

doi 10.18699/vjgb-26-04

Редактирование оснований в гене *AUTS2* и высокопроизводительное NGS-генотипирование клонов: стратегия создания клеточной модели

А.П. Ян ^{1, 2, 3} , П.А. Сальников ^{1, 2}, А.А. Буздин^{4, 5, 6}, В.А. Ковальская ⁷, Е.В. Мусатова⁸, П.С. Орлов ^{1, 9}, О.П. Рыжкова⁷, А.И. Субботовская ⁹, М.В. Сунцова^{4, 5}, А.Ю. Христинченко ⁵, А.А. Хабарова ¹ 

¹ Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия

² Новосибирский национальный исследовательский государственный университет, Новосибирск, Россия

³ Научно-технологический университет «Сириус», федеральная территория «Сириус», Краснодарский край, Россия

⁴ Первый Московский государственный медицинский университет им. И.М. Сеченова Министерства здравоохранения Российской Федерации, Москва, Россия

⁵ Национальный медицинский исследовательский центр эндокринологии им. академика И.И. Дедова, Москва, Россия

⁶ Институт биоорганической химии им. академиков М.М. Шемякина и Ю.А. Овчинникова Российской академии наук, Москва, Россия

⁷ Медико-генетический научный центр им. академика Н.П. Бочкова, Москва, Россия

⁸ Центр генетики и репродуктивной медицины «Genetico», Москва, Россия

⁹ Федеральный исследовательский центр фундаментальной и трансляционной медицины, Новосибирск, Россия

 a.yan@alumni.nsu.ru, khabarova@bionet.nsc.ru

Аннотация. Изучение молекулярных механизмов, лежащих в основе расстройств аутистического спектра (РАС), требует создания клеточных моделей, способных отражать цис-регуляторные эффекты и аллель-специфичную экспрессию генов. В настоящем исследовании мы представляем подход к получению индуцированных плюрипотентных стволовых клеток (ИПСК), модифицированных с использованием аденинового редактора оснований (ABE), для введения синонимичных однонуклеотидных замен в ген *AUTS2* – кандидата на участие в патогенезе РАС. Эти замены позволяют маркировать аллели и отслеживать различия в экспрессии нормального и реорганизованного аллелей в цис-контексте. Мы разработали стратегию высокоэффективного генотипирования клонов с использованием секвенирования продуктов ПЦР (ампликонов) на платформе нового поколения (NGS). Анализ более 100 субклонов показал, что предложенный подход превосходит секвенирование по Сэнгеру по масштабируемости, чувствительности и экономичности. Мы отобрали клоны с целевыми гетерозиготными заменами, оценили уровень мозаицизма и провели фазирование с герминальными гетерозиготными вариантами, позволяющее убедиться в моноклональном происхождении клеточной линии или идентифицировать аллель, ассоциированный с мутацией. Полученные линии ИПСК маркируют разные аллели гена *AUTS2*, что открывает перспективу анализа влияния цис-регуляторных элементов на экспрессию гена в различных типах клеток. Результаты работы подчеркивают практическую значимость редакторов оснований и целевого NGS-генотипирования при создании клеточных моделей с однонуклеотидными заменами для фундаментальных и прикладных исследований.

Ключевые слова: индуцированные плюрипотентные стволовые клетки (ИПСК); CRISPR/Cas9; однонуклеотидные замены; аллель-специфичная оценка экспрессии; хромосомная перестройка

Для цитирования: Ян А.П., Сальников П.А., Буздин А.А., Ковальская В.А., Мусатова Е.В., Орлов П.С., Рыжкова О.П., Субботовская А.И., Сунцова М.В., Христинченко А.Ю., Хабарова А.А. Редактирование оснований в гене *AUTS2* и высокопроизводительное NGS-генотипирование клонов: стратегия создания клеточной модели. *Вавиловский журнал генетики и селекции*. 2026;30(1):72-84. doi 10.18699/vjgb-26-04

Финансирование. Работа поддержана грантом РФФ 24-25-00152.

Благодарности. Секвенирование библиотек ДНК поддержано Министерством науки и высшего образования Российской Федерации (Соглашение № 075-15-2022-310 от 20 апреля 2022 г.). Культивирование клеток выполнено на базе ЦКП «Коллекция плюрипотентных культур клеток человека и млекопитающих общебиологического и биомедицинского направления» ИЦиГ СО РАН. Биоинформационный анализ данных проведен при финансовой поддержке бюджетного проекта (№ FWNR-2025-0017). Работа А.А. Буздина, М.В. Сунцовой и А.Ю. Христинченко поддержана проектом МЗ РФ «Разработка системы жидкостной биопсии для мониторинга ответа на терапию и прогрессии рака легкого».

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

Индивидуальный вклад авторов. А.А. Хабарова разработала концепцию исследования и руководила проектом. А.А. Хабарова, А.П. Ян, П.А. Сальников, А.А. Буздин, В.А. Ковальская, О.П. Рыжкова, М.В. Сунцова, А.Ю. Христинченко, Е.В. Мусатова, П.С. Орлов, А.И. Субботовская провели эксперименты. П.А. Сальников выполнил биоинформационную обработку результатов секвенирования. А.А. Хабарова, П.А. Сальников, А.П. Ян проанализировали результаты и подготовили текст рукописи. Все авторы редактировали рукопись и одобрили ее окончательный вариант.

Base editing in the *AUTS2* gene and high-throughput NGS genotyping of clones: a strategy for generating a cellular model

A.P. Yan ^{1, 2, 3} , P.A. Salnikov ^{1, 2}, A.A. Buzdin^{4, 5, 6}, V.A. Kovalskaia ⁷, E.V. Musatova⁸, P.S. Orlov ^{1, 9}, O.P. Ryzhkova⁷, A.I. Subbotovskaia ⁹, M.V. Suntsova^{4, 5}, A.U. Khristichenko ⁵, A.A. Khabarova ¹ 

¹ Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia

² Novosibirsk State University, Novosibirsk, Russia

³ Sirius University of Science and Technology, Sirius Federal Territory, Krasnodar region, Russia

⁴ I.M. Sechenov First Moscow State Medical University of the Ministry of Healthcare of the Russian Federation, Moscow, Russia

⁵ National Medical Research Center for Endocrinology named after Academician I.I. Dedov, Moscow, Russia

⁶ M.M. Shemyakin–Yu.A. Ovchinnikov Institute of Bioorganic Chemistry of the Russian Academy of Sciences, Moscow, Russia

⁷ Research Centre for Medical Genetics, Moscow, Russia

⁸ Center of Genetics and Reproductive Medicine “Genetico”, Moscow, Russia

⁹ Federal Research Center of Fundamental and Translational Medicine, Novosibirsk, Russia

 a.yan@alumni.nsu.ru, khabarova@bionet.nsc.ru

Abstract. Studying the molecular mechanisms underlying autism spectrum disorders (ASD) requires cellular models capable of capturing *cis*-regulatory effects and allele-specific gene expression. In this study, we present an approach for generating induced pluripotent stem cells (iPSCs) modified using an adenine base editor (ABE) to introduce synonymous single-nucleotide substitutions in the *AUTS2* gene – a candidate involved in ASD pathogenesis. These substitutions serve as allele-specific markers, enabling the tracking of expression differences between normal and rearranged alleles in a *cis*-regulatory context. We developed a high-efficiency strategy for genotyping clones using amplicon-based next-generation sequencing (NGS). Analysis of over 100 subclones demonstrated that this approach surpasses Sanger sequencing in scalability, sensitivity, and cost-effectiveness. We selected clones with targeted heterozygous substitutions, assessed mosaicism levels, and performed phasing with germline heterozygous variants to confirm monoclonal origin and identify the allele carrying the chromosomal rearrangement. The resulting iPSC lines mark distinct *AUTS2* alleles, providing a foundation for analyzing the impact of *cis*-regulatory elements on gene expression across different cell types. Our findings highlight the practical value of base editors and targeted NGS genotyping in creating cellular models with single-nucleotide substitutions for both basic and applied research.

