

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ  
НАУЧНОЕ УЧРЕЖДЕНИЕ «ФЕДЕРАЛЬНЫЙ  
ИССЛЕДОВАТЕЛЬСКИЙ ЦЕНТР ИНСТИТУТ ЦИТОЛОГИИ И  
ГЕНЕТИКИ СИБИРСКОГО ОТДЕЛЕНИЯ РОССИЙСКОЙ  
АКАДЕМИИ НАУК»

*На правах рукописи*

**САЛЬНИКОВ ПАВЕЛ АЛЕКСАНДРОВИЧ**

**Связь пространственной организации генома и регуляции  
экспрессии генов в локусах *Slc29a3/Unc5b* и *Erb41l4a* мыши**

**1.5.7. Генетика**

Диссертация на соискание ученой степени

кандидата биологических наук

Научный руководитель

доктор биологических наук

Фишман Вениамин Семенович

Новосибирск 2026

## Оглавление

<b>Список сокращений.....</b>	<b>4</b>
<b>Введение .....</b>	<b>6</b>
Актуальность темы исследования и степень ее разработанности .....	6
Цели и задачи исследования .....	8
Научная новизна работы .....	9
Теоретическая и практическая ценность работы.....	9
Основные положения, выносимые на защиту .....	10
Вклад автора .....	10
Апробация результатов .....	11
Структура и объем работы .....	12
Благодарности .....	12
<b>Обзор литературы.....</b>	<b>14</b>
Цис-регуляция генной экспрессии .....	14
Развитие методологии изучения пространственной архитектуры генома .....	17
Пространственная организация хроматина.....	21
Механизмы организации пространственной архитектуры генома на уровне ТАДов .....	24
Технологии редактирования генома с помощью направляемых нуклеаз .....	27
Влияние пространственной архитектуры на регуляцию экспрессии генов.....	38
Модельные локусы <i>Slc29a3/Unc5b</i> и <i>Erb4114a</i> .....	48

<b>Материалы и методы .....</b>	<b>56</b>
Материалы .....	56
Линия мышей.....	56
Олигонуклеотиды.....	56
Методы .....	57
Работа с мышами.....	57
Получение и подготовка CRISPR/Cas9 хнРНК.....	58
Выделение тотальной РНК .....	58
Обратная транскрипция.....	59
Выделение геномной ДНК.....	59
Полимеразная цепная реакция (ПЦР) .....	60
Цифровая ПЦР с флуоресцентными зондами.....	60
Анализ данных цифровой ПЦР с флуоресцентными зондами.....	61
Электрофорез в агарозном геле .....	61
Секвенирование по Сэнгеру .....	62
ПЦР-генотипирование мышей на делецию dM1R1 .....	63
ПЦР-ПДФ-генотипирование мышей на мутации в сайтах L2, M2, R2 .....	63
Приготовление NGS-библиотек для генотипирования.....	64
Генотипирование мышей на мутации в локусе <i>Erb41l4a</i> .....	64
Секвенирование иммунопреципитации хроматина (ChIP-Seq) .....	65
Статистический анализ результатов количественной ПЦР.....	66
Приготовление библиотек Hi-C.....	66
Приготовление библиотек для UMI-опосредованного таргетного РНК-секвенирования .....	68

Секвенирование РНК.....	69
<b>Результаты.....</b>	<b>70</b>
1. Стратегия выбора модельных локусов .....	70
2. Локус <i>Slc29a3/Unc5b</i> .....	78
Получение мышей, несущих делеции сайтов связывания CTCF в локусе <i>Slc29a3/Unc5b</i> .....	78
Изменения пространственной архитектуры локуса <i>Slc29a3/Unc5b</i> у мышей, гомозиготных по аллелю dM1R1L2M2 .....	87
Изменения экспрессии генов <i>Psap</i> , <i>Vsir</i> , <i>Cdh23</i> , <i>Slc29a3</i> , <i>Unc5b</i> и <i>Sgpl1</i> у мышей, гомозиготных по аллелю dM1R1L2M2, оцененные методом цифровой ПЦР .....	91
Изменения экспрессии генов <i>Psap</i> , <i>Vsir</i> , <i>Cdh23</i> , <i>Slc29a3</i> , <i>Unc5b</i> и <i>Sgpl1</i> у мышей, гомозиготных по аллелю dM1R1 и dM1R1L2M2, оцененные UMI-опосредованным методом секвенирования целевой РНК .....	94
3. Локус <i>Erb4114a</i> .....	99
Получение мышей, несущих хромосомные перестройки, затрагивающие границу ТАДов в локусе <i>Erb4114a</i> .....	99
Транскрипционный эффект хромосомных перестроек, затрагивающих границу ТАДов в локусе <i>Erb4114a</i> .....	102
<b>Обсуждение.....</b>	<b>107</b>
<b>Выводы.....</b>	<b>121</b>
<b>Список литературы .....</b>	<b>122</b>
<b>Приложения .....</b>	<b>136</b>

## Список сокращений

3C (Chromosome Conformation Capture) – класс методов, позволяющих определить пространственную архитектуру хроматина

ChIP-Seq (Chromatin Immunoprecipitation Sequencing) – технология, позволяющая анализировать распределение в хроматине ДНК-связывающих белков

COT-1 ДНК – реактив, использующийся для блокирования неспецифической гибридизации ДНК, представляет собой библиотеку фрагментов ДНК длиной от 50 до 300 п.н., обогащенную повторяющимися геномными последовательностями

CRISPR/Cas9 (clustered regularly interspaced short palindromic repeats/CRISPR associated protein 9) – технология редактирования генома, основанная на бактериальной системе адаптивного иммунитета

CTCF (CCCTC-binding factor) – транскрипционный фактор, инсулятор, важный организатор пространственной архитектуры хроматина

dM1R1 – делеция сайтов связывания CTCF M1 и R1 в локусе *Slc29a3/Unc5b* мыши

DSB (Double-Strand Break) – двуцепочечный разрыв ДНК

FDR (False Discovery Rate) – это ожидаемая доля ложных открытий среди всех результатов, признанных значимыми.

FISH (Fluorescence in situ Hybridization) – флуоресцентная ин ситу гибридизация, метод, позволяющий визуализировать пространственное расположение локусов ДНК с помощью гибридизации с флуоресцентными зондами на микропрепаратах

HDR (Homology-Directed Repair) – репарация разрывов ДНК по пути гомологичной рекомбинации

INDEL – мутации по типу потери или приобретения фрагментов ДНК длиной в несколько нуклеотидов

MMEJ (Microhomology-mediated end joining) – путь репарации с использованием соединения концов, опосредованного микрогомологией

NGS (Next Generation Sequencing) – технология секвенирования нового поколения

SINE (Short interspersed nuclear elements) – короткие диспергированные повторы, повторяющиеся последовательности геномной ДНК длиной от 100 до 700 п.н.

SNP (Single Nucleotide Polymorphism) – однонуклеотидные полиморфизмы, однонуклеотидный полиморфизм

SPF (Specific Pathogen Free) – стандарт содержания животных в отсутствие патогенных микроорганизмов и инфекционных агентов

ssODN (single-stranded oligo deoxynucleotide) – синтетический одноцепочечный ДНК-олигонуклеотид

UMI (unique molecular identifier) – уникальные молекулярные баркоды п.н. – пары нуклеотидов

ПЦР – полимеразная цепная реакция

ПЦР-рв (ПЦР в реальном времени) – количественная ПЦР, позволяющая количественно оценивать содержание определенных молекул ДНК в образце

ТАД – топологически ассоциированный домен

хнРНК – химерная направляющая РНК системы CRISPR/Cas

ЦКП – центр коллективного пользования

## Введение

### **Актуальность темы исследования и степень ее разработанности**

Развитие технологий захвата конформации хромосом и картирования сайтов связывания белков хроматина привело к формированию современного представления об организации генома в пространстве клеточного ядра [1]. Новые методы открыли возможности для изучения процессов компактизации хроматина с точки зрения их молекулярно-генетических механизмов, их взаимодействия и участия в процессах регуляции генной экспрессии. Выяснилось, что активность архитектурных механизмов приводит к формированию неравномерного распределения пространственных контактов между участками генома, связанного с функцией и активностью различных генетических элементов [2,3]. Так, цис-регуляторным элементам, таким как энхансеры, для осуществления своей функции необходима пространственная близость к промотору гена-мишени [4,5]. Таким образом, архитектура хроматина является одним из уровней регуляции генной экспрессии и других процессов в ядре клетки, который осуществляется путем контроля частот пространственных взаимодействий между регионами генома.

Исследование топологии хроматина методами высокого разрешения привело к открытию топологически ассоциированных доменов (ТАДов), регионов генома, внутри которых участки чаще взаимодействуют друг с другом, чем с элементами за пределами своего ТАДа [3]. У человека границы ТАДов в большинстве случаев формируются кластерами сайтов связывания инсуляторного фактора CTCF, препятствующими взаимодействию между генетическими элементами разных ТАДов [6]. В результате энхансеры значительно чаще взаимодействуют с промоторами, расположенными с ними в одном домене, а процессы ремоделирования хроматина, как правило, не способны преодолеть границы ТАДов. Считается, что ТАДы являются единицами эпигенетической информации, при этом пересечение границ

ТАДов может сопровождаться резкими изменениями эпигенетических сигнатур хроматина [7,8].

Нарушение структуры ТАДов, вызванное мутациями сайтов связывания CTCF или хромосомными перестройками, может приводить к нарушению цис-регуляции экспрессии близлежащих генов и, как следствие, развитию генетических патологий. Согласно последним данным, этот патогенетический механизм характерен для значительной доли (около 10%) описанных хромосомных перестроек человека [9]. Изменение архитектуры хроматина способствует онкогенезу: некодирующие мутации нарушают регуляцию экспрессии (в том числе активируют онкогены) и затрагивают элементы, контролирующие цикл, апоптоз и дифференцировку [10–13].

