9 или 27,





ди- и триметилированный лизин гистона H3 (H3K4me2-me3) в положении 4 и др.) считаются маркерами открытого хроматина [31, 32].

Современным, доступным и чувствительным методом полногеномного анализа участков ДНК, свободных от нуклеосом, считается анализ ATAC-seq. Данные, полученные ранее с помощью других методов исследования структуры хроматина, – DNase-seq (DNase I hypersensitive sites sequencing – анализ гиперчувствительных к расщеплению ДНКазой районов генома), MNase-seq (Micrococcal Nuclease digestion with deep sequencing – анализ участков, чувствительных к микрококковой нуклеазе), ChIP-seq (Chromatin Immunoprecipitation Sequencing – анализ ДНК-белковых взаимодействий, основанный на иммунопреципитации хроматина для анализа распределения модифицированных гистонов) – полностью согласуются с результатами ATAC-seq. С помощью метода ATAC-seq можно выявлять области открытого хроматина, уточнять позиции нуклеосом или сайтов связывания регуляторных элементов с высокой точностью [23, 30, 33].

Мы исследовали нуклеазную активность SpCas9 и AsCpf1 в клетках HEK293FT в safe harbor локусах на 19-й хромосоме (AAVS1) и на 1-й хромосоме (GSH-Ch1). Изучение данных экспериментов ATAC-seq для клеток HEK293FT показало, что локус AAVS1 можно отнести к открытому хроматину, а локус GSH-Ch1 – к закрытому. Ген кардиального белка MyBPC3 в клетках HEK293FT имеет промежуточные характеристики хроматина. Проведенное исследование эффективности

внесения разрывов в исследуемые локусы хроматина клеток HEK293FT нуклеазами позволило обнаружить, что SpCas9 справляется с хроматином любой плотности, тогда как AsCpf1 эффективно вносит разрывы в ДНК только в локусах с открытым хроматином – AAVS1 и MYBPC3. В локусе с закрытым хроматином GSH-Ch1 события редактирования в данном случае происходят на очень низком уровне.

Как было показано нами в данной работе, а также другими исследователями, на нуклеазную активность SpCas9 *in vivo* структура хроматина не оказывает влияния [21]. Обнаруженная низкая активность AsCpf1 в GSH-Ch1 может быть свидетельством того, что плотный хроматин данного локуса в клетках HEK293FT препятствует внесению разрывов в ДНК данной нуклеазой *in vivo*. Существует специфический паттерн распределения нуклеосом в целевых районах в разных типах клеток. Анализируя эксперименты ATAC-seq, проведенные на кардиомиоцитах, дифференцированных из ИПСК, мы обнаружили, что MYBPC3 хроматина становится более открытым не только в транскрипционном активном гене, но и в локусе GSH-Ch1.

Таким образом, чтобы использовать для редактирования более точную программируемую нуклеазу AsCpf1, следует тщательно выбирать район протоспейсера, учитывая эпигенетический ландшафт хроматина и степень его открытости в разных типах клеток, либо применять в эксперименте вещества и условия культивирования клеток, влияющие на структуру хроматина. ☺

## Дополнительная информация

### Финансирование

Работа проведена без привлечения дополнительного финансирования со стороны третьих лиц.

### Конфликт интересов

Авторы декларируют отсутствие явных и потенциальных конфликтов интересов, связанных с публикацией настоящей статьи.

### Участие авторов

С.В. Павлова – дизайн и проведение экспериментальной части исследования, интерпретация результатов, написание текста; Е.А. Елисафенко – проведение биоинформатического анализа данных, анализ результатов исследования, редактирование текста

статьи; Л.Ш. Шаяхметова – проведение экспериментальной части исследования, редактирование рукописи; С.П. Медведев – концепция и дизайн исследования, интерпретация результатов, написание текста, утверждение итогового варианта текста рукописи. Все авторы прочли и одобрили финальную версию статьи перед публикацией, согласны нести ответственность за все аспекты работы и гарантируют, что ими надлежащим образом были рассмотрены и решены вопросы, связанные с точностью и добросовестностью всех частей работы.

### Благодарности

Работа поддержана Российским фондом фундаментальных исследований и Новосибирской областью, номер проекта 19-44-540002 р\_а. Бюджетный проект 0259-2021-0011.

## Литература / References

- Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, Norville JE, Church GM. RNA-guided human genome engineering via Cas9. *Science*. 2013;339(6121):823–826. doi: 10.1126/science.1232033.
- Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, Zhang F. Multiplex genome engineering using CRISPR/Cas systems. *Science*. 2013;339(6121):819–823. doi: 10.1126/science.1231143.
- Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, Romero DA, Horvath P. CRISPR provides acquired resistance against viruses in prokaryotes. *Science*. 2007;315(5819):1709–1712. doi: 10.1126/science.1138140.