Key words: induced pluripotent stem cells (iPSCs); CRISPR/Cas9; single-nucleotide substitutions; allele-specific expression analysis; chromosomal rearrangement

For citation: Yan A.P., Salnikov P.A., Buzdin A.A., Kovalskaia V.A., Musatova E.V., Orlov P.S., Ryzhkova O.P., Subbotovskaia A.I., Suntsova M.V., Khristichenko A.U., Khabarova A.A. Base editing in the *AUTS2* gene and high-throughput NGS genotyping of clones: a strategy for generating a cellular model. *Vavilovskii Zhurnal Genetiki i Seleksii* = *Vavilov J Genet Breed.* 2026;30(1):72-84. doi 10.18699/vjgb-26-04

Введение

По данным Global Burden of Disease, в 2021 г. с расстройствами аутистического спектра жили примерно 61.8 млн человек по всему миру, а совокупные потери лет жизни, скорректированных по нетрудоспособности, составили 11.5 млн – т.е. 147 лет жизни на каждые 100 000 человек (Global Burden of Disease Study 2021..., 2025). Выяснение генетических причин аутизма так же, как и других врожденных неврологических заболеваний, зачастую требует исследования механизмов патологических процессов, происходящих в головном мозге. С учетом малой доступности этого органа для прямого получения клеточного материала особенно важно создание клеточных моделей, которые позволят отслеживать изменения в экспрессии интересующих генов. Незаменимым инструментом для этого являются индуцированные плюрипотентные стволовые клетки (ИПСК), которые получают из большого спектра клеточного материала пациента. Такие клетки можно направленно дифференцировать в различные клеточные типы, что открывает большую свободу моделирования процессов, происходящих в разных тканях организма (Rowe, Daley, 2019; De Masi et al., 2020).

Для реконструкции патогенных мутаций создания в клеточных моделях широко применяется система геномного редактирования CRISPR/Cas9. Однако, несмотря на ее распространенность, эта система имеет ограничения, включающие неспецифическую активность и неточность репарации двуцепочечных разрывов, что может приводить к нежелательным мутациям (Смирнов и др., 2016; Uddin et al., 2020). В настоящее время развивается ряд более точных методов, основанных на технологии CRISPR/Cas9, лишенных основных недостатков исходной технологии. Один из таких методов – редакторы оснований, позволяющие специфично заменять один нуклеотид на другой (A→G или C→T) без индукции двуцепочечных разрывов (Gaudelli et al., 2017).

Внесение модификаций в геном с одонуклеотидной точностью не только открывает возможность моделирования генетических заболеваний, вызванных точечными мутациями (Lu, Huang, 2018; Geurts et al., 2023), но и позволяет маркировать аллели, что дает уникальную возможность проследить влияние *cis*-регуляторных вариантов на экспрессию и функциональную активность генов на моноаллельном уровне. Сравнение уровней экспрессии

аллелей повторяет дизайн гибридологического подхода, позволяя оценить цис-регуляторные эффекты при нивелировании любых транс-влияний (Salnikov et al., 2025). Кроме того, редакторы оснований активно используются для разработки генно-терапевтических стратегий для лечения заболеваний, связанных с вариантами в ядерной и митохондриальной ДНК (Billon et al., 2017; Liang et al., 2023).

В системе редактирования оснований используется гибридный белок, состоящий из лишённого нуклеазной активности белка Cas9 и фермента нуклеотид-дезаминазы. Такой химерный белок позволяет вносить замены А→G или С→Т без создания двуцепочечных разрывов ДНК. В случае аденинового редактора оснований (Adenin base editor, ABE) дезаминирование происходит на одной цепи ДНК в узком окне активности дезаминазы (4 нуклеотида), а одноцепочечный разрыв на другой цепи стимулирует систему репарации восстанавливать последовательность по модифицированной цепи (Rees, Liu, 2018). Система ABE индуцирует конверсию А→G путем дезаминирования аденина с превращением его в гипоксантин, который затем распознается ДНК-полимеразой как гуанин (Gaudelli et al., 2017; Chen et al., 2023).

Таким образом, редакторы оснований катализируют прямое химическое преобразование нуклеотидов без создания двуцепочечных разрывов, что минимизирует риск хромосомных нарушений и исключает необходимость в доставке экзогенной ДНК как матрицы для гомологичной репарации. Это существенно снижает частоту нецелевых инсерций, делеций и хромосомных перестроек по сравнению с классическим Cas9.

Несмотря на ряд очевидных преимуществ данной системы редактирования, редакторы оснований в своем применении имеют ряд ограничений. Ключевой недостаток – каталитическая специфичность дезаминазных ферментов, позволяющая осуществлять только транзиции (замены С→Т или А→G (замена пурина на пурин или пиримидина на пиримидин), но не трансверсии (замена пурина на пиримидин или наоборот), что значительно сужает спектр корректируемых мутаций (Komor et al., 2016; Gaudelli et al., 2017; Rees, Liu, 2018). Другую серьезную проблему представляет нецелевая активность. Она проявляется как в редактировании не полностью комплементарных сайтов в ДНК в пределах «окна редактирования», так и, что особенно важно, в неспецифической модификации клеточной РНК. Дезаминазы, особенно цитозиновые, могут случайно дезаминировать цитозины в РНК, потенциально нарушая транскриптом и вызывая токсические эффекты, причем эта активность часто не зависит от направляющей РНК (Grünwald et al., 2019; Jin et al., 2019; Zuo et al., 2019; Yu et al., 2020).

Настоящая статья посвящена использованию системы редактирования оснований ABE и инновационной стратегии генотипирования клеточных клонов при создании клеточных моделей на основе ИПСК. При анализе кариотипа пациента с расстройством аутистического спектра была обнаружена хромосомная перестройка на одном из аллелей, граница которой находится в межгенной области

вблизи гена *AUTS2* (Gridina et al., 2025). Мы полагаем, что эта перестройка может нарушать регуляцию экспрессии гена *AUTS2*, что, в свою очередь, может быть связано с наблюдаемым фенотипом пациента. Для экспериментальной проверки этой гипотезы необходимо создать модель, позволяющую оценивать различия в экспрессии нормального аллеля и находящегося в цис-положении с хромосомной перестройкой в разных типах клеток. Однако для этого пациента не найдено однонуклеотидных замен в кодирующей области гена *AUTS2*, которые бы позволили дискриминировать транскрипты с разных аллелей. Из ранее полученных ИПСК данного пациента с использованием системы ABE мы создали модельные линии, несущие гетерозиготные однонуклеотидные замены в 10-м экзоне гена *AUTS2*. Эти линии дают возможность получать нейроны/нейральные предшественники *in vitro* и количественно оценивать аллель-специфичную транскрипцию *AUTS2* с перестроенной и интактной хромосомами.

Генотипирование клонов и оценку эффективности редактирования рутинно оценивают секвенированием по Сэнгеру ампликонов целевого локуса. Наиболее удобный и точный подход – секвенирование ампликона, содержащего целевую замену. Для секвенирования ампликонов применимы два подхода: метод Сэнгера и секвенирование нового поколения (NGS) (рис. 1, в, з). При секвенировании ампликонов выбор между методом Сэнгера и NGS определяется конкретными задачами исследования. Метод Сэнгера остается оптимальным для анализа отдельных продуктов ПЦР при небольшом количестве образцов. Его ограниченная пропускная способность делает его экономически нецелесообразным при работе с большим количеством образцов. В отличие от этого, NGS-технологии позволяют проводить одновременное секвенирование множества образцов, что значительно снижает стоимость анализа на один образец и делает метод незаменимым для массового анализа. Однако следует учитывать, что NGS требует биоинформатического анализа. Таким образом, для рутинного анализа единичных ампликонов метод Сэнгера сохраняет свои преимущества, тогда как при необходимости масштабировать анализ для большого количества образцов NGS демонстрирует явное превосходство.

В рамках этого исследования с помощью NGS проанализировано 117 субклонов и отобраны линии, подходящие для дифференцировки и анализа экспрессии гена *AUTS2*. Мы подробно описываем протокол пробоподготовки для NGS-генотипирования и биоинформационного анализа полученных данных, который может быть адаптирован для других задач генотипирования.

Материалы и методы

Создание вектора MLM3636-BstV2I. В качестве основы для вектора мы взяли плазмиду MLM3636 (Addgene #43860), содержащую последовательность направляющей РНК (нРНК) с сайтом клонирования спейсера, рассчитанного на использование фермента рестрикции BsmBI (New England Biolabs, США). Были задействованы следующие последовательности протоспейсеров: 5'-CAA AAGTTGACCCATTCTAC-3' для H044 и 5'-TCACACT