В то же время роль ТАДов в регуляции генной экспрессии не настолько однозначна. С одной стороны, как сайты связывания CTCF, так и границы топологических доменов эволюционно стабильны и находятся под селективным давлением. В ряде работ показано, что удаление кластеров сайтов связывания CTCF ведет к развитию патологий [14–17]. С другой стороны, удаление отдельных сайтов связывания далеко не всегда ведет к серьезным нарушениям генной экспрессии, а деплеция белков CTCF и когезина приводит к значимым изменениям в экспрессии лишь около 10% генов [18–22].

При этом нужно учитывать, что на сегодняшний день все работы, исследовавшие последствия нарушений структуры ТАДов, имеют существенные ограничения. Так, практически все они концентрировались на локусах генов развития, с уже имеющимися для них клиническими данными о патогенности некодирующих мутаций [14,15,23,24]. Этот класс генов характеризуется повышенным эволюционным давлением и сложной регуляцией, что отражается на структуре некодирующих регионов вокруг таких генов [4,25]. Кроме того, вышеприведенные работы проведены на

культурах клеток *in vitro* или на отдельных тканях *in vivo*. Таким образом, функциональные последствия модификаций отдельных сайтов связывания CTCF на организменном уровне остаются малоизученными, и, возможно, будут затрагивать различные механизмы регуляции генной экспрессии в зависимости от эпигенетического статуса локуса и траектории дифференцировки клетки.

В силу описанных обстоятельств, для углубления понимания роли ТАДов в регуляции генной экспрессии необходимо расширение круга исследованных случаев нарушений локальной архитектуры генома, и рассмотрение их в разных эпигенетических контекстах с использованием животных моделей.

### **Цели и задачи исследования**

В рамках данной работы мы ставим перед собой **цель** оценить влияние нарушений пространственной архитектуры хроматина в локусах *Slc29a3/Unc5b* и *Erb41l4a* на регуляцию генной экспрессии и ее тканеспецифичность.

В связи с этим, нами были поставлены следующие **задачи**:

1. Получить генетически модифицированных животных с делециями сайтов связывания белка CTCF в локусе *Slc29a3/Unc5b*.
2. Сравнить структуру трехмерных контактов в локусе *Slc29a3/Unc5b* мышцы у животных дикого типа и у животных с делециями сайтов связывания белка CTCF.
3. Сравнить уровень экспрессии аллелей генов *Psap*, *Vsir*, *Cdh23*, *Slc29a3*, *Unc5b* и *Sgpl1*, несущих в цис-положении делеции сайтов связывания CTCF, и аллелей дикого типа в различных органах и тканях мышцы.
4. Получить генетически модифицированных животных с мутациями в регионе, содержащем границу ТАДов в локусе *Erb41l4a* и сравнить уровень экспрессии генов локуса *Erb41l4a* у мышечей дикого типа и мышечей, несущих мутации границы ТАДов в этом локусе.

## Научная новизна работы

В данной работе впервые были описаны эффекты мутаций, приводящих к нарушению структуры ТАДов для локусов *Slc29a3/Unc5b* и *Epb41l4a*. На примере локуса *Slc29a3/Unc5b* были получены мышинные модели и исследованы нарушения пространственной организации хроматина для локусов, содержащих гены с широким профилем экспрессии среди тканей. Было показано, что для конститутивно активных генов нарушения архитектуры хроматина могут быть тканеспецифичными как по величине, так и по направлению эффекта. Было впервые показано, что хромосомные aberrации, затрагивающие границу ТАДов в терминальной области гена *Epb41l4a*, способны приводить к нарушению экспрессии гена *Nrep*, что, возможно, лежит в основе патогенеза неврологического фенотипа пациентов с хромосомными перестройками, затрагивающими эту границу. Кроме того, впервые было показано, что возможность взаимодействия двух эпигенетических ландшафтов локуса *Epb41l4a* при слиянии ТАДов имеет тканевую зависимость, и может реализовываться в одном органе, но не наблюдаться в другом.

## Теоретическая и практическая ценность работы

В данной работе предложен инновационный подход к исследованию цис-эффектов мутаций, основанный на использовании гибридных животных и моноаллельном анализе экспрессии генов методами секвенирования нового поколения. Такой подход позволяет однозначно различить цис- и транс-эффекты мутаций и имеет высокую чувствительность и точность измерений, сравнимые с цифровой ПЦР.

Кроме того, при получении генетически модифицированных линий животных, мы показали практическую возможность и высокую эффективность внесения коротких случайных мутаций с помощью системы CRISPR/Cas9 в сайты, близко расположенные друг к другу. В нашей работе

мы наблюдали отсутствие нежелательных исходов мутагенеза в виде делеций фрагментов ДНК между этими сайтами, что указывает на применимость такого подхода (при соответствующей оптимизации) в экспериментах геномного редактирования.

### **Основные положения, выносимые на защиту**

1. Функциональные последствия мутаций, нарушающих CTCF-зависимую пространственную архитектуру хроматина, зависят от эпигенетического контекста и не сводятся к универсальному однонаправленному эффекту – в негемеозисном локусе *Slc29a3/Unc5b* у мышей такие мутации приводят к тканеспецифическим изменениям в транскрипции целого ряда генов (*Psap*, *Vsir*, *Cdh23*, *Slc29a3*, *Unc5b*, *Sgpl1*), при этом уровни их экспрессии различаются как по направлению изменения, так и по амплитуде.

2. Хромосомные перестройки, затрагивающие сильную и эволюционно консервативную границу ТАДов в локусе *Erb4114a*, разделяющую регуляторные ландшафты гена *Arc* и гена *Nrep*, ассоциированного у человека с неврологическими заболеваниями, приводят у мышей к снижению экспрессии гена белка *Nrep* (Neuronal regeneration related protein) в тканях головного мозга.

### **Вклад автора**

Автором выполнены все экспериментальные процедуры, представленные в данной работе, а также анализ и обработка их результатов, за исключением микроинъекций в зиготы мыши и всей сопутствующей эмбриологической работы (данная часть работы выполнена Кораблевым А.Н. и Серовой И.А.), подготовки ChIP-Seq библиотек и их биоинформатической обработки (выполнены Лукьянчиковой В.А. и Белокопытовой П.С. соответственно), и биоинформатической обработки результатов Hi-C (выполненной Торгунаковым Н.Ю. и Весной Э.).

## Апробация результатов

Результаты данной работы были представлены на конференциях:

1. Bioinformatics of Genome Regulation and Structure / Systems Biology (BGRS/SB-2024 14-я Международная Мультиконференция) — 5–10.08.2024
2. Хромосома 2023. Международная конференция — 5–10.09.2023
3. Международный конгресс CRISPR-2023 — 11–13.09.2023
4. Bioinformatics of Genome Regulation and Structure / Systems Biology (BGRS/SB-2022 13-я Международная Мультиконференция) — 4–8.07.2022
5. МНСК-2020: Биология, 58-я Международная научная студенческая конференция — 10–13.04.2020
6. МНСК-2019: Биология, 57-я Международная научная студенческая конференция — 14–19.04.2019
7. МНСК-2018: Биология, 56-я Международная научная студенческая конференция — 22–27.04.2018

По материалам данной работы были опубликованы статьи в международных рецензируемых журналах:

1. Fishman VS, Salnikov PA, Battulin NR. Interpreting Chromosomal Rearrangements in the Context of 3-Dimensional Genome Organization: A Practical Guide for Medical Genetics. *Biochemistry (Mosc)*. 2018 Apr;83(4):393-401. doi: 10.1134/S0006297918040107. PMID: 29626926.
2. Kabirova E, Nurislamov A, Shadskiy A, Smirnov A, Popov A, Salnikov P, Battulin N, Fishman V. Function and Evolution of the Loop Extrusion Machinery in Animals. *Int J Mol Sci*. 2023 Mar 6;24(5):5017. doi: 10.3390/ijms24055017. PMID: 36902449; PMCID: PMC10003631.
3. Salnikov P, Korablev A, Serova I, Belokopytova P, Yan A, Stepanchuk Y, Tikhomirov S, Fishman V. Structural variants in the *Epb41l4a* locus: TAD

disruption and *Nrep* gene misregulation as hypothetical drivers of neurodevelopmental outcomes. *Sci Rep.* 2024 Mar 4;14(1):5288. doi: 10.1038/s41598-024-52545-y. PMID: 38438377; PMCID: PMC10912600.

4. 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 Aug 12;18(1):55. doi: 10.1186/s13072-025-00618-1. PMID: 40796890; PMCID: PMC12341078.

### **Структура и объем работы**

Работа включает в себя следующие разделы: введение, обзор литературы, материалы и методы, результаты, их обсуждение, выводы, список литературы, приложения. Общий объем работы содержит 148 страниц, 26 рисунков и 1 таблицу, библиография содержит 211 источников.

### **Благодарности**

Искренне благодарю своего научного руководителя, Вениамина Семеновича Фишмана, за внимательное наставничество и создание теплой рабочей атмосферы лаборатории. Спасибо за терпение и поддержку, за чуткое руководство и неизменный оптимизм, за щедрость на идеи и предоставленную свободу научного творчества.

Я глубоко признателен старшим коллегам, а именно Гридиной Марии Михайловне, Баттулину Нариману Рашитовичу и Хабаровой Анне Александровне за неизменное участие и готовность прийти на помощь. Отдельную благодарность выражаю соавторам этой работы, чья вовлеченность и профессионализм сделали возможной реализацию этой работы, а именно Кораблеву Алексею Николаевичу, Серовой Ирине Александровне, Лукьянчиковой Варваре Алексеевне, Белокопытовой Полине Станиславовне, Эмиль Весне и Торгунакову Никите Юрьевичу.