4. Mojica FJ, Díez-Villaseñor C, Soria E, Juez G. Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria. *Mol Microbiol.* 2000;36(1):244–246. doi: 10.1046/j.1365-2958.2000.01838.x.
5. Mojica FJ, Díez-Villaseñor C, García-Martínez J, Soria E. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J Mol Evol.* 2005;60(2): 174–182. doi: 10.1007/s00239-004-0046-3.
6. Koonin EV, Makarova KS, Zhang F. Diversity, classification and evolution of CRISPR-Cas systems. *Curr Opin Microbiol.* 2017;37:67–78. doi: 10.1016/j.mib.2017.05.008.
7. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science.* 2012;337(6096): 816–821. doi: 10.1126/science.1225829.
8. Doudna JA, Charpentier E. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science.* 2014;346(6213): 1258096. doi: 10.1126/science.1258096.
9. Fu Y, Foden JA, Khayter C, Maeder ML, Reyon D, Joung JK, Sander JD. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat Biotechnol.* 2013;31(9): 822–826. doi: 10.1038/nbt.2623.
10. Kleinstiver BP, Tsai SQ, Prew MS, Nguyen NT, Welch MM, Lopez JM, McCaw ZR, Aryee MJ, Joung JK. Genome-wide specificities of CRISPR-Cas Cpf1 nucleases in human cells. *Nat Biotechnol.* 2016;34(8):869–874. doi: 10.1038/nbt.3620.
11. Zetsche B, Gootenberg JS, Abudayyeh OO, Slaymaker IM, Makarova KS, Essletzbichler P, Volz SE, Joung J, van der Oost J, Regev A, Koonin EV, Zhang F. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell.* 2015;163(3):759–771. doi: 10.1016/j.cell.2015.09.038.
12. Kim D, Kim J, Hur JK, Been KW, Yoon SH, Kim JS. Genome-wide analysis reveals specificities of Cpf1 endonucleases in human cells. *Nat Biotechnol.* 2016;34(8):863–868. doi: 10.1038/nbt.3609.
13. Gao L, Cox DBT, Yan WX, Manteiga JC, Schneider MW, Yamano T, Nishimasu H, Nureki O, Crosetto N, Zhang F. Engineered Cpf1 variants with altered PAM specificities. *Nat Biotechnol.* 2017;35(8):789–792. doi: 10.1038/nbt.3900.
14. Rivière I, Dunbar CE, Sadelain M. Hematopoietic stem cell engineering at a crossroads. *Blood.* 2012;119(5):1107–1116. doi: 10.1182/blood-2011-09-349993.
15. Gaspar HB, Cooray S, Gilmour KC, Parsley KL, Zhang F, Adams S, Björkregren E, Bayford J, Brown L, Davies EG, Veys P, Fairbanks L, Bordon V, Petropoulou T, Kinnon C, Thrasher AJ. Hematopoietic stem cell gene therapy for adenosine deaminase-deficient severe combined immunodeficiency leads to long-term immunological recovery and metabolic correction. *Sci Transl Med.* 2011;3(97):97ra80. doi: 10.1126/scitranslmed.3002716.
16. Cartier N, Hacein-Bey-Abina S, Bartholomae CC, Veres G, Schmidt M, Kutschera I, Vidaud M, Abel U, Dal-Cortivo L, Caccavelli L, Mahlaoui N, Kiermer V, Mittelstaedt D, Bellesme C, Lahlou N, Lefrère F, Blanche S, Audit M, Payen E, Leboulch P, l'Homme B, Bougnères P, Von Kalle C, Fischer A, Cavazzana-Calvo M, Aubourg P. Hematopoietic stem cell gene therapy with a lentiviral vector in X-linked adrenoleukodystrophy. *Science.* 2009;326(5954):818–823. doi: 10.1126/science.1171242.
17. Sadelain M, Papapetrou EP, Bushman FD. Safe harbours for the integration of new DNA in the human genome. *Nat Rev Cancer.* 2011;12(1): 51–58. doi: 10.1038/nrc3179.