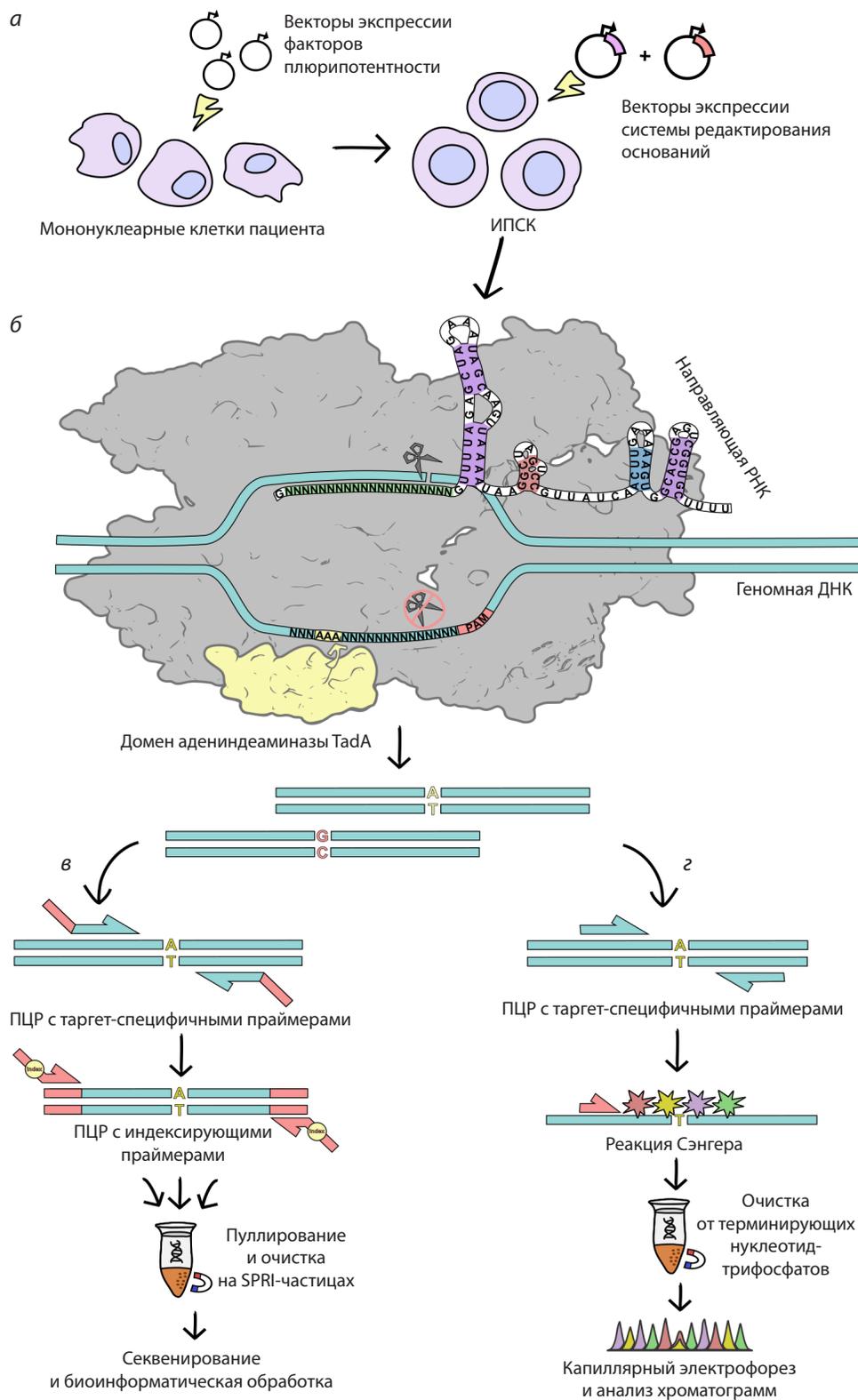


Рис. 1. Общая схема эксперимента по получению однонуклеотидных замен в ИПСК человека и сравнение этапов протоколов генотипирования клонов методами NGS и секвенирования по Сэнгеру.

а – электропорация моноцитарных клеток крови пациента векторами экспрессии факторов репрограммирования, получение ИПСК и затем электропорация ИПСК плазмидами, кодирующими систему редактирования оснований: целевые направляющие РНК и модифицированный белок Cas9, слитый с адениндеаминазой TadA; *б* – схематичное взаимодействие комплекса модифицированного белка Cas9 (серый цвет), слитого с доменом адениндеаминазы TadA (светло-желтый), и направляющей РНК с последовательностью протоспейсера в геномной ДНК. Аденины в окне редактирования выделены желтым; *в* – схема приготовления NGS-библиотеки для генотипирования клонов путем добавления адаптерных последовательностей и индексирования с помощью двух раундов ПЦР; *з* – схема стандартного протокола секвенирования по Сэнгеру.

GTGCCGGTAGAAT-3' для H068. Данный фермент не производится на территории РФ, поэтому для упрощения дальнейших манипуляций мы заменили оригинальный сайт клонирования на аналогичный сайт из плазмиды pSpCas9(BB)-2A-GFP (PX458) (Addgene #48138), распознаваемый ферментом BstV2I, выпускаемым российским производителем ООО «СибЭнзим».

Для этого мы амплифицировали фрагмент, содержащий сайт клонирования pSpCas9(BB)-2A-GFP (PX458) с праймеров pSpCas9_clonF и pSpCas9_clonR, и почти всю последовательность вектора MLM3636 с применением праймеров MLM3636_clonF и MLM3636_clonR. Полученный продукт обработали 1 ед. фермента DpnI (New England Biolabs, США) и очистили с помощью суспензии SPRI-частиц VAHTS DNA Clean Beads (Vazyme, Китай). Сборку конструкции проводили с помощью NEBuilder HiFi (New England Biolabs) в соответствии с протоколом производителя.

Для получения векторов, экспрессирующих нРНК, 1 мкг плазмиды MLM3636-BstV2I линейризовали 25 ед. фермента BstV2I в 50 мкл реакционной смеси при 55 °С в течение ночи. После рестрикции ДНК очищали 0.8-кратным объемом суспензии SPRI-частиц VAHTS DNA Clean Beads (Vazyme). Одноцепочечные олигонуклеотиды H044_F и H044_R (по 2 мкМ каждого) фосфорилировали в 1× буфере для T4-ДНК-лигазы с 100 нМ АТФ и 10 ед. T4-полинуклеотидкиназы («СибЭнзим»). Для пары H068 применяли аналогичную процедуру. Затем олигонуклеотиды нагревали до 95 °С и гибридизовали медленным охлаждением со скоростью 0.1 °С/с до 4 °С.

Полученный раствор дц-олигонуклеотида H044 (H068) F+R на следующем этапе лигировали в линейризованную векторную конструкцию MLM3636-BstV2I. Для этого смешивали на льду следующие компоненты: 100 нМ раствора дц-олигонуклеотида до 20 нг линейризованной ранее плазмиды, 100 нМ АТФ, 1× SE-буфер T4 ДНК лигазы, 20 ед. T4 ДНК лигазы. Реакционную смесь инкубировали в течение ночи при 4 °С. Очищали продукт 0.8 объемом суспензии SPRI-частиц VAHTS DNA Clean Beads (Vazyme), элюировали в 5 мкл H₂O. Для последующей трансформации бактериальных клеток использовали половину полученного объема продукта.

Трансформацию *Escherichia coli* проводили с помощью метода электропорации, плазмидную ДНК выделяли набором Midi-prep («Евроген», Россия) и дополнительно очищали 0.8 объемом суспензии VAHTS DNA Clean Beads (Vazyme). Элюцию проводили в минимальном объеме воды, чтобы достигнуть концентрации ДНК не менее 1 мкг/мкл. Для селекции клеток с применением пурамицина использовали плазмиду pURC_puro (предоставлена сотрудником отдела молекулярных механизмов онтогенеза ИЦиГ СО РАН А. Нурисламовым). Данная конструкция экспрессирует встроенный в плазмиду под промотором CMV ген устойчивости к пурамицину, что позволяет отобрать такие клетки и обогатить популяцию клеток событиями редактирования.

Культивирование и электропорация клеток с помощью системы Neon. ИПСК культивировали на пласти-

ковых планшетах, дно которых предварительно покрывали раствором культурального матрикса Matrigel (Corning, США), с использованием среды mTesr1 (StemCell, Канада) с добавлением смеси антибиотиков пенициллина и стрептомицина («ПанЭко», Россия). Для дезагрегации клеток применяли реагент TripLE (Thermo Fisher, США), рассаживали в соотношении 1:3–1:5 с добавлением ингибитора Y-27632 (10 мкМ, ингибитор Rho-киназы) (StemCell, Канада). За день до электропорации линии ИПСК пересаживали так, чтобы клеточная культура покрывала дно планшета не больше чем на 70 % от площади лунки. Перед переносом трансформированных клеток дно культуральных планшетов покрывали матриксом Matrigel (1 мл на 10 см²) и инкубировали в термостате при 37 °С в течение 20–30 мин. За 1 ч до трансформации к клеткам добавляли 10 мкМ Y-27632.

Клеточную культуру дезагрегировали с помощью 700 мкл TripLE (Thermo Fisher). Затем добавляли в каждую лунку по 1 мл промывочной среды, мягко ресуспендировали клетки и собирали суспензию в центрифужную пробирку. Подсчет клеток осуществляли в камере Горяева. Суспензию центрифугировали 5 мин при 300 g. После подсчета необходимое количество клеточной суспензии (150–200 тыс. клеток на одну реакцию электропорации) отмывали один раз в фосфатном буфере. В пробирке на 1500 мкл смешивали три плазмиды (каждой по 0.5 мкг): плазмиды экспрессии нРНК (MLM3636-BstV2I), редактора оснований (addgene #108382) и pURC_puro (устойчивость к пурамицину) и добавляли 150–250 тыс. клеток, ресуспендированных после отмывания в фосфатном буфере в буфере R из набора для электропорации Neon. Суммарный объем суспензии клеток и смеси плазмид не должен превышать 10 мкл.