Благодарю друзей, которые сделали добрее рабочие будни: Александру Петровну Ян за регулярные обеды, Эдуарда Александровича Чуйко за ужины, Артема Рустамовича Нурисламова за доставку морепродуктов, Яну Константиновну Степанчук за прекрасный музыкальный вкус и Мунгалова Романа Владимировича за научное чутье и карьерную смелость.

За поддержку на всем пути благодарю свою семью и друзей.

## Обзор литературы

### Цис-регуляция генной экспрессии

Несмотря на десятилетия пристального изучения, вопрос о механизмах и принципах регуляции пространственно-временной экспрессии генов остается одним из ключевых вопросов не только генетики, но и биологии вообще. Сегодня мы знаем, что в сложной регуляции активности генов немаловажную роль играют цис-регуляторные элементы. Наиболее влиятельные и глубоко изученные классы регуляторных элементов — промоторы и энхансеры [26]. Также выделяют репрессивные цис-элементы, такие как сайленсеры [27] и инсуляторы — участки ДНК, связывающие белки и блокирующие действие энхансера, если расположены между ним и целевым промотором [28,29]. На сегодняшний день единственный известный инсуляторный белок млекопитающих — СССТС-binding factor (CTCF), обладающий одиннадцатью ДНК-связывающими доменами цинковых пальцев [30]. Однако биохимические механизмы его инсуляторной активности были раскрыты только в последнее время [7,31].

Принято определять промоторы и энхансеры как участки ДНК, размеры которых варьируют от нескольких десятков до тысяч пар оснований, обладающие способностью активировать транскрипцию [4]. Общей чертой энхансеров и промоторов является наличие множества сайтов связывания транскрипционных факторов, которые облегчают транскрипцию. Тем не менее, промоторы обладают некоторыми уникальными характеристиками. Промоторы млекопитающих состоят из коровой и проксимальной частей [32,33]. Структура коровой части промоторов не уникальна и состоит из сайтов связывания РНК-полимеразы и общих транскрипционных факторов. Таким образом, этот элемент служит для правильного формирования преинициаторного комплекса РНК-полимеразы II в пределах 50-100 пар нуклеотидов от начала транскрипции (TSS) (34). Коровые промоторы обычно имеют низкую базовую активность, общие мотивы, такие как ТАТА-бокс и

СрG-островки [34,35], и их активность в широких пределах может быть модулирована другими более удаленными цис-регуляторными элементами.

В отличие от этого, проксимальные части промоторов расположены на большем удалении от TSS и служат платформой для связывания тканеспецифических транскрипционных факторов. Функционально неясно, как проксимальный промотор взаимодействует с основным и как он способствует активации гена. Одна из возможностей — его взаимодействие с окружающими энхансерными элементами, выступающими посредниками в передаче сигнала, необходимого для инициации процесса элонгации полимеразными комплексами, собранными на коровом промоторе [32,36].

В то время как промоторы расположены строго в начале гена, энхансеры могут находиться на расстоянии многих тысяч пар нуклеотидов от целевого промотора [37]. Зачастую энхансеры используются для регуляции уровня экспрессии генов в ходе клеточной дифференцировки, поэтому профиль их активности специфичен для типа клеток. Более того, из-за их разнообразного положения и отсутствия привязки к TSS поиск энхансеров остается нетривиальной задачей. В самом строгом варианте, они определяются по своей функции, то есть способности повышать уровень транскрипции гена, показанной с помощью молекулярно-генетических (STARR-Seq) [38] или трансгенных репортерных методов [39,40].

Идентификация энхансеров с помощью геномных данных может быть проведена с использованием данных о доступности хроматина (DNaseI-гиперчувствительность или ATAC-seq) и модификациях гистонов (H3K27ac + H3K4me1) [41]. Кроме того, зачастую последовательности энхансеров проявляют значительную степень консервативности, сохраняемые отбором на протяжении миллионов лет в силу своей функциональной значимости [42,43].

Эти подходы показали, что энхансеры проявляют узкоспецифичные пространственно-временные сигнатуры и могут действовать комбинаторно

для достижения сложных паттернов генной экспрессии [44,45]. Интересно, что, хотя гены домашнего хозяйства обычно не имеют дистально расположенных энхансеров, гены развития или тканеспецифичные структурные гены часто регулируются сразу несколькими энхансерами, иногда образуя крупные кластеры, называемые суперэнхансерами [46,47].

Точные механизмы взаимодействия энхансеров и промоторов остаются в значительной степени неясными. Считается, что энхансеры функционируют независимо от положения и ориентации, активируя свои целевые гены и могут располагаться на значительном расстоянии от своего целевого промотора [37,48], тем не менее требуя физической близости для осуществления своей функции [26]. Общепринято представление о том, что промотор и дистальный энхансер сближены в ядерном пространстве, образуя основание хроматиновой петли [49]. Были предложены различные модели о возникновении и поддержании петлевого взаимодействия между промотором и энхансером. В настоящее время считается, что установление таких контактов облегчается с помощью механизмов, организующих пространственную архитектуру ядра, в том числе процессами фазовой сепарации хроматина [50].

Так, подобно промоторам, энхансеры содержат специфичные мотивы связывания транскрипционных факторов, которые, в свою очередь, могут устанавливать мультивалентные белок-белковые взаимодействия. За счет этого свойства достигается увеличение локальной концентрации белков транскрипционного аппарата, а также других транскрипционных факторов, связывающих ДНК и привлекающих определенные геномные последовательности в такого рода «транскрипционную фабрику» [51,52]. Образование мембранонезависимых конденсатов в ядре, содержащих транскрипционные факторы, коактиваторный белок Медиатор и РНК-полимеразу, было описано для фактора плюрипотентности OCT4 [53] и гомеозисного гена *NoxD13* [54].

В заключение, цис-регуляторные элементы в сочетании с транскрипционными факторами необходимы для установления сложных пространственно-временных паттернов генной экспрессии для определения судьбы клеток и развития многоклеточных организмов. Тем не менее, остаются не до конца ясными механизмы осуществления функций энхансеров, а также установление и взаимоспецифичность промотор-энхансерного взаимодействия в сложных эпигенетических ландшафтах, объединяющих большое количество генов и энхансеров в одном локусе. В настоящее время открываются все новые данные о важном участии в этих процессах механизмов пространственной архитектуры хроматина [55,56].

### **Развитие методологии изучения пространственной архитектуры генома**

Развитие молекулярных методов значительно расширило наше понимание трехмерной архитектуры ядра. Первые прорывные исследования в этой области были достигнуты благодаря выходу за рамки традиционной световой микроскопии. Первые исследования пространственной организации генетического материала в интерфазном ядре были выполнены с помощью технологии флуоресцентной гибридизации *in situ* (FISH). С помощью этого метода впервые были открыты ключевые закономерности, такие как существование хромосомных территорий и радиальную асимметрию в распределении эухроматина и гетерохроматина [57]. Так, выяснилось, что хромосомы занимают обособленные территории внутри ядра клетки, при этом длинные хромосомы обычно располагаются на периферии, а меньшие тяготеют к его центру [58]. Несмотря на это разделение, границы соседних территорий могут диффузионно перекрываться [59]. Кроме того, хромосомный материал внутри одной территории дополнительно сегрегирует на области, условно богатые и бедные генами [60].

Кроме того, оказалось, что расположение генетических локусов в пространстве ядра связано с уровнем их активности. Например, некоторые

гены способны значительно изменять свое положение относительно своих хромосомных территорий при изменении статуса активности [61,62]. Аналогично, экспрессия генов коррелирует с их расположением относительно радиальной оси ядра, причём неактивные гены обычно тяготеют ближе к периферическим его областям [63].

Более того, уже при помощи FISH было установлено, что относительная пространственная близость генов в ядре также связана с процессами их эпигенетической регуляции. Так, некоторые группы генов, связанных общей функцией и временем активации, колокализуются в пространстве ядра при активации [64]. При этом клеточная визуализация *in vivo* показывает, что участки генома обычно имеют ограниченную подвижность, и не способны свободно диффундировать [65] без контроля специфических молекулярных механизмов. Пространственная ориентация локусов хроматина может модифицироваться транскрипционными факторами, связывающимися с определёнными последовательностями ДНК [66]. Так, было показано, что специфические факторы, такие как белки, ассоциированные с ламинной (внутренней белковой оболочкой ядра), могут перемещать геномные локусы к периферии ядра, что одновременно сопровождается снижением транскрипции [67].

Вторым революционным шагом в исследовании пространственной архитектуры генома стало изобретение метода захвата конформации хромосом (Chromosome Conformation Capture, 3C), использующего биохимические реакции, опосредованные близостью сегментов ДНК, для изучения их пространственной организации. Пионерская работа [68], представившая миру метод 3C, стимулировала многочисленные достижения в геномных исследованиях. Чтобы оценить эти разработки, важно понимать базовые принципы метода 3C [1].

Процесс начинается с фиксации хроматина формальдегидом для сохранения трехмерной структуры ДНК. Затем ДНК сегментируется с помощью сайт-специфичных эндонуклеаз, таких как ферменты бактериальной рестрикции *HindIII* (6-нуклеотидный сайт узнавания) или более часто щелящих, например, *DpnII* (4-нуклеотидный сайт). После этого происходит лигирование получившихся ДНК-фрагментов. При этом ковалентно сшиваются участки генома, которые оказались в пространственной близости друг к другу, но при этом не обязательно находившиеся в линейной близости в молекуле ДНК. Для анализа пространственной конфигурации генетических элементов в 3С количественно определяются события лигирования между двумя интересующими сегментами ДНК с помощью полуколичественной или количественной ПЦР.