18. Pellenz S, Phelps M, Tang W, Hovde BT, Sinit RB, Fu W, Li H, Chen E, Monnat RJ Jr. New Human Chromosomal Sites with "Safe Harbor" Potential for Targeted Transgene Insertion. *Hum Gene Ther.* 2019;30(7):814–828. doi: 10.1089/hum.2018.169.
19. Horlbeck MA, Witkowski LB, Guglielmi B, Replogle JM, Gilbert LA, Villalta JE, Torigoe SE, Tjian R, Weissman JS. Nucleosomes impede Cas9 access to DNA in vivo and in vitro. *Elife.* 2016;5:e12677. doi: 10.7554/eLife.12677.
20. Strohkendl I, Saifuddin FA, Gibson BA, Rosen MK, Russell R, Finkelstein IJ. Inhibition of CRISPR-Cas12a DNA targeting by nucleosomes and chromatin. *Sci Adv.* 2021;7(11):eabd6030. doi: 10.1126/sciadv.abd6030.
21. Barkal AA, Srinivasan S, Hashimoto T, Gifford DK, Sherwood RI. Cas9 Functionally Opens Chromatin. *PLoS One.* 2016;11(3):e0152683. doi: 10.1371/journal.pone.0152683.
22. Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. Genome engineering using the CRISPR-Cas9 system. *Nat Protoc.* 2013;8(11):2281–2308. doi: 10.1038/nprot.2013.143.
23. Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods.* 2013;10(12):1213–1218. doi: 10.1038/nmeth.2688.
24. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods.* 2015;12(4):357–360. doi: 10.1038/nmeth.3317.
25. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics.* 2012;28(12): 1647–1649. doi: 10.1093/bioinformatics/bts199.
26. Smit AFA, Hubley R, Green P. RepeatMasker Open-4.0. 2013–2015. Available from: <http://www.repeatmasker.org>.
27. Ustyantseva EI, Medvedev SP, Vetchinova AS, Minina JM, Illarionovskiy SN, Zakian SM. A Platform for Studying Neurodegeneration Mechanisms Using Genetically Encoded Biosensors. *Biochemistry (Mosc).* 2019;84(3):299–309. doi: 10.1134/S000629791903012X.
28. Kimura Y, Shofuda T, Higuchi Y, Nagamori I, Oda M, Nakamori M, Onodera M, Kanematsu D, Yamamoto A, Katsuma A, Suemizu H, Nakano T, Kanemura Y, Mochizuki H. Human Genomic Safe Harbors and the Suicide Gene-Based Safeguard System for iPSC-Based Cell Therapy. *Stem Cells Transl Med.* 2019;8(7): 627–638. doi: 10.1002/sctm.18-0039.
29. Schep AN, Buenrostro JD, Denny SK, Schwartz K, Sherlock G, Greenleaf WJ. Structured nucleosome fingerprints enable high-resolution mapping of chromatin architecture within regulatory regions. *Genome Res.* 2015;25(11):1757–1770. doi: 10.1101/gr.192294.115.
30. Buenrostro JD, Wu B, Chang HY, Greenleaf WJ. ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. *Curr Protoc Mol Biol.* 2015;109:21.29.1–21.29.9. doi: 10.1002/0471142727.mb2129s109.
31. Boyle AP, Davis S, Shulha HP, Meltzer P, Margulies EH, Weng Z, Furey TS, Crawford GE. High-resolution mapping and characterization of open chromatin across the genome. *Cell.* 2008;132(2):311–322. doi: 10.1016/j.cell.2007.12.014.
32. Igolkina AA, Zinkevich A, Karandasheva KO, Popov AA, Selifanova MV, Nikolaeva D, Tkachev V, Penzar D, Nikitin DM, Buzdin A. H3K4me3, H3K9ac, H3K27ac, H3K27me3 and H3K9me3 Histone Tags Suggest Distinct Regulatory Evolution of Open and Condensed Chromatin Landmarks. *Cells.* 2019;8(9):1034. doi: 10.3390/cells8091034.
33. Voong LN, Xi L, Wang JP, Wang X. Genome-wide Mapping of the Nucleosome Landscape by Micrococcal Nuclease and Chemical Mapping. *Trends Genet.* 2017;33(8):495–507. doi: 10.1016/j.tig.2017.05.007.