Электропорацию проводили с помощью системы Neon при условиях 1100 V – 30 мс – 1 импульс и переносили клетки в лунку культурального планшета (6 лунок) в 3 мл среды mTesr1, не содержащей антибиотиков. Через сутки в среду добавляли антибиотики (1× смесь пенициллина и стрептомицина) и также добавляли селективный антибиотик пурамицин в концентрации 1 мкг/мл на 48 ч (Gibco, США). Среду меняли полностью через 1 день. Через 7–10 дней колонии клеток индивидуально переносили в отдельные лунки 24-луночного планшета (снимали около 30 колоний). Как только колонии достигали размеров 1/4 от площади лунки, их дезагрегировали с помощью TripLE, затем после центрифугирования часть клеточной суспензии оставляли для выделения ДНК и последующего генотипирования, а часть суспензии замораживали в KSR (Gibco) с добавлением 10 % диметилсульфоксида.

Генотипирование клонов с помощью NGS-секвенирования ампликонов. Для постановки ПЦР использовали набор реактивов от компании «МБС-Технология» (Россия). Общий объем 50 мкл содержал: 50 нг ДНК, 1× БР MBUision, 0.2 мМ каждого dNTP; 0.2 мкМ праймера AUTS_gen_F; 0.2 мкМ праймера AUTS_gen_R (табл. 1); 0.4 мкл MBUision полимеразы («МБС-Технология»); H₂O до 50 мкл. ПЦР проводили при следующих условиях: 95 °С 3 мин, 20 циклов: 95 °С 15 с, 60 °С 30 с, 72 °С 30 с, и затем

Таблица 1. Праймеры и олигонуклеотиды

Название	Последовательность 5'–3'
MLM_genF	TCCGGCAGGAAGAGGGCCTATTTTC
MLM_genR	CCTCGAGCGGCCCAAGCTTAAAAA
H044_F	CACCGCAAAAGTTGACCCATTCTAC
H044_R	AAACGTAGAATGGGTCAACTTTTGC
H068_F	CACCGTCACACTGTGCCGGTAGAAT
H068_R	AAACATTCTACCGGCACAGTGTGAC
AUTS_gen_F	CTGGAGTTCAGACGTGTGCTCTCCGATCTTGACAGCTTAATGACAGGGAAGC
AUTS_gen_R	TCTTTCCCTACACGACGCTCTCCGATCTCTGATCTGTGGCAAGAGCACAA
P5_IP{index_ID}	AATGATACGGCGACCACCGAGATCTACAC{8bp_Index}ACACTCTTTCCCTACACGAC
P7_IP{index_ID}	CAAGCAGAAGACGGCATACGAGAT{8bp_Index}GTGACTGGAGTTCAGACGTGT
pSpCas9_clonF	GGCCTATTTCCCATGATTCCT
pSpCas9_clonR	CGACTCGGTGCCACTTTT
MLM3636_clonF	AGGAATCATGGGAAATAGGCC
MLM3636_clonR	AAAAGTGGCACCGAGTCG
AUTS2_phase_F1	CTGGAGTTCAGACGTGTGCTCTCCGATCTACTAGCTTTTGCTTTGATCCC
AUTS2_phase_R1	TCTTTCCCTACACGACGCTCTCCGATCTGCCTGACTGCCACTAAAGAG
AUTS2_phase_F2	CTGGAGTTCAGACGTGTGCTCTCCGATCTCCACTGTAGCAGTGAACAAA
AUTS2_phase_R2	TCTTTCCCTACACGACGCTCTCCGATCTATGCCTGACTGCCACTAAAG

72 °C 1 мин, 4 °C. Полученные продукты ПЦР разводили в 100 раз и использовали 2 мкл данного раствора в качестве матрицы для последующей ПЦР, которую проводили с индексированием и включением полных последовательностей, необходимых для секвенирования, P5_IP {index_ID} и P7_IP {index_ID} (см. табл. 1).

Аmplification осуществляли на данных условиях: 95 °C 3 мин; 8 циклов: 95 °C 15 с; 60 °C 30 с; 72 °C 30 с и затем 72 °C 1 мин, 4 °C. Приблизительную концентрацию продуктов ПЦР оценивали с помощью их визуализации на электрофореze в агарозном геле, затем все образцы смешивали в примерно равных количествах и очищали одним объемом суспензии VAHTS DNA Clean Beads (Vazyme), затем элюировали в 20 мкл H₂O.

Секвенирование выполняли в режиме парных прочтений по 150 п. н. (в среднем ~10000 прочтений на образец). Из полученных данных убрали последовательности адаптеров (cutadapt), после чего выровняли на индексированную референс-последовательность целевого региона инструментом bowtie2, параметры по умолчанию. Полученные данные в формате BAM визуализировали с помощью Integrative Genomics Viewer (IGV, США).

Для идентификации вариантов формировали VCF-файлы командой bcftools mpileup с записью глубины аллелей (флаг AD). Далее собственным скриптом на Python рассчитывали частоту представленности мутантного аллеля по каждой позиции, агрегировали результаты в единую матрицу и строили линейные графики для групп образцов, что позволяло оценить распределение аллельных частот для замен по всем возможным позициям. Все использо-

ванные скрипты с примерами доступны на GitHub <https://github.com/Somatich/NGS-genotyping-Yan-et-al.-2025>.

Фазирование внесенных замен с герминальным вариантом. Под фазированием однонуклеотидных замен мы подразумеваем определение того, находятся ли два варианта разных замен в цис-конфигурации или в транс-конфигурации. Мы подобрали пару праймеров, фланкирующих как сайт индуцированного мутагенеза, так и герминальный вариант (chr7:70768282 hg38, rs3829006 G/A), с ожидаемой длиной продукта ПЦР длиной 363 п. н.; к 5'-концам праймеров были добавлены технические последовательности для NGS. Данный герминальный вариант находился на расстоянии 238 п. н. от сайта индуцируемого мутагенеза и был выбран при анализе данных секвенирования экзона пациента (Gridina et al., 2025). Удаление на расстояние не более 200–300 нуклеотидов позволяет получить продукт ПЦР, длина которого оптимальна для секвенирования с помощью системы NGS Illumina (секвенирование образца с длиной фрагмента ДНК более 500 п. н. значительно снижало бы его эффективность). Были использованы две разные пары праймеров: AUTS2_phase_F1/R1 и AUTS2_phase_F2/R2 (см. табл. 1).

Для постановки ПЦР задействовали набор реактивов от компании «МБС-Технология». Индексирование ПЦР и очистку проводили так же, как указано выше. Библиотеки секвенировали в режиме парных прочтений по 150 п. н., получив около 5000 пар прочтений на образец. Чтения картировали утилитой Bowtie2 с настройками по умолчанию. Фазирование SNP выполняли вручную, анализируя выравнивания в геномном браузере IGV.

Результаты

Дизайн нРНК и оценка эффективности редактирования в 10-м экзоне гена *AUTS2*

Целью данного исследования является создание клеточной модели на основе ИПСК, содержащих гетерозиготную однонуклеотидную синонимичную замену в 10-м экзоне гена *AUTS2*, с применением аденинового редактирования оснований (ABE) (см. рис. 1, а). Выбор 10-го экзона был обусловлен тем, что он присутствует в большинстве описанных и предсказанных транскриптов. Кроме этого, именно рядом с 10-м экзоном обнаружен герминальный однонуклеотидный вариант, позволяющий нам фазировать внесенную замену относительно аллелей.

На этапе дизайна нРНК была проанализирована нуклеотидная последовательность 10-го экзона гена *AUTS2*. В результате этого анализа найдены позиции аденина, расположенные на третьем месте в составе кодонов и попадающие в рабочее окно ABE (нуклеотиды 4–7 протоспейсера). Таким образом, замена таких аденинов на гуанин является синонимичной заменой. В 10-м экзоне идентифицировано два таких аденина: chr7:70768044 A→G и chr7:70768068 T→C (hg38). Замена аденина на позиции chr7:70768044 приводит к синонимичной замене аденина в триплете AAA на AAG (p.Lys570Lys). Данный вариант встречается в популяции с вероятностью 0.000005596, не предсказан как влияющий на сплайсинг или создающий новый сайт сплайсинга (при анализе последовательности в базах данных SpliceAI и MobiDetails). Замена на позиции chr7:70768068 T→C – синонимичная замена AGT на AGC, данный вариант не описан в базе данных gnomAD и также не предсказан как вариант, влияющий на сплайсинг или создающий новый сайт сплайсинга (SpliceAI и MobiDetails) (Reese et al., 1997; Jaganathan et al., 2019; Vaux et al., 2021).