С момента своего введения 3С становится ключевым методом для изучения пространственной организации хромосом. С его помощью было открыто, что регуляторные ДНК-элементы для осуществления своей активности пространственно контактируют с промоторами своих целевых генов. В локусе генов  $\beta$ -глобина регуляторный энхансерный элемент LCR (Locus Control Region) формирует хроматиновый хаб, в который входят промоторы активных генов локуса, состав которых динамично меняется в ходе внутриутробного развития человека [69]. Также образование регуляторных петель было продемонстрировано на примере регуляции генов локусов *H19–Igf2* [70], интерлейкина  $T_H2$  [71] и  $\alpha$ -глобина [72].

С развитием технологий появились модификации метода 3С, позволяющие интегрировано анализировать пространственную архитектуру геномных локусов с разрешением, намного превышающим возможности микроскопии. Распространение технологии секвенирования нового поколения (NGS) стимулировало разработку метода полногеномного анализа архитектуры хроматина, получившего название Hi-C [2]. В Hi-C протокол приготовления 3С-библиотеки слегка модифицирован. После того как

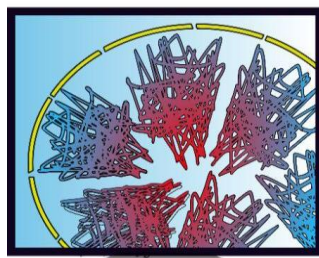
фиксированная ДНК фрагментируется, липкие концы ДНК достраиваются до тупых с включением биотин-меченных нуклеотидов, и затем лигируются. В дальнейшем ДНК очищается и фрагментируется воздействием ультразвука. С помощью наночастиц, покрытых биотин-связывающим белком стрептавидином, производится выделение участвовавших в реакции фрагментов ДНК, содержащих стык лигирования. Эти фрагменты ДНК секвенируются, и парные прочтения, приходящиеся на разные рестрикционные фрагменты, оцениваются как пространственные взаимодействия. Таким образом, можно построить матрицу частот пространственных взаимодействий между рестрикционными фрагментами по всему геному [2,3].

Разрешение (т.е. минимальный размер фрагмента, для которого можно проанализировать распределение контактов хроматина в ядре) для ранних экспериментов Hi-C на клетках млекопитающих составляло около 1 миллиона оснований. Данные такого качества уже позволили исследователям подтвердить ранее существовавшие представления о сегрегации участков генома на два компартмента, в зависимости от статуса их генной активности [2,73]. В настоящее время хорошим разрешением считается 5 тысяч п.н., однако данные даже такого качества требуют очень глубокого секвенирования Hi-C библиотек [74]. Тем не менее, исследователями был разработан ряд технологических приемов, позволяющих снизить стоимость Hi-C эксперимента, например, таргетное обогащение с помощью панели гибридизационных зондов. Данные таких экспериментов позволили исследовать пространственную архитектуру генома на уровне регуляторных петель между промоторами и энхансерами. Эти эксперименты привели к открытию новых структурных и функциональных явлений в ядре млекопитающих [16,75–78].

## **Пространственная организация хроматина**

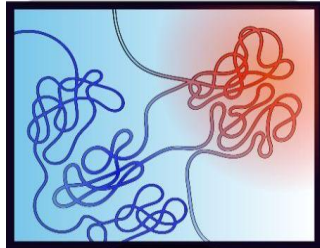
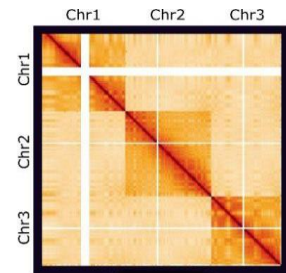
Эксперименты Hi-C генерируют обширные данные о частотах попарных контактов между локусами всего генома, которые обычно визуализируют с помощью карты пространственных контактов. Эта карта представляет собой матрицу, где строки и столбцы соответствуют определенным фрагментам генома, а частоты контактов отображаются в ячейках с использованием цветовой шкалы для облегчения визуального анализа. Такие карты демонстрируют организацию хроматина внутри ядра, давая возможность анализировать структуры, включающие в свой состав от нескольких тысяч до миллионов пар оснований.

В первую очередь Hi-C карты генома человека и других млекопитающих дают возможность на основе паттерна пространственных контактов выделить геномные компартменты (Рис. 1). Принято разделять геном на активный (А) и неактивный (В) компартменты, характеризующиеся тем, что регионы внутри одного компартмента часто контактируют друг с другом, а регионы из разных компартментов — редко [2,79]. Это приводит к проявлению на пространственных картах так называемого «шахматного» узора, где яркие ячейки, указывающие на обильные контакты внутри компартмента, чередуются с темными ячейками, соответствующими контактам между регионами из разных компартментов. Длина протяженных регионов, входящих в один компартмент, по геномным меркам достаточно велика, и варьирует от миллиона до десятков миллионов пар оснований.



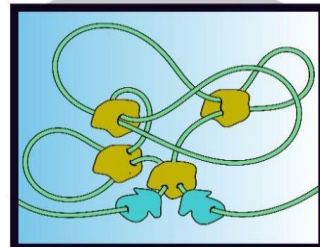
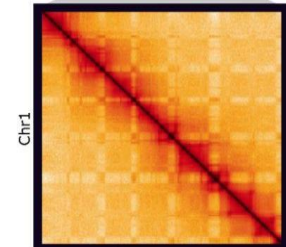
### Хромосомные территории

Каждая хромосома чаще контактирует сама с собой, чем с другими хромосомами



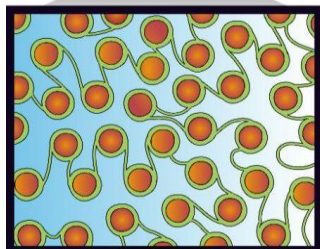
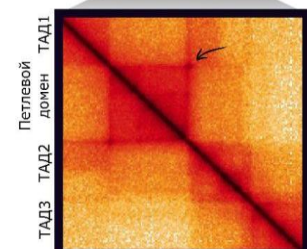
### А и В компартменты

Участки хроматина контактируют преимущественно с участками, находящимися в том же эпигенетическом и транскрипционном статусе



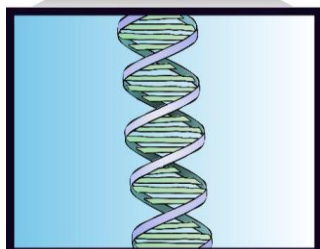
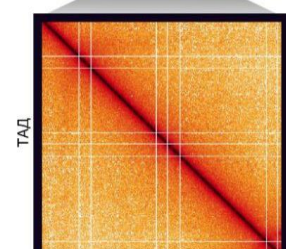
### ТАДы

Между сайтами CTCF возникают петли с повышенной частотой внутренних контактов по отношению к наружным



### Нуклеосомная фибрилла

Комплекс ДНК и гистонов (~11 нм в диаметре) образует структуру, объясняющую экспоненциальный рост контактов при приближении по первичной структуре ДНК



### дцДНК

Двойная спираль молекулы дезоксирибонуклеиновой кислоты диаметром 2,2 - 2,4 нм. Нерегулярный биополимер, обеспечивает хранение генетической информации о развитии и функционировании живых организмов

Рисунок 1. Иерархические уровни пространственной организации генома и их отражение на картах Hi-C. Слева схематически показаны основные уровни упаковки и пространственной организации хроматина. Справа приведены соответствующие примеры карт пространственных контактов Hi-C, отражающие характерные паттерны взаимодействий для каждого уровня организации (из [74]).

Хотя разделение локусов на компартменты основано на распределении пространственных контактов, оно тесно коррелирует с их генетическими и эпигенетическими характеристиками [80]. Так, компартмент А обычно

характеризуется высокой плотностью генов, транскрипционной активностью и активными эпигенетическими метками хроматина, в то время как компартмент В, соответствующий гетерохроматину, ассоциирован с противоположными характеристиками [2]. В некоторых исследованиях были предприняты попытки дополнительно подразделить компартменты на дополнительные категории, каждая из которых характеризуется отличительными эпигенетическими метками, наподобие системы типов хроматина дрозофилы. Однако в этом направлении не удалось достичь обоснованного решения, и консенсуальным в современной науке является разделение хроматина на А и В [3].

Топологически ассоциированные домены (ТАДы, Рис. 1) выделяются на картах Hi-C высокого разрешения как области генома, инсулированные от своего окружения [81,82]. Эти домены характеризуются более высокой частотой внутридоменных взаимодействий по сравнению с междоменными взаимодействиями, и потому формируют видимые треугольники вокруг главной диагонали на картах Hi-C. Их размер варьируется от 20 тысяч пар нуклеотидов до одного миллиона, и они являются значимыми структурными единицами в геноме. Границы ТАДов часто совпадают с границами других геномных и эпигенетических единиц, включая компартменты А и В, регионы различного времени репликации [83]. Кроме того, показано, что гены в пределах одного ТАДа показывают коррелирующие уровни экспрессии среди клеточных типов [25].

Границы ТАДов обогащены генами домашнего хозяйства [84], элементами SINE и сайтами связывания транскрипционного фактора CTCF [3,85]. ТАДы демонстрируют высокую степень консервативности между различными типами клеток и даже между видами в синтенических регионах их геномов, что указывает на фундаментальную роль ТАДов в структуре генома. Наряду с консервативностью локальная структура ТАДов способна

изменяться в процессе дифференциации клеток, и может быть нарушена мутациями, затрагивающими границы ТАДов [80].

Формирование хроматиновых петель — еще один важный аспект пространственной архитектуры хроматина. Они проявляются как яркие точки или локальные максимумы на матрице Hi-C, расположенные на некотором отдалении от главной диагонали [3]. Петли хроматина часто участвуют в регуляции генов, пространственно сближая цис-регуляторные элементы между собой [86,87]. Петли обычно закрепляются в местах связывания CTCF, расположенных у их основания, и ориентация и целостность этих мест жизненно важны для стабильности и функционирования петли [30,88].