## Efficiency of SpCas9 and AsCpf1 (Cas12a) programmable nucleases at genomic safe harbor loci in HEK293 cells

S.V. Pavlova<sup>1</sup> • E.A. Elisaphenko<sup>1</sup> • L.Sh. Shayakhmetova<sup>1, 2</sup> • S.P. Medvedev<sup>1</sup>

**Rationale:** The development of eukaryote genome engineering tools based on CRISPR-Cas programmable bacterial nucleases systems opens wide horizons for gene therapies, human disease cell modeling, as well as investigation into manifestation of disease phenotypes and visualization of cellular processes. The safety and approximation of experiments both at the cellular and organismal levels depend on the accuracy of introducing double-stranded breaks into the target DNA regions. The search for new variants of more accurate CRISPR-Cas nucleases and evaluation of their ability to hydrolyze nucleosome DNA *in vivo* is considered a critical task for the development of the genome engineering technologies.

**Aim:** To analyze the activity of the programmable nuclease AsCpf1 (Cas12a), with low level of off-target activity, in the human genome loci that are safe for the introduction of transgenic constructs (“safe harbor”) and to compare its efficiency with that of the widely used SpCas9 nuclease in HEK293 cells.

**Materials and methods:** We performed the bioinformatics analysis of the association between target regions with nucleosomes and other proteins in the safe harbor loci AAVS1 and GSH-Ch1 and the transcriptionally inactive gene *MYBPC3* (cardiac myosin binding protein 3) based on ATAC-seq data for the HEK293FT cells obtained from the NCBI SRA database. Plasmids encoding SpCas9 and AsCpf1 nucleases and guide RNA to the target regions were constructed and transfected into the HEK293FT cells. Events in the target regions of the HEK293FT cell genome were studied in the sequenograms with the TIDE algorithm.

**Results:** The results of the ATAC-seq experiments for HEK293FT cells have shown that the AAVS1 locus can be referred as open chromatin with a low nucleosome density, while the GSH-Ch1 locus can be attributed to closed chromatin. In HEK293FT cells, the cardiac *MYBPC3* gene has intermediate chromatin density. Assessment of the efficiency of introducing breaks into the studied HEK293FT cell chromatin loci by nucleases has shown that SpCas9 is able to cope with chromatin of any nucleosome density, while AsCpf1 can effectively introduce DNA breaks only at loci with open chromatin, such as AAVS1 and *MYBPC3*. Editing events occur at a very low rate at the GSH-Ch1 locus with a high nucleosome density.

**Conclusion:** We have found low efficiency of the AsCpf1 nuclease in the genomic safe harbor locus GSH-Ch1, which is characterized by a high nucleosome density. When planning an experiment on AsCpf1 nuclease genome editing, the epigenetic chromatin landscape and the nucleosome density should be considered, as well as chromatin opening substances should be used.

**Key words:** gene editing, CRISPR-Cas nuclease SpCas9 and AsCpf1, guide RNA, nucleosomes, open chromatin

**For citation:** Pavlova SV, Elisaphenko EA, Shayakhmetova LSh, Medvedev SP. Efficiency of SpCas9 and AsCpf1 (Cas12a) programmable nucleases at genomic safe harbor loci in HEK293 cells. Almanac of Clinical Medicine. 2021;49(6):385–395. doi: 10.18786/2072-0505-2021-49-037.

Received 29 June 2021; accepted 28 July 2021; published online 15 September 2021

**Sophia V. Pavlova** – PhD (in Biol.), Research Fellow, Laboratory of Developmental Epigenetics<sup>1</sup>; ORCID: <https://orcid.org/0000-0003-1095-3692>  
✉ 1–16 Mal'tseva ul., Novosibirsk, 630090, Russian Federation. Tel.: +7 (913) 903 13 62.  
E-mail: sonpavlova@gmail.com

**Evgeny A. Elisaphenko** – PhD (in Biol.), Senior Research Fellow, Laboratory of Developmental Epigenetics<sup>1</sup>; ORCID: <https://orcid.org/0000-0002-3204-8178>. E-mail: antares@bionet.nsc.ru

**Lilia Sh. Shayakhmetova** – Laboratory Assistant<sup>1</sup>; Bachelor<sup>2</sup>

**Sergey P. Medvedev** – PhD (in Biol.), Leading Research Fellow, Laboratory of Developmental Epigenetics<sup>1</sup>; ORCID: <https://orcid.org/0000-0002-1520-5549>

### Conflict of interests

The authors declare that there is no conflict of interests regarding the publication of this article.

### Authors' contributions

S.V. Pavlova, the study design and its experimental part, interpretation of the results, text writing; E.A. Elisaphenko, bioinformatics data analysis, analysis of the study results, editing of the manuscript; L.Sh. Shayakhmetova, conduction of the experiments, editing of the manuscript; S.P. Medvedev, the study concept and design, interpretation of the results, text writing, approval of the final version of the manuscript. All the authors have read and approved the final version of the manuscript before submission, agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work have been appropriately investigated and resolved.

### Acknowledgements

The study was supported by the Russian Foundation for Fundamental Research and by the Novosibirsk Oblast, project # 19-44-540002\_p\_a., budget project 0259-2021-0011.

<sup>1</sup>Institute of Cytology and Genetics, Siberian Branch of Russian Academy of Sciences; 10 Akademika Lavrent'yeva prospekt, Novosibirsk, 630090, Russian Federation

<sup>2</sup>Novosibirsk State University; 1 Pirogova ul., Novosibirsk, 630090, Russian Federation