Для обоих этих вариантов сконструированы две направляющие РНК: Н044 и Н068 (см. табл. 1). Оценка эффективности направляющих РНК в условиях *in vivo* проведена на культуре ИПСК с использованием плазмиды, экспрессирующей белок Cas9 (addgene 62988), и тоже показала примерно одинаковую эффективность обеих нРНК: для позиции Н044 (chr7:70768044) составила 9.2 %, для Н068 (chr7:70768068) – 7.6 %.

Анализ выполняли с помощью инструмента Cas-Analyser (Crisper Rgen) (Hwang et al., 2018). Важно подчеркнуть, что эффективность геномного редактирования определяется множеством других трудно контролируемых факторов, таких как локальное состояние хроматина конкретной линии, поэтому не следует полностью полагаться на расчетное предсказание активности и специфичности направляющей РНК, но, чтобы повысить вероятность получения требуемой мутации, целесообразно параллельно использовать несколько нРНК с разными протоспейсерами. Такой подход повышает вероятность получения требуемой мутации. Поэтому в дальнейших экспериментах мы применяли обе нРНК.

В качестве пробного эксперимента на популяции клеток (после трансфекции для анализа были взяты все клет-

ки, без отбора тех, в которые действительно попали конструкции), была оценена эффективность обеих направляющих РНК в комбинации с редактором оснований Cas9(ABE7.10). После электропорации клеток генетическими конструкциями, экспрессирующими белок Cas9 и направляющую РНК, на третий день из клеток выделяли ДНК и проводили пробоподготовку образцов для NGS. Анализ осуществляли с помощью инструмента BE-Analyser (Crisper Rgen, Южная Корея) (Hwang et al., 2018). Эффективность внесения замены аденина с помощью редактора оснований для Н044 (chr7:70768044) составила 5.1 %.

На соседних позициях (chr7:70768041–70768043) с левым аденином (chr7:70768044) находятся четыре аденина. Мы оценили вероятность редактирования редактором оснований: так, для аденина в позиции chr7:70768043 она составила 3 %, частота внесения замены аденина еще в двух соседних позициях (chr7:70768041–70768042) была 0.4 и 0.7 % соответственно. Для Н068 в позиции chr7:70768068 эффективность составила 3 %. Полученные данные свидетельствуют о том, что, несмотря на наличие кластера из четырех аденинов (chr7:70768041–70768044) вблизи целевой позиции, протоспейсер Н044 обеспечивает более высокую эффективность редактирования в заданной позиции по сравнению с Н068. Достаточно высокая частота нецелевой замены аденина в позиции chr7:70768043 в нашем случае не являлась критическим фактором для использования данной нРНК, но вынуждает задействовать большую выборку модифицированных клонов для поиска нужного нам генотипа с заменой в необходимой позиции.

Также для повышения доли клеток с необходимой генетической модификацией в популяции мы решили применить методы обогащения, такие как котрансфекция клеток с плазмидой, несущей селективный маркер (например, устойчивости к пурамицину), что позволит эффективно отобрать клетки, успешно трансфицированные генетическими конструкциями, и значительно повысит выход модифицированных клеток. Таким образом, результаты пилотного эксперимента подтверждают работоспособность системы редакторов оснований в локусе гена *AUTS2*; нРНК Н044 демонстрирует более высокую эффективность редактирования целевой позиции по сравнению с нРНК Н068, несмотря на потенциальные проблемы, связанные с кластером аденинов.

Получение и отбор клонов ИПСК

с внесенной заменой после редактирования оснований

Для внесения целевой замены в 10-м экзоне гена *AUTS2* применили электропорацию с использованием системы Neon™. Электропорацию двух исходных линий человеческих ИПСК проводили смесью трех плазмид, кодирующих направляющую РНК, адениновый редактор оснований и маркер устойчивости к пурамицину. После электропорации клетки рассаживали в низкой плотности на матриксе и культивировали на среде mTesr1, содержащей селективный антибиотик в течение трех суток. На 7–10-й день после трансфекции были визуально идентифицированы и механически отобраны индивидуальные колонии, де-

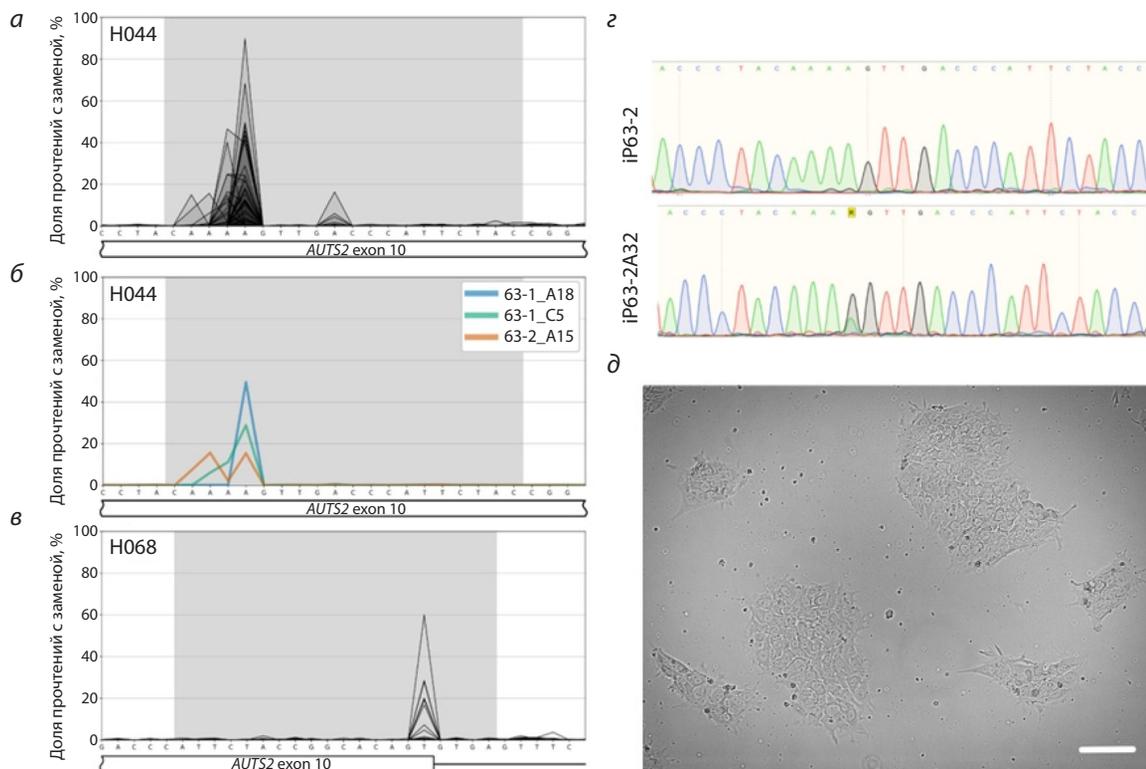


Рис. 2. Примеры результатов представленного протокола.

а – визуализация суммарных результатов генотипирования для 69 клеточных клонов, модифицированных с помощью направляющей РНК Н044; *б* – пример визуализации результатов генотипирования для трех избранных образцов. По оси абсцисс – геномная последовательность chr7:70768037–70768062, hg38 в нуклеотидах, по оси ординат – доля прочтений секвенирования, содержащих замену; *в* – визуализация суммарных результатов генотипирования для 48 клеточных клонов, модифицированных с помощью направляющей РНК Н068. По оси абсцисс – геномная последовательность chr7:70768048–70768077, hg38 в нуклеотидах. Серая область соответствует последовательности направляющей РНК, использованной в эксперименте; *з* – секвеннограммы образцов iP-63-2 (исходная линия ИПСК) и iP-63-2A32 (линия ИПСК с синонимичной заменой триплета AAA/AAG) для сайта мутагенеза; *д* – морфология клеточной культуры ИПСК в проходящем свете; белая линия – 100 мкм.

монстрирующие характерную для ИПСК морфологию (рис. 2, *д*). В результате проведенного отбора для последующей детальной характеристики было сохранено 117 клонов, которые показали стабильный рост и поддерживали типичный для плюрипотентных клеток морфологический фенотип.

На следующем этапе было необходимо решить экспериментальную задачу по генотипированию более чем 100 образцов. Для оценки общей эффективности и результатов геномного редактирования традиционно используют секвенирование по Сэнгеру ампликонов редактируемого сайта (см. рис. 1, *в*, *з*). Однако этот метод трудозатратен, а при анализе большого количества образцов (больше 80) становится нерентабельным (рис. 3). Мы предлагаем альтернативное решение – массовое генотипирование клеточных клонов с использованием секвенирования нового поколения. Этот подход обеспечивает высокую масштабируемость, существенно ускоряет процесс анализа, позволяя одновременно получить информацию как о составе аллелей, так и об их частотах в популяции клеток.