Интеграция данных о пространственной организации генома с данными других областей дает возможность глубже понять механизмы регуляции генов и причины фенотипических манифестаций генетических мутаций, не затрагивающих кодирующие участки генов. Открытие сложной и динамичной архитектуры хроматина подчеркивает важность учета не только самих генетических последовательностей, но и их эпигенетических модуляций в понимании функции генома в норме и патологии [74].

### **Механизмы организации пространственной архитектуры генома на уровне ТАДов**

Долгое время молекулярные механизмы формирования ТАДов и хроматиновых петель у млекопитающих были ключевыми вопросами в области изучения пространственной организации генома. Исходя из данных о колокализации Когезина и транскрипционного фактора CTCF на границах ТАДов была предложена модель протягивания петли (Loop Extrusion) [89]. Модель предполагает, что Когезин, известный по своей традиционной функции удержания сестринских хроматид во время митоза, в интерфазном ядре млекопитающих проявляет моторную активность в отношении ДНК, вытягивая петли через свою кольцевую структуру.

Интересно, что Когезин относится к семейству белков SMC [90] — эволюционно консервативных ДНК-связывающих АТФаз, встречающихся у всех живых организмов. Показано, что уже у бактерий эти белки как участвуют в процессах когезии и сегрегации генетического материала, так и могут образовывать петли ДНК. У эукариот белки SMC входят в комплексы Когезин [91], Конденсин I и II [92], и комплекс SMC5/6 [93], функционирующие как гетеродимеры (например, Когезин состоит из белков SMC1-SMC3). Белки SMC связаны с белками из суперсемейства Клейзина, например, RAD21 в Когезине и CAP-H в Конденсинах, которые действуют как мостики, связывающие АТФазные головки SMC [94]. Последним обязательным участником комплекса является STAG1/STAG2, который связывается с RAD21 и обязателен для поддержания структуры и функции когезина в связывании ДНК [95].

Загрузка Когезина на хромосомы требует участия специального комплекса-загрузчика NIPBL-MAU2, связывающегося с RAD21 [96,97]. Механизм входа ДНК в кольцо Когезина достоверно не известен. В настоящее время считается, что для осуществления архитектурных функций, Когезин в комплексе с загрузчиком связывается с нитью интерфазной хромосомы в двух местах, и затем протягивает ее сквозь себя. Ранее предполагалось, что обе нити ДНК, соединяемые Когезином, оказываются продеты внутрь кольца, однако новейшие представления состоят в том, что либо хотя бы одна, либо обе молекулы ДНК связываются с Когезином снаружи электростатическими взаимодействиями [97]. Благодаря такому нетопологическому связыванию Когезин во время экстрезии хроматина может преодолевать крупные молекулярные препятствия на своем пути [98].

Движение Когезина вдоль ДНК ограничивается белками CTCF, которые связываются с ДНК в специфических непалиндромных сайтах [99]. Последние исследования предполагают, что N-концевой домен CTCF связывается с 86-аминокислотным мотивом комплекса STAG1/2-RAD21 [100]. Это

взаимодействие приводит к стабилизации петли и остановке процесса экструзии. Важно, что это взаимодействие происходит независимо от пространственной ориентации CTCF относительно движения когезинового комплекса вдоль ДНК. Было показано, что наиболее стабильные петли образуются между парами конвергентно направленных сайтов связывания CTCF [101]. Способность CTCF останавливать Когезин, формируя стабильные петли, является ключевой для формирования как топологических доменов, так и петель хроматина внутри них.

Установление механизма моторного действия Когезина в отношении ДНК остается отдельно стоящей задачей в силу сложности как самого белкового комплекса, так и всего процесса. В настоящее время считается, что Когезин способен активно протягивать сквозь себя нить ДНК, гидролизуя молекулы АТФ. Теоретически, процесс протягивания петли Когезином может быть как однонаправленным, то есть активно через комплекс транслоцируется только одна нить ДНК, так и двунаправленным, обеспечивающим симметричный рост петли с обеих сторон [102]. В настоящее время получены данные, подтверждающие первую гипотезу у примитивных организмов (таких как дрожжи [103] и *C. elegans* [104]) для комплекса Конденсина, организующего компактизацию хромосом в митозе. С другой стороны, фиксация когезина на одном определенном месте ДНК должна была бы характерно отражаться на данных Hi-C млекопитающих. Предполагается, что видимая симметрия хроматиновых петель и ТАДов может быть результатом взаимодействия двух асимметричных комплексов Когезина, участвующих в протягивании одной петли [105].

Интересен вопрос об интерференции процесса протягивания петли с другими белками и явлениями в клеточном ядре. Известно, что петлеобразование хроматина имеет антагонистическое действие по отношению к процессам компартментализации и фазовой сепарации ядерного материала [20,106]. Кроме того, известно, что активная транскрипция и

продвижение РНК-полимеразы препятствует образованию и стабилизации хроматиновых петель, диссоциируя ДНК и связанный с ней CTCF [107]. Не исключено, что и некоторые другие компоненты могут динамично взаимодействовать с архитектурными белками, влияя на динамику вытягивания петли либо как препятствия [108], либо, возможно, даже способствуя процессу.

### **Технологии редактирования генома с помощью направляемых нуклеаз**

Для дальнейшего погружения в тему участия пространственной архитектуры хроматина в регуляции генной экспрессии, нам необходимо рассмотреть современные возможности генетической инженерии клеток животных. Технологии на основе направляемых нуклеаз являются одним из наиболее распространенных в данной области подходов и широко применяются для высокоточного редактирования генома в живых клетках. Их действие основано на возможности узнавания некоторой заданной последовательности ДНК и индукции двуцепочечных разрывов (DSB) в этом сайте, после чего происходит репарация разрыва клеточными системами. В зависимости от механизма восстановления ДНК, помимо восстановления исходной последовательности, результатом могут быть небольшие инсерции и делеции, точечные замены при наличии донорной матрицы, а также более крупные структурные изменения. Существует несколько классов таргетируемых нуклеаз, позволяющих индуцировать направленные изменения генома в живых клетках.

ZFN (Zinc-finger Nuclease) — технология создания искусственных таргетируемых нуклеаз, включающих неспецифический нуклеазный домен FokI и ДНК-связывающие домены типа «цинковых пальцев», обеспечивающие распознавание заданной последовательности ДНК (каждый такой домен узнает определенный нуклеотидный триплет) [109]. Методами генной инженерии можно получить белок, специфически узнающий

выбранный участок генома. При связывании двух таких белков с соседними сайтами домены FokI димеризуются и вносят DSB. Однако разработка ZFN остается технически сложной и трудоемкой. Кроме того, все подходы, основанные на таргетируемых нуклеазах, имеют общий существенный недостаток, а именно риск внесения разрыва ДНК не в целевой, а случайный сайт генома (так называемый *off-target*-эффект). Они могут возникать из-за вариабельности генома или недостаточной специфичности системы. Поэтому применение направляемых нуклеаз требует дополнительных мер контроля, включая возвратные скрещивания при получении линий мышей [110,111], а также комплексную оценку безопасности: дизайн направляющих РНК, *in silico* предсказание потенциальных *off-target*-сайтов, биохимические и клеточные методы, targeted deep sequencing, WGS/long-read sequencing и специализированные подходы для выявления структурных вариантов. Отдельной проблемой при использовании таких систем *in vivo* является возможная иммуногенность компонентов редактирования или систем доставки.

TALEN (Transcription Activator-Like Effector Nucleases) — это искусственные таргетируемые нуклеазы, созданные на основе ДНК-связывающих доменов белков TAL бактерий рода *Xanthomonas* и нуклеазного домена FokI. Специфичность TALEN определяется синтетическим ДНК-связывающим доменом, состоящим из повторяющихся модулей. Каждый такой модуль обычно распознаёт один нуклеотид ДНК. Благодаря этому TALEN можно проектировать под заданную последовательность, последовательно подбирая модули, узнающие нужные нуклеотиды [109]. По сравнению с ZFN эта технология более удобна для проектирования и получила более широкое применение. Однако создание TALEN остается трудоемким и требует более высокого уровня технической подготовки по сравнению с системой CRISPR/Cas9 [112].

CRISPR/Cas (Clustered Regularly Interspaced Short Palindromic Repeats / CRISPR associated protein) — семейство систем редактирования генома, первая из которых (CRISPR/Cas9) впервые была успешно адаптирована для направленного внесения изменений в ДНК эукариотических клеток в 2012 году [113]. С тех пор CRISPR/Cas9 стала одной из доминирующих платформ геномного редактирования в фундаментальных исследованиях и клинических приложениях. В 2023 году FDA одобрило Casgevy как первую терапию, использующую CRISPR/Cas9, для лечения пациентов с серповидноклеточной анемией [114–116].

Широкое распространение CRISPR/Cas9 во многом связано с простотой проектирования системы: для задания новой специфичности достаточно изменить 20-нуклеотидный спейсер направляющей РНК, комплементарный целевому участку генома [117]. Это существенно упростило проведение экспериментов по сравнению с ZFN и TALEN и способствовало внедрению CRISPR/Cas9 практически во все направления генетических исследований.