Для анализа NGS целевой фрагмент, содержащий внесенную замену, амплифицировали с помощью сайт-специфичных праймеров, фланкирующих сайт мутагенеза.

Ключевым требованием для успешного генотипирования (при использовании парных прочтений длиной 150 пар оснований) является расположение праймеров на расстоянии от сайта мутагенеза в пределах 50–130 пар оснований до 5'-концов праймеров (см. рис. 1, *в*). При таком дизайне эксперимента обеспечивается покрытие внесенной замены обоими прочтениями в рамках одной пары. Для интеграции в NGS, на 5'-концы сайт-специфичных праймеров необходимо добавить технические последовательности, представляющие собой 3'-регионы ДНК-адаптеров для секвенирования. Для завершения последовательностей адаптеров проводится второй раунд ПЦР с набором универсальных праймеров, содержащих индексные последовательности для дальнейшего пулирования и демультиплексирования библиотек.

Оценка эффективности редактирования и мозаичности колоний

После электропорации получили и проанализировали в общей сложности 117 клонов: 69 клонов с использованием нРНК Н044 (см. рис. 2, *а*) и 48 для Н068 (см. рис. 2, *в*). Общая эффективность редактирования для направляющей РНК Н044, рассчитанная как доля клонов с любыми об-

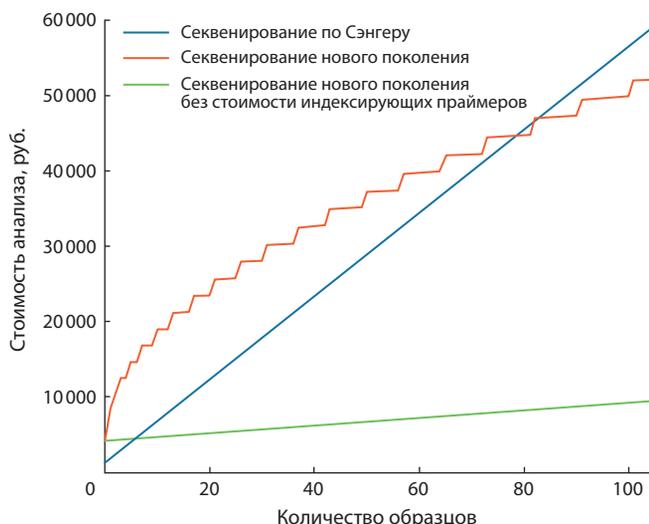


Рис. 3. Зависимость стоимости секвенирования методом секвенирования по Сэнгеру (синий цвет) и методом NGS с учетом (оранжевый) и без учета (зеленый) стоимости синтеза индексирующих праймеров.

наруженными заменами с частотой встречаемости аллеля более 2 %, составила 54 % (37/69). Поскольку для дальнейших исследований нам требовалось получить линии, несущие замены в гетерозиготном состоянии, мы провели фильтрацию и отобрали клоны с долей отредактированного аллеля ≥ 30 %. При таком пороге эффективность редактирования составила 13 % (9/69), при этом два клон демонстрировали частоты внесенных замен 68 и 89 %, что говорит об их потенциальной гомозиготности (доля отредактированного аллеля ≥ 30 %). Для направляющей РНК Н068 суммарная эффективность составила 31 % (15/48) при пороге 2 % и 4 % (2/48) при пороге 30 %.

Видно, что при теоретически возможных только гомо- и гетерозиготных состояниях локуса наблюдаемые частоты встречаемости альтернативного аллеля могут существенно отклоняться от ожидаемых 0, 50 и 100 %. Часть наблюдаемых расхождений, вероятно, связана с особенностями подготовки NGS-библиотек (многочисленные циклы амплификации и риск кросс-контаминации). Однако ключевую роль, по-видимому, играет вероятность кросс-контаминации культур на этапе субклонирования после электропорации. Исключить ее полностью затруднительно, поскольку клетки не всегда диссоциируют до одиночных, могут повторно агрегировать и мигрировать в ходе совместного культивирования. Дополнительную вариабельность вносит асинхронность процесса редактирования: между трансфекцией и началом экспрессии и работы компонентов системы редактирования оснований проходит время, за которое некоторые клетки успевают пройти цикл деления. В результате в разных клетках одной и той же колонии могут появиться различные мутации или не появиться вовсе, что формирует мозаичную по внесенным заменам культуру. Таким образом, независимо от исходной эффективности мутагенеза, необходимо

исследовать значительное количество клонов, чтобы исключить мозаичные колонии.

В окно редактирования (4–6-й нуклеотиды спейсера) попадает сразу несколько аденинов. Результаты свидетельствуют о том, что замены происходят по всем «валентным» позициям, но с различной эффективностью: наибольшей – по пятой позиции (chr7:70768044, около 45 % позитивных клонов при пороге в 2 % частоты аллеля) и вдвое реже по четвертой позиции (chr7:70768043, около 17 % при том же пороге) (см. рис. 2, а). Шестая позиция представлена гуанином и не поддается оценке эффективности замен, хотя и попадает в теоретическое окно редактирования. Частота редактирования по остальным аденинам региона не превышала 5 %. Примечательно, что случаев одновременных замен двух нуклеотидов в пределах одного и того же аллеля не зафиксировано. Также не зафиксировано случаев внесения замены на нецелевой цепи (конверсии T→C).

Для наглядности на рис. 2 приведен предлагаемый способ визуализации результатов (см. рис. 2, а–в). Представленный метод визуализации позволяет быстро сравнивать образцы между собой, оценивать уровень шума и отбирать позитивные образцы для получения модельных клеточных линий для дальнейшей работы.

Мы отобрали пять колоний с подходящими заменами в потенциально гетерозиготном и генетически гомогенном состоянии. Для оценки того, какой аллель маркирует внесенная замена, а также для оценки мозаичности, необходимо фазировать внесенные замены с герминальными гетерозиготными вариантами в геномном окружении. Так, если наблюдается всего два комплаунд-аллеля, то можно считать, что полученные гетерозиготные линии моноклональны. Мы фазировали полученные замены для клонов с однонуклеотидным вариантом (SNV chr7:70768282 hg38, rs3829006 G/A), существование которого было обнаружено при полногеномном секвенировании пациента, маркирующий аллель, находящийся в цис-окружении с хромосомной аберрацией. Было отобрано три немозаичных клонов (с частотой встречаемости мозаичных аллелей менее 5 %), два из которых несли синонимичную замену в гетерозиготном состоянии, при этом маркирующую разные аллели у разных клонов (табл. 2).

Таблица 2. Фазирование и оценка мозаичности модифицированных по гену *AUTS2* клонов

Клон	Позиция замены, hg38	Частота замен, %	Аллель	Мозаицизм
63-2_A32	chr7:70768044	41.5	SNV	–
63-1_C16	chr7:70768043	40.2	WT	–
63-1_C21	chr7:70768044	48.4	WT	–
63-2_A30	chr7:70768043	46.5	WT	+
63-2_A30	chr7:70768044	39.5	SNV	+
63-2_C19	chr7:70768064	28.60	SNV	+

Примечание. WT – дикий тип; SNV – вариант с заменой.

Таблица 3. Сравнение генотипирования с помощью секвенирования по Сэнгеру и с помощью NGS

Критерий	Секвенирование по Сэнгеру	NGS-генотипирование
Масштабируемость	Позволяет секвенировать по восемь образцов одновременно, а анализировать результаты по одному образцу	Позволяет секвенировать и анализировать десятки-сотни образцов одновременно
Трудоемкость пробоподготовки	Один раунд ПЦР, реакция секвенирования и очистка реакции секвенирования. Требуется индивидуальная пробоподготовка (ПЦР и очистка образцов)	Два раунда ПЦР с целевыми и индексированными праймерами индивидуально для каждого образца. После индексирования в результате второго раунда ПЦР их можно объединить в один и очистить с помощью магнитных SPRI-частиц
Скорость анализа результатов	Индивидуальный анализ хроматограмм. Требуется ручная интерпретация хроматограмм	Быстрый и автоматизированный анализ
Чувствительность	Анализирует усредненный сигнал; сложно интерпретировать при наличии нескольких аллелей. Порог обнаружения – около 15–20 %; низкочастотные мутации не выявляются	Позволяет различать и количественно оценивать частоту отдельных аллелей (аллель-специфичный анализ). До 0.1 % – позволяет обнаруживать низкочастотные варианты и мозаицизм
Качество секвенирования последовательности	Крайние участки последовательности (до 100 п. о.), близкие к местам посадки праймеров, секвенируются в низком качестве и часто не пригодны для анализа	Крайние участки секвенируемой последовательности хорошего качества
Требования к праймерам	Обычные	Длинные праймеры, чувствительны к ошибкам, требуют очистки на ПААГ
Выбор клонов	Не позволяет надежно различать мозаичные или субклональные события	Позволяет выбирать клоны с заданной частотой аллеля; выявляет субклоны и исключает мозаицизм
Применимость метода	Подходит для проверки отдельных мутаций с высокой частотой в образце. Низкую эффективность редактирования (менее ~20 %) не удастся обнаружить	Подходит для широкого круга задач, включая количественную оценку редактирования, отслеживание субпопуляций и др. Обнаруживает даже 5–9 % редактирования в популяции клеток
Стоимость	Около 700 рублей на образец	Около 100 рублей на образец

После анализа колоний на мозаицизм выбранные клеточные линии культивировались для дальнейшего анализа аллельной экспрессии гена *AUTS2*. Для избежания вероятных ошибок в ходе массового культивирования клеточных линий мы повторно проанализировали наличие гетерозиготных замен, результаты анализа представлены на секвенограмме (см. рис. 2, з). Таким образом, мы получили необходимые линии и инструмент для анализа влияния хромосомной перестройки пациента на экспрессию гена *AUTS2*.

Обсуждение

В данной работе мы получили три генетически модифицированные линии ИПСК пациента, несущие однонуклеотидные замены в экзоне 10 гена *AUTS2*, а также продемонстрировали эффективность и работоспособность стратегии отбора генетически модифицированных колоний с использованием генотипирования с помощью NGS секвенирования ампликонов. Эти линии характеризуются генетической гомогенностью, получены из двух разных линий ИПСК одного пациента, две из них несут синонимичные замены и маркируют разные аллели. Таким образом, они полностью подходят для дальнейшего изучения влияния хромосомной перестройки пациента на экспрессию гена *AUTS2* и ее дальнейшего функционального влияния.

Генетическая маркировка аллелей позволяет оценить различия в их цис-регуляции в одной и той же клетке. Такая экспериментальная схема минимизирует влияние

трудно контролируемых факторов на биологическую модель. Поскольку оба аллеля находятся под действием одного и того же набора транс-регуляторов, отражающих клеточное состояние и внешние условия, обнаруживаемые различия в экспрессии обусловлены исключительно их цис-окружением. По этой причине в анализ можно включать даже линии с несинонимичными заменами: если изменение аминокислотной последовательности влияет на транскрипцию через механизмы обратной связи, этот эффект одинаково влияет на оба аллеля и, таким образом, не искажает оценку различий в цис-регуляции.

Применение NGS для генотипирования значительно сокращает время лабораторных манипуляций и упрощает процесс анализа по сравнению с секвенированием первого поколения. В отличие от секвенирования отдельных образцов методом Сэнгера, NGS позволяет масштабировать процесс и анализировать более 100 образцов за один раз (табл. 3). Для приготовления NGS-библиотеки необходимы рутинное выделение ДНК и проведение двух раундов ПЦР с использованием праймеров, специфичных к целевой последовательности нужного района генома и вносящих необходимые технические последовательности для секвенирования.

Интеграция NGS с биоинформатическими инструментами обеспечивает автоматизированную идентификацию нуклеотидных вариантов и устраняет ошибки, обусловленные качеством хроматограмм в методе Сэнгера. Помимо этого, применение NGS при анализе результатов

Таблица 4. Оценка стоимости генотипирования образцов при использовании секвенирования методом Сэнгера и с помощью NGS

Этап	Наименование (реагент/процесс)	Количество	Цена за шт. (руб.)	Итого (руб.)			
Секвенирование по Сэнгеру							
На сайт	Сайт-специфичные праймеры	Синтез олигонуклеотидов 20 оснований	2*	540	1080		
На образец	ПЦР	MBUision полимераза	1	20	20	555***	
	Реакция Сэнгера	GenSeq-1000	2	159.5	319		
	Очистка	Переосаждение спиртом		2	0		0
		Sephadex G-50		2 на 20 мг	4		8
		SeqMag-1000		2	33		66
		VAHTS DNA Clean Beads		2	55		110
		Секвенирование	Генетический анализатор «НАНОФОР 05»	2	104**		208
NGS							
На сайт	Сайт-специфичные праймеры	Синтез олигонуклеотидов 55 оснований	2	2035	4070		
Зависит от количества образцов в пуле	Индексирующие праймеры*	Синтез олигонуклеотидов 55 оснований	$2\sqrt{N}$, где N – кол-во образцов	2035	$4060\sqrt{N}$, где N – кол-во образцов		
На образец	ПЦР	MBUision полимераза	2	20	40	49.744	
	Секвенирование	GenoLab MV2.0 (ячейка FCM, 300 циклов, 250M прочтений)	0.01M прочтений	243 600	9.744		
На пул	Очистка	VAHTS DNA Clean Beads	1	110	110		

* В идеале следует использовать систему из четырех праймеров: два – для наработки продукта с помощью ПЦР и еще два – для секвенирования в обоих направлениях.

** Стоимость анализа одного образца в пересчете на расход полимера в 35-сантиметровом капилляре.

*** С учетом очистки на спин-колонках с сорбентом Sephadex G-50.

редактирования генома расширяет экспериментальный арсенал, позволяя решать задачи, недостижимые при традиционном секвенировании по Сэнгеру. Если метод Сэнгера фиксирует усредненный сигнал по каждой позиции и способен выявлять мутации лишь при их доле $\geq 20\%$ (по визуальному анализу хроматограмм) (Davidson et al., 2012; Bennett et al., 2020), то NGS предоставляет информацию по каждому аллелю, обеспечивая точность детекции до 2% и тем самым исключая двусмысленность интерпретации. Используя этот подход, мы оценили эффективность направляющей РНК в клеточной популяции.

Анализ данных NGS показал низкую частоту модификации ($\approx 5\%$ от общей популяции), что обусловило корректировку стратегии получения мутантных линий: была выбрана конструкция с селективным маркером для обогащения клеток, в которые успешно интегрирован вектор.

На следующих этапах данные NGS-анализа позволили нам выявить клеточные линии с нужной мутацией и без мозаичности, обеспечивая чистоту генетического материала, или выбрать клоны с нужной частотой аллеля, что критически важно для получения стабильных генетических линий. Благодаря высокой глубине покрытия (секвенирование ампликона обычно выполняли с покрытием 10000 прочтений на фрагмент около 300 п. н.) можно об-

наружить даже низкочастотные варианты, что невозможно при использовании менее чувствительных методов.

Распространено мнение, что NGS секвенирование – слишком дорогое и тем самым нивелирует свои преимущества. Однако важно понимать, что в нашем случае используется целевое секвенирование ампликонов, что снижает стоимость затрат на само секвенирование до < 10 руб. за образец. Основная статья затрат – индексирующие праймеры, число которых растет пропорционально \sqrt{N} , где N – количество образцов. Несмотря на это, данные праймеры универсальны: их можно применять как в амплификационном протоколе подготовки библиотек, описанном в работе, так и в стандартных схемах с лигированием универсальных адаптеров. Также возможно использование готовых коммерческих наборов индексирующих праймеров, однако зачастую это дороже индивидуального синтеза. В конечном итоге при учете цены заказа индексирующих праймеров, NGS-генотипирование начинает превосходить метод Сэнгера по стоимости при объеме > 80 образцов; без учета этой статьи (при уже имеющемся наборе индексирующих праймеров) экономия и вовсе достигается уже при анализе > 7 образцов (см. рис. 3), что делает предложенный метод перспективным для широкого круга исследователей (табл. 4).

NGS-секвенирование ампликонов может быть применено и при анализе эффективности других систем редактирования, основанных на системе CRISPR/Cas9. Если индуцируемая мутация, например, при получении протяженных делеций или инверсий, превышает возможности покрытия NGS (максимальная длина пары ридов 300 пар оснований), достаточно секвенировать только границы мутации, чтобы проверить целостность прилегающих участков и выявить нежелательные события мутагенеза. Метод применим, пока мутагенез не затронул сайты праймеров; тем же ограничением обладает и любой другой способ ПЦР-генотипирования.

Однако следует принимать во внимание ограничения и недостатки предлагаемого метода. Так как происходит большое количество циклов ПЦР, в данных секвенирования зачастую появляются амплификационные артефакты. Также из-за большого количества манипуляций с ампликонами до индексирования и значительной чувствительности метода высока чувствительность и к контаминации. Следует обязательно придерживаться правил правильной работы с пост-амплификационным материалом и исключать любые его контакты с местами приготовления смесей ПЦР и выделения ДНК. Кроме того, использование праймеров, содержащих длинные некомplementарные геномной ДНК последовательности на 5'-конце, приводит к общему снижению эффективности амплификации, что требует индивидуальной оптимизации условий ПЦР для каждой пары.