Системы CRISPR/Cas возникли у бактерий и архей как механизм адаптивного иммунитета против бактериофагов и других мобильных генетических элементов [118]. При адаптации короткие фрагменты чужеродной ДНК интегрируются в CRISPR-локус в виде спейсеров. С этого локуса транскрибируется pre-crRNA, которая далее процессируется в crRNA, содержащую повторенную последовательность (необходимую для узнавания Cas-белками) и переменный спейсер, комплементарный чужеродной ДНК. Соответствующий спейсеру участок в геноме фага называется протоспейсером; рядом с ним обязательно находится PAM, короткий мотив, необходимый для распознавания мишени Cas-нуклеазой и отличающий чужеродную ДНК от собственного CRISPR-локуса клетки (кроме того, он необходим для процесса адаптации, то есть приобретения новых спейсеров системой CRISPR). Cas9 образует комплекс с crRNA и tracrRNA, распознаёт протоспейсер, комплементарный спейсеру crRNA, и вносит DSB, как правило,

на расстоянии трёх нуклеотидов от PAM с образованием тупых концов. Это приводит к деградации фаговой ДНК и обеспечивает защиту клетки от инфекции.

Технологическая версия CRISPR/Cas9 (Рис. 2) включает РНК-направляемую нуклеазу Cas9, способную взаимодействовать с ДНК без дополнительных белковых кофакторов [119]. Вместо двух молекул РНК (crRNA и tracrRNA, необходимых для работы природной системы) часто используют искусственную химерную направляющую РНК (хнРНК), объединяющую спейсер, повторную часть crRNA и последовательность tracrRNA. Такое упрощение облегчает применение системы и повышает её экспериментальную эффективность.

Помимо SpCas9, активно применяются и развиваются другие CRISPR-эффекторы и производные платформы. Современная классификация включает 2 класса, 7 типов и 46 подтипов CRISPR/Cas-систем [118], отличающихся последовательностью PAM, субстратом связывания и каталитической функцией. Однако, несмотря на разнообразие CRISPR-инструментов, наиболее широко используемым вариантом остаётся CRISPR/Cas9 из *Streptococcus pyogenes* (SpCas9), что во многом связано с коротким и часто встречающимся PAM-мотивом (5'-NGG-3'). Это обеспечивает высокую гибкость выбора целевых сайтов в геноме, однако сопровождается рисками нецелевого редактирования похожих последовательностей и других побочных эффектов.

Ранние эксперименты по применению CRISPR/Cas9 выявили относительно высокую частоту нецелевых мутаций [120,121]. Это связано с тем, что отдельные нуклеотиды 20-нуклеотидного спейсера вносят неодинаковый вклад в специфическое распознавание протоспейсера Cas9. Наиболее критична PAM-проксимальная область, тогда как несовпадения в более удалённых от PAM позициях переносятся системой значительно лучше.

В частности, разрез может происходить в сайтах, содержащих несовпадения в PAM-дистальной части спейсера, включая 15–20-е позиции. При этом чем чаще встречается PAM-мотив Cas9 в геноме, тем выше и число возможных нецелевых сайтов. Для снижения риска *off-target*-эффектов могут использоваться Cas-нуклеазы с более длинными или редкими PAM, чем у SpCas9, а также инженерные варианты SpCas9 с повышенной специфичностью. Такие модификации уменьшают вероятность разреза в участках с неполной комплементарностью направляющей РНК [122].

Более существенной проблемой является вариабельность исходов репарации DSB, так называемые нежелательные *on-target*-эффекты. Показано [111,123], что при внесении геномных делеций, результатом вмешательства могут быть также инверсии и дупликации удаляемых участков. Эти данные подчеркивают необходимость тщательного дизайна CRISPR/Cas9-экспериментов и последующей валидации полученных генотипов. По мере перехода CRISPR-терапий к клиническому применению особое значение приобретает анализ не только классических *off-target*-эффектов, но и *on-target*-аббераций, транслокаций и других нарушений геномной целостности [124]. Следует учитывать, что детекция делеций и инверсий с помощью ПЦР недостаточна для полноценного генотипирования, поскольку она не позволяет надёжно выявлять все возможные структурные варианты. Кроме того, при редактировании зигот и ранних эмбрионов вмешательство может происходить на фоне репликации ДНК и первых клеточных делений, что повышает вероятность мозаицизма и сложных перестроек.

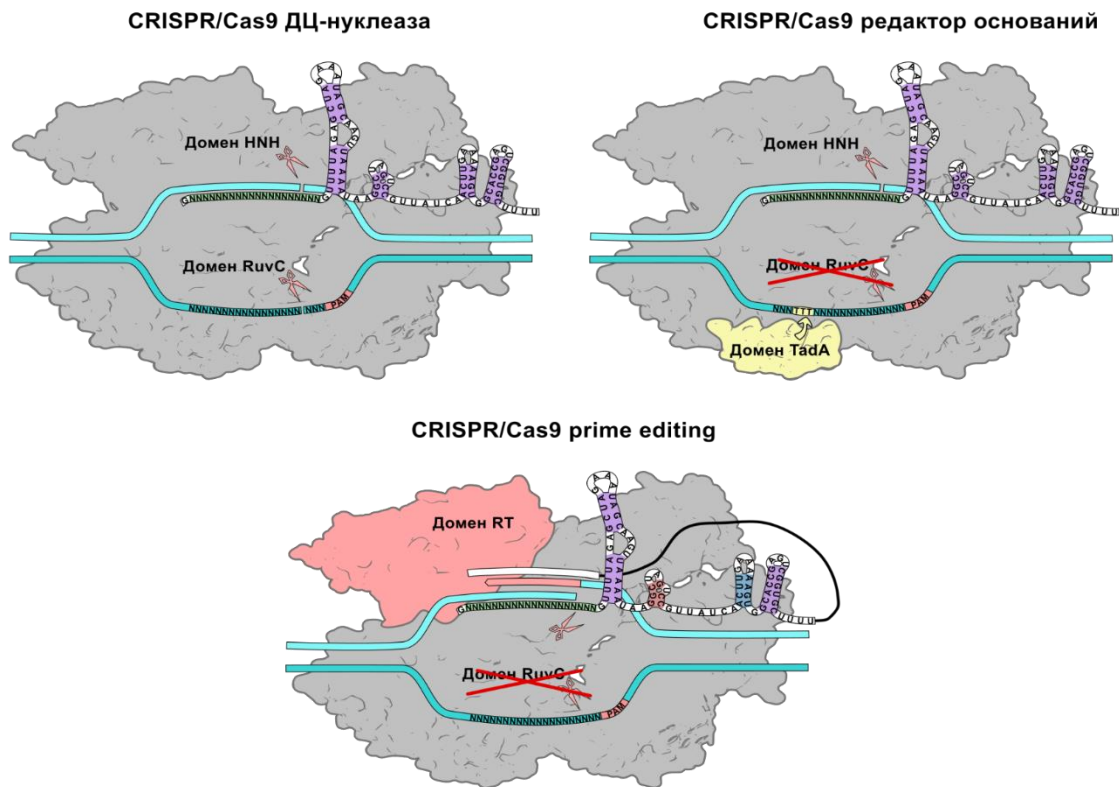


Рисунок 2. Сравнение компонентов систем классической CRISPR/Cas9-нуклеазы, редакторов оснований и prime editing. В классической системе CRISPR/Cas9 нуклеаза Cas9 распознает целевой участок ДНК за счет комплементарного взаимодействия направляющей РНК с протоспейсером и вносит DSB при участии нуклеазных доменов HNH и RuvC. В системах редактирования оснований один из нуклеазных доменов Cas9 инактивирован, а к Cas9 присоединен ферментативный домен дезаминазы, что позволяет вносить отдельные нуклеотидные замены без образования двуцепочечного разрыва. В системе prime editing Cas9-никаза слита с доменом обратной транскриптазы; специальная направляющая РНК одновременно задает целевой участок и служит матрицей для внесения заданного изменения в ДНК. Адаптировано из [125].

Внесение DSB само по себе не обеспечивает получения требуемого генетического изменения. Вообще говоря, исход редактирования определяется конкуренцией нескольких путей репарации. NHEJ (non-homologous end joining) напрямую соединяет тупые концы двуцепочечной ДНК без какой-либо проверки на соответствие получившейся

последовательности исходной. Репарация по этому пути часто приводит к небольшим инсерциям или делециям в месте разрыва, что используется для нокаута генов [126]. При индукции двух или более DSB, их концы могут соединиться в любой комбинации, формируя протяженные делеции, инверсии или хромосомные транслокации [119]. MMEJ (microhomology-mediated end joining, или alt-EJ) требует коротких участков микрогомологии (длиной всего лишь в несколько нуклеотидов) по краям разрыва и обычно приводит к предсказуемым делециям между такими участками, что тоже может быть использовано для нокаута генов или, как будет показано в в этой работе, для удаления функциональных элементов генома [127]. HDR (homology-directed repair) является самым точным путем репарации, поскольку использует гомологичную матрицу в виде обычно сестринской хроматиды для восстановления исходной последовательности. В качестве матрицы может быть использован экзогенный донорный ДНК-шаблон, что позволяет вносить точные заранее спроектированные изменения в последовательность геномной ДНК [128].

Эти три пути репарации конкурируют между собой после внесения DSB системой CRISPR/Cas9 в клетках млекопитающих, и исход этой конкуренции, зависящий от клеточного типа, фазы клеточного цикла, локуса, хроматинового состояния, способа доставки компонентов системы CRISPR/Cas и прочих факторов, определяет исход генетического редактирования [129]. Существуют технологические методы повлиять на эту конкуренцию и сместить преимущество в сторону нужного пути репарации. Для повышения эффективности HDR используют донорные матрицы с гомологичными плечами, выбирают сайты разреза как можно ближе к редактируемой позиции и по возможности обогащают клетки в S/G2-фазах клеточного цикла. Также применяют фармакологическое или генетическое подавление NHEJ, например через ингибирование DNA-ПКcs, и подавление MMEJ через ингибирование Polθ/POLQ или PARP1. Напротив, если требуется получить делеции без

донорной матрицы, можно использовать предсказуемые профили NHEJ/ММЕJ, подбирая sgRNA с учетом сайтов микрогомологии вокруг разрыва. Однако такие меры не делают редактирование полностью контролируемым, поскольку разные пути репарации динамично конкурируют между собой, и вклад какого-либо из них невозможно свести к нулю.

Нежелательные *on-target*-эффекты возникают именно из-за неполной предсказуемости репарации. NHEJ и ММЕJ могут приводить не только к небольшим INDEL-мутациям, но и крупным делециям, инверсиям, дупликациям и транслокациям, что следует обязательно учитывать при разработке стратегии генотипирования. HDR, несмотря на большую точность, также может сопровождаться неполным копированием донорной матрицы, случайной интеграцией донорной ДНК, а также продуктами репарации по другим путям. Кроме того, искусственное подавление отдельных путей репарации может повышать эффективность желаемого исхода, но одновременно увеличивать токсичность и клеточный стресс, что увеличивает риск отбора клеток с нарушенными механизмами контроля геномной стабильности.

В связи с этим в настоящее время развивается ряд более точных подходов, основанных на технологии CRISPR/Cas9 и направленных на преодоление ограничений исходной системы. Одним из них являются редакторы оснований (Рис. 2), позволяющие вносить замены А→G или С→Т без индукции двуцепочечных разрывов ДНК [130]. Высокая точность редактирования открывает возможности для моделирования заболеваний, обусловленных точечными мутациями [131,132], кроме того, редакторы оснований уже активно применяются при разработке генно-терапевтических стратегий для заболеваний, связанных с вариантами в ядерной и митохондриальной ДНК [133,134].

Система редактирования оснований использует химерный белок, состоящий из Cas9 с нарушенной нуклеазной активностью и нуклеотид-дезаминазы. Последовательность белка Cas9 в составе редактора оснований модифицирована для инактивации одного из нуклеазных доменов (RuvC). Полученный белок сохраняет способность распознавать и связывать участок ДНК, комплементарный спейсеру хнРНК, но вносит только одноцепочечный разрыв ДНК. В общем случае, слияние такого белка (или белка с инактивированными обоими нуклеазными доменами) с различными функциональными доменами позволяет адресно реализовывать функцию данного домена в выбранном локусе генома [135], специфичность к которому определяется направляющей CRISPR РНК, при этом избегая негативных побочных эффектов внесения DSB. Такие системы могут использоваться не только для геномного редактирования, но и для редактирования эпигеномного, например, для активации экспрессии генов [136], их репрессии, направленного изменения гистоновых модификаций и моделирования искусственных топологических структур хроматина.

В случае с редакторами оснований Cas9-никаза слита с дезаминазами нуклеотидов, которых существует два типа, принимающих в качестве субстрата либо цитозин, либо аденин. В случае аденинового редактора оснований (adenine base editor, ABE) дезаминирование происходит в узком окне активности дезаминазы (Рис. 2), а одноцепочечный разрыв на противоположной цепи стимулирует репарацию с использованием модифицированной цепи в качестве матрицы [137]. ABE индуцирует конверсию А→G за счет дезаминирования аденина с образованием гипоксантина, который затем распознается ДНК-полимеразой как гуанин [138]. Таким образом, редакторы оснований позволяют изменять нуклеотидную последовательность без DSB и без доставки экзогенной донорной ДНК, что снижает риск *on-target*-эффектов и степень опасности *off-target*-эффектов.

Несмотря на эти преимущества, применение редакторов оснований имеет ряд ограничений. Ключевым из них является каталитическая специфичность дезаминаз, позволяющая выполнять исключительно транзиции, то есть замены С→Т или А→G, без смены химического класса азотистого основания. Это существенно сужает спектр клинического и фундаментального применения редакторов оснований [137,138]. Другой важной проблемой остается нецелевая активность, проявляющаяся как в редактировании off-target сайтов, так и неспецифических соседних оснований внутри целевого сайта. Кроме того, дезаминазы, особенно цитозиновые, могут случайно дезаминировать цитозины в РНК, потенциально нарушая транскриптом и вызывая токсические эффекты; такая активность часто не зависит от направляющей РНК, а только от присутствия эффекторного белка [139].

Вершиной развития CRISPR-технологий на данный момент является prime editing — метод высокоточного редактирования генома, впервые описанный в 2019 году [140]. Подобно редакторам оснований, эта система основана на Cas9-никазе; однако в данном случае Cas9-никаза слита с обратной транскриптазой и функционирует совместно со специальной направляющей РНК, называемой pegRNA (prime editing guide RNA). pegRNA одновременно определяет целевой участок генома, как и в классической системе CRISPR/Cas9, и несет матрицу для внесения заданной генетической модификации (Рис. 2).

После связывания комплекса с ДНК Cas9-никаза вносит одноцепочечный разрыв, а обратная транскриптаза синтезирует новую последовательность, используя участок pegRNA в качестве матрицы, а свободный 3'-конец поврежденной цепи — в качестве затравки. Именно с этим связано название метода: prime указывает на праймирование синтеза ДНК. Затем в процесс вовлекаются клеточные системы репарации ДНК, которые могут включить синтезированную последовательность в состав генома. Таким образом, prime editing позволяет вносить все типы однонуклеотидных замен, а также

небольшие инсерции и делеции без образования DSB и сопутствующих рисков.

Ограничения prime editing связаны прежде всего с тем, что изначально этот метод был разработан для высокоточного внесения точечных изменений, таких как нуклеотидные замены и короткие делеции и инсерции. Поэтому он хуже подходит для создания протяженных делеций и удаления функциональных элементов генома. Было показано, что комбинация prime editing с внесением дополнительного одноцепочечного разрыва в редактируемую цепь позволяет получать делеции длиной до 80 п. н. и может использоваться для удаления функциональных сайтов генома с эффективностью до 60%. Однако по трудоемкости такой подход существенно превосходит классический CRISPR/Cas9, основанный на внесении двух DSB.

Производные от prime editing методы, такие как PRIME-Del [141] и PEDAR [142], позволяют высокоточно получать делеции длиной до 10 тыс. п. н. Тем не менее эффективность редактирования при использовании этих подходов в типичном случае оказывается порядка 1%. Хотя такие методы потенциально более безопасны с точки зрения риска крупных перестроек, транслокаций и других последствий DSB, что открывает возможности для их терапевтического применения, технологическая зрелость prime editing пока уступает классическим CRISPR/Cas9-подходам. Интересно, что один препарат на основе prime editing уже перешел к ранним клиническим испытаниям: геннотерапевтический продукт PM359 для лечения р47phox-дефицитной хронической гранулематозной болезни [143].

С учетом сравнительно низкой эффективности этих подходов, а также менее строгих требований к безопасности при создании модельных линий мышей, для которых нежелательные аллели могут быть удалены путем возвратных скрещиваний, в настоящей работе используются более

конвенциональные методы. Эти методы основаны на применении CRISPR/Cas9-нуклеазы, внесении DSB и последующей репарации ДНК.

### **Влияние пространственной архитектуры на регуляцию экспрессии генов**

Считается, что пространственная организация хроматина контролирует пространственные взаимодействия функциональных элементов генома, таких как промоторы и энхансеры [3,82]. В пределах одного ТАДа энхансеры с большей вероятностью устанавливают пространственные контакты с промоторами потенциальных генов-мишеней. Границы ТАДов, часто ассоциированные с сайтами связывания CTCF и когезином, ограничивают распространение регуляторных взаимодействий между соседними доменами. Нарушение границы ТАДа или изменение расположения сайтов CTCF может приводить к эктопическим контактам между энхансерами и промоторами, потере нормальных регуляторных взаимодействий или изменению уровня экспрессии генов. (Рис. 3). Для исследования точных механизмов и степени участия топологии хроматина в генетической регуляции, распространен подход обратной генетики, заключающийся в генно-инженерном удалении архитектурных факторов (в первую очередь, сайтов связывания CTCF). Нарушение расположения и/или ориентации CTCF-сайтов может приводить к локальной реорганизации топологии хроматина и, в ряде случаев, нарушению регуляции генной экспрессии и развитию наследственных заболеваний. В настоящее время произведено большое количество работ в этом русле, на ряде примеров показавших связь пространственной организации генома и регуляции генетических процессов, однако возможность предсказания последствий нарушения топологии хроматина до сих пор остается ограниченной.