Заключение

В настоящей статье представлен метод получения модельных линий на основе ИПСК с маркированным синонимичной мутацией аллелем гена *AUTS2* с применением технологии редакторов оснований. Полученные линии ИПСК будут использованы для изучения влияния хромосомных aberrаций на экспрессию *AUTS2* и для исследования функциональных последствий этого нарушения.

Предложенный подход к генотипированию большого количества клонов с помощью NGS-секвенирования ампликонов демонстрирует высокую эффективность и масштабируемость при снижении стоимости анализа в сравнении с традиционными методами. Большая пропускная способность метода позволила провести анализ значительного количества образцов, а также оценить мозаицизм, состав и частоты аллелей в клеточной популяции.

Таким образом, представленная стратегия дает возможность эффективно получать клеточные модели для исследования влияния цис-регуляторных вариантов на транскрипционную активность генов, при этом контролируя генетическую гомогенность модифицированных клеточных моделей, что может быть востребовано как в фундаментальных исследованиях, так и в решении прикладных задач.

Список литературы / References

Смирнов А.В., Юнусова А.М., Лукьянчикова В.А., Баттулин Н.Р. Система CRISPR/Cas9 – универсальный инструмент геномной инженерии. *Вавиловский журнал генетики и селекции*. 2016; 20(4):493-510. doi 10.18699/VJ16.175

- [Smirnov A.V., Yunusova A.M., Lukyanchikova V.A., Battulin N.R. CRISPR/Cas9, a universal tool for genomic engineering. *Vavilovskii Zhurnal Genetiki i Seleksii = Vavilov J Genet Breed*. 2016;20(4): 493-510. doi 10.18699/VJ16.175 (in Russian)]
- Baux D., Van Goethem C., Ardouin O., Guignard T., Bergougnoux A., Koenig M., Roux A.-F. MobiDetails: online DNA variants interpretation. *Eur J Hum Genet*. 2021;29(2):356-360. doi 10.1038/s41431-020-00755-z
- Bennett E.P., Petersen B.L., Johansen I.E., Niu Y., Yang Z., Chamberlain C.A., Met Ö., Wandall H.H., Frödin M. INDEL detection, the 'Achilles heel' of precise genome editing: a survey of methods for accurate profiling of gene editing induced indels. *Nucleic Acids Res*. 2020;48(21):11958-11981. doi 10.1093/nar/gkaa975
- Billon P., Bryant E.E., Joseph S.A., Nambiar T.S., Hayward S.B., Rothstein R., Ciccio A. CRISPR-mediated base editing enables efficient disruption of eukaryotic genes through induction of STOP codons. *Mol Cell*. 2017;67(6):1068-1079.e4. doi 10.1016/j.molcel.2017.08.008
- Chen X., McAndrew M.J., Lapinaite A. Unlocking the secrets of ABEs: the molecular mechanism behind their specificity. *Biochem Soc Trans*. 2023;51(4):1635-1646. doi 10.1042/BST20221508
- Davidson C.J., Zeringer E., Champion K.J., Gauthier M.-P., Wang F., Boonyaratankornkit J., Jones J.R., Schreiber E. Improving the limit of detection for Sanger sequencing: a comparison of methodologies for *KRAS* variant detection. *BioTechniques*. 2012;53(3):182-188. doi 10.2144/000113913
- De Masi C., Spitalieri P., Murdocca M., Novelli G., Sangiuolo F. Application of CRISPR/Cas9 to human-induced pluripotent stem cells: from gene editing to drug discovery. *Hum Genomics*. 2020; 14(1):25. doi 10.1186/s40246-020-00276-2
- 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*. 2017;551(7681): 464-471. doi 10.1038/nature24644
- Geurts M.H., Gandhi S., Boretto M.G., Akkerman N., Derks L.L.M., Van Son G., Celotti M., ... Andersson-Rolf A., Chuvá De Sousa Lopes S.M., Van Es J.H., Van Boxtel R., Clevers H. One-step generation of tumor models by base editor multiplexing in adult stem cell-derived organoids. *Nat Commun*. 2023;14(1):4998. doi 10.1038/s41467-023-40701-3
- Global Burden of Disease Study 2021 Autism Spectrum Collaborators. The global epidemiology and health burden of the autism spectrum: findings from the Global Burden of Disease Study 2021. *Lancet Psychiatry*. 2025;12(2):111-121. doi 10.1016/S2215-0366(24) 00363-8
- Gridina M., Lagunov T., Belokopytova P., Torgunakov N., Nuriddinov M., Nurislamov A., Nazarenko L.P., ... Filipenko M., Rogaev E., Shilova N.V., Lebedev I.N., Fishman V. Combining chromosome conformation capture and exome sequencing for simultaneous detection of structural and single-nucleotide variants. *Genome Med*. 2025;17(1):47. doi 10.1186/s13073-025-01471-3
- Grünewald J., Zhou R., Garcia S.P., Iyer S., Lareau C.A., Aryee M.J., Joung J.K. Transcriptome-wide off-target RNA editing induced by CRISPR-guided DNA base editors. *Nature*. 2019;569(7756):433-437. doi 10.1038/s41586-019-1161-z
- Hwang G.-H., Park J., Lim K., Kim S., Yu J., Yu E., Kim S.-T., Eils R., Kim J.-S., Bae S. Web-based design and analysis tools for CRISPR base editing. *BMC Bioinformatics*. 2018;19(1):542. doi 10.1186/s12859-018-2585-4
- Jaganathan K., Kyriazopoulou Panagiotopoulou S., McRae J.F., Fazl Darbandi S., Knowles D., Li Y.I., Kosmicki J.A., ... Gao H., Kia A., Batzoglou S., Sanders S.J., Farh K.K.-H. Predicting splicing from primary sequence with deep learning. *Cell*. 2019;176(3):535-548.e24. doi 10.1016/j.cell.2018.12.015
- Jin S., Zong Y., Gao Q., Zhu Z., Wang Y., Qin P., Liang C., Wang D., Qiu J.-L., Zhang F., Gao C. Cytosine, but not ade-

- nine, base editors induce genome-wide off-target mutations in rice. *Science*. 2019; 364(6437):292-295. doi 10.1126/science.aaw7166
- 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;533(7603):420-424. doi 10.1038/nature17946
- Liang Y., Chen F., Wang K., Lai L. Base editors: development and applications in biomedicine. *Front Med*. 2023;17(3):359-387. doi 10.1007/s11684-023-1013-y
- Lu Z., Huang X. Base editors: a powerful tool for generating animal models of human diseases. *Cell Stress*. 2018;2(10):242-245. doi 10.15698/cst2018.10.156
- Rees H.A., Liu D.R. Base editing: precision chemistry on the genome and transcriptome of living cells. *Nat Rev Genet*. 2018;19(12):770-788. doi 10.1038/s41576-018-0059-1
- Reese M.G., Eeckman F.H., Kulp D., Haussler D. Improved splice site detection in Genie. In: Proceedings of the First Annual International Conference on Computational Molecular Biology (RECOMB '97). 1997;232-240. doi 10.1145/267521.267766
- Rowe R.G., Daley G.Q. Induced pluripotent stem cells in disease modelling and drug discovery. *Nat Rev Genet*. 2019;20(7):377-388. doi 10.1038/s41576-019-0100-z
- Salnikov P., Belokopytova P., Yan A., Viesná E., Korablev A., Serova I., Lukyanchikova V., Stepanchuk Y., Torgunakov N., Tikhomirov S., Fishman V. Direction and modality of transcription changes caused by TAD boundary disruption in *Slc29a3/Unc5b* locus depends on tissue-specific epigenetic context. *Epigenetics Chromatin*. 2025; 18(1):55. doi 10.1186/s13072-025-00618-1
- Uddin F., Rudin C.M., Sen T. CRISPR gene therapy: applications, limitations, and implications for the future. *Front Oncol*. 2020;10:1387. doi 10.3389/fonc.2020.01387
- Yu Y., Leete T.C., Born D.A., Young L., Barrera L.A., Lee S.-J., Rees H.A., Ciaramella G., Gaudelli N.M. Cytosine base editors with minimized unguided DNA and RNA off-target events and high on-target activity. *Nat Commun*. 2020;11(1):2052. doi 10.1038/s41467-020-15887-5
- Zuo E., Sun Y., Wei W., Yuan T., Ying W., Sun H., Yuan L., Steinmetz L.M., Li Y., Yang H. Cytosine base editor generates substantial off-target single-nucleotide variants in mouse embryos. *Science*. 2019;364(6437):289-292. doi 10.1126/science.aav9973

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Поступила в редакцию 14.06.2025. После доработки 05.11.2025. Принята к публикации 11.11.2025.