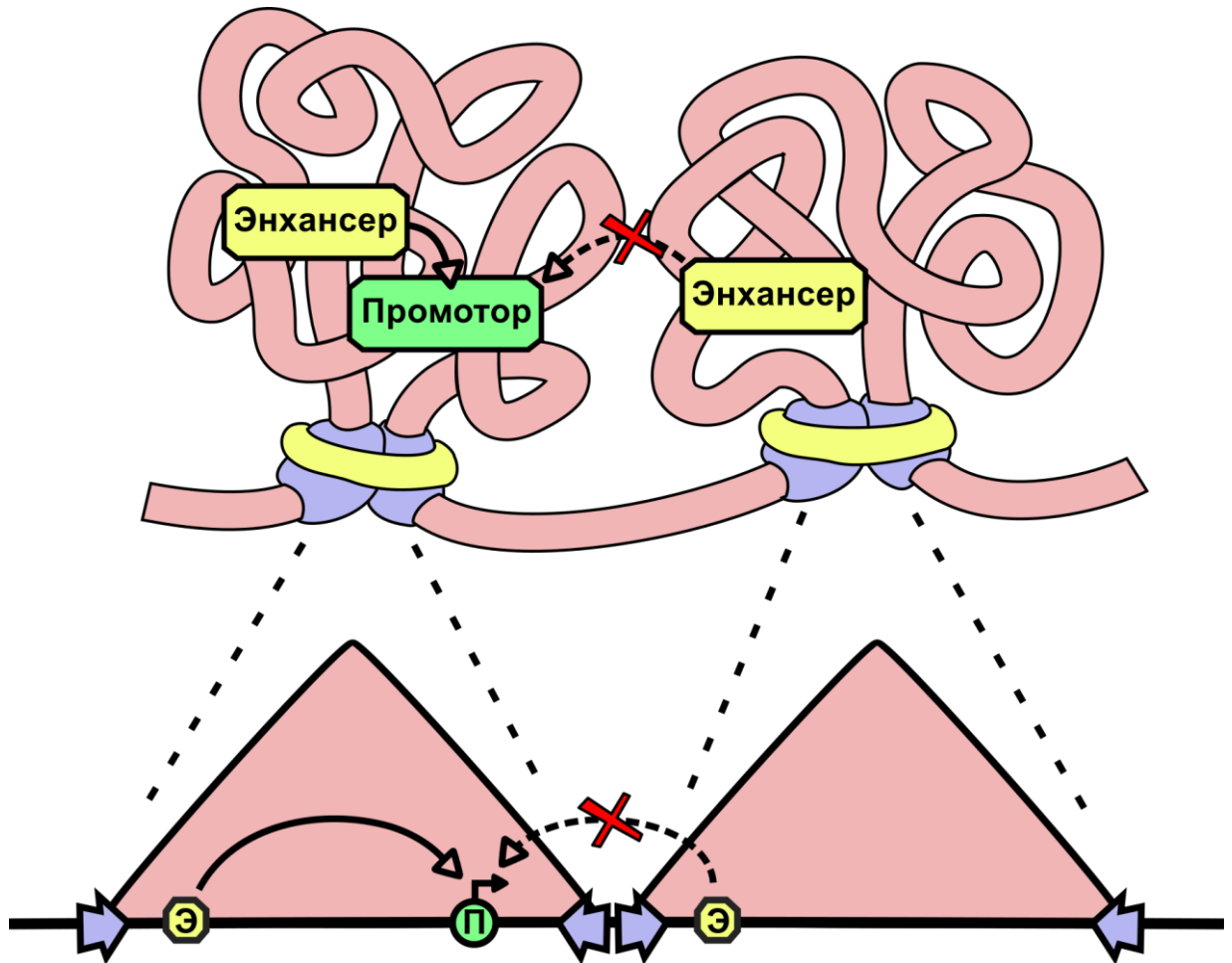


Рисунок 3. Схематическое представление влияния границ ТАДов на энхансер-промоторные взаимодействия. Сверху показана пространственная организация хроматина в виде ТАДов, сформированных петлями хроматина при участии когезина и сайтов связывания CTCF. Энхансеры и промоторы преимущественно взаимодействуют в пределах одного ТАДа, тогда как граница ТАДа ограничивает контакты между регуляторными элементами соседних доменов. В нижней части представлена условная проекция той же организации в геномных координатах: ТАДы отображены как области повышенной частоты внутренних контактов, аналогичные ярко окрашенным треугольным структурам на Hi-C-картах, возвышающиеся над осью геномных координат.

Первые работы в этой области были осуществлены при изучении существующих генетических заболеваний человека, патогенез которых не объяснялся напрямую из ассоциированных с патологией генетических вариантов. Такие варианты характеризуются тем, что мутации не нарушают кодирующие участки генов, потенциально вовлеченных в развитие патологии, но вместо этого затрагивают некодирующие элементы генома, такие как сайты связывания CTCF, или значимо меняют геномное расстояние между регуляторными элементами. Так, в первой статье [14], касающейся этого вопроса, рассматривался ряд хромосомных перестроек локуса *WNT6/INH/EPHA4/PAX3* хромосомы 2 человека, приводящих к таким нарушениям развития конечности, как брахидактилия, полидактилия и F-синдром, при котором происходит сращивание костей запястья и предплюсны, а также нарушения развития пальцев рук и ног. Этот локус содержит 3 ТАД, один из которых содержит гены *WNT6* и *INH*, другой — ген *EPHA4* с комплексом его энхансеров, а третий — ген *PAX3*. Инверсии, делеции и дупликации, затрагивающие граничные элементы между ними, приводят к развитию патологий конечностей (Рис. 4). Точно к таким же патологиям приводят и смоделированные на синтетическом локусе мыши с помощью системы CRISPR/Cas9 аналогичные хромосомные перестройки. Однако другие перестройки того же масштаба, но не затрагивающие граничные элементы, не имеют фенотипического проявления. С помощью технологии 4C было показано, что затрагивающие границы ТАДов перестройки приводят к перераспределению пространственных контактов между энхансером *EPHA4* и другими генами локуса. Эктопические взаимодействия регуляторных элементов приводят к тому, что гены *WNT6*, *INH* и *PAX3* начинают проявлять паттерн экспрессии, характерный для гена *EPHA4* в развивающихся конечностях, что и приводит к нарушениям.

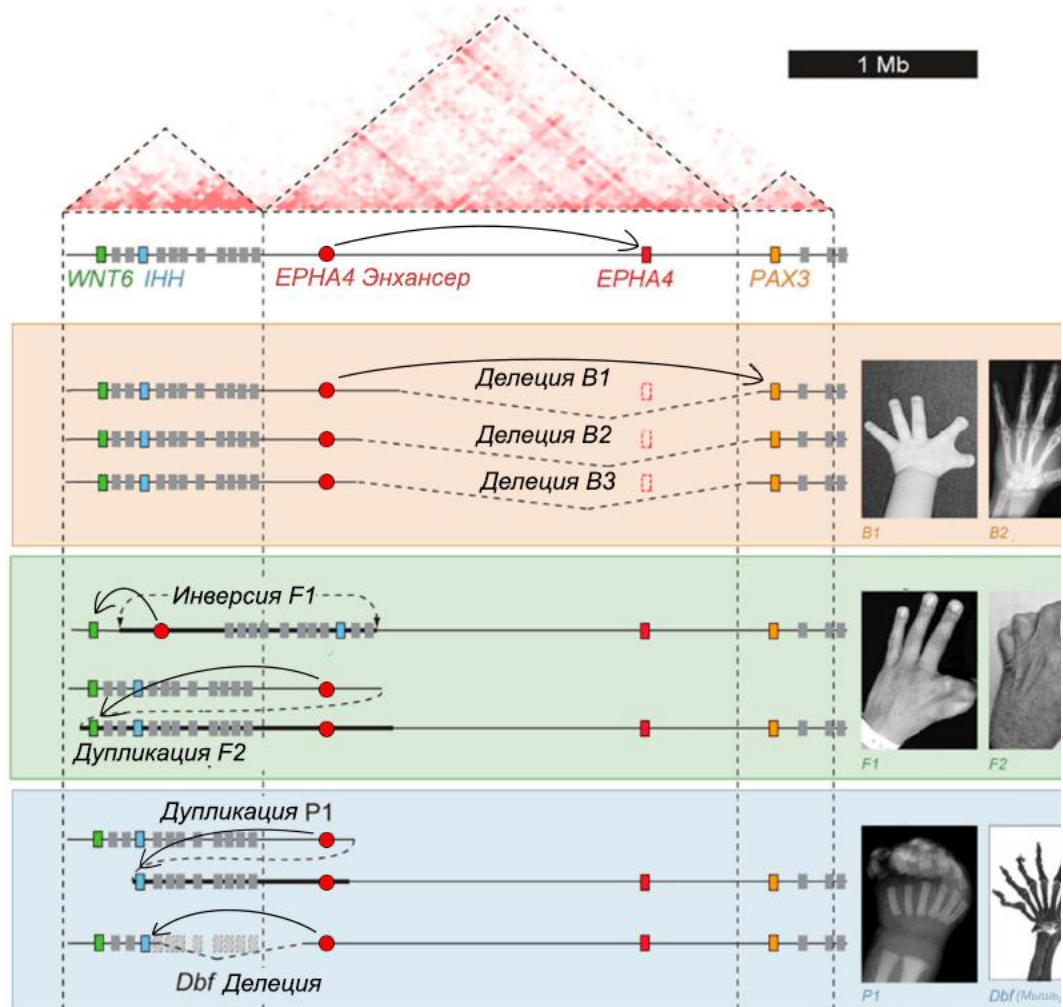


Рисунок 4. Схема исследованных aberrаций в локусе *WNT6/INH/ERNA4/PAX3*. Наверху показана карта пространственных контактов локуса в норме. Выделяются три ТАДа. Ниже — перестройки, приводящие к фенотипам: брахидактилии (случаи B1, B2, B3), F-синдрому (F1, F2), полидактилии (P1, Dbf). Стрелка показывает взаимодействие между энхансером *ERNA4* в норме и в патологии. Из [14], с изменениями.

Следующий показательный пример связан с локусом *KCNJ2/SOX9* [15], разделенным на два ТАДа, один из которых содержит гены *KCNJ2* и *KCNJ16*, а второй – *SOX9* и регулирующие его активность энхансеры. Ген *SOX9* является регулятором развития пола: экспрессия гена *SRY* через сигнальные каскады

запускает экспрессию *SOX9* и запускает развитие гонад по мужскому типу. Известны дупликации в локусе *KCNJ2/SOX9*, которые приводят к переключению программы развития гонад мужского типа у хромосомных женщин. Однако более протяженные дупликации, включающие ген *KCNJ2*, не приводят к этому фенотипическому проявлению, но взамен ассоциированы с синдромом Кука, который связан с патологиями развития конечностей (Рис. 5А). Исследователи воспроизвели ряд подобных дупликаций различного размера в синтенном локусе мышей. Оказалось, что небольшие дупликации в энхансерной области гена *Sox9* все так же приводят к патологии развития пола, но как только дупликация увеличивается настолько, что начинает включать в себя границу ТАДов этого локуса, данный фенотип пропадает. Более того, такая перестройка вообще не имеет никакого фенотипического проявления. Однако, как только в дупликацию включается еще и ген *Kcnj2*, начинают проявляться патологии в

развитии

конечностей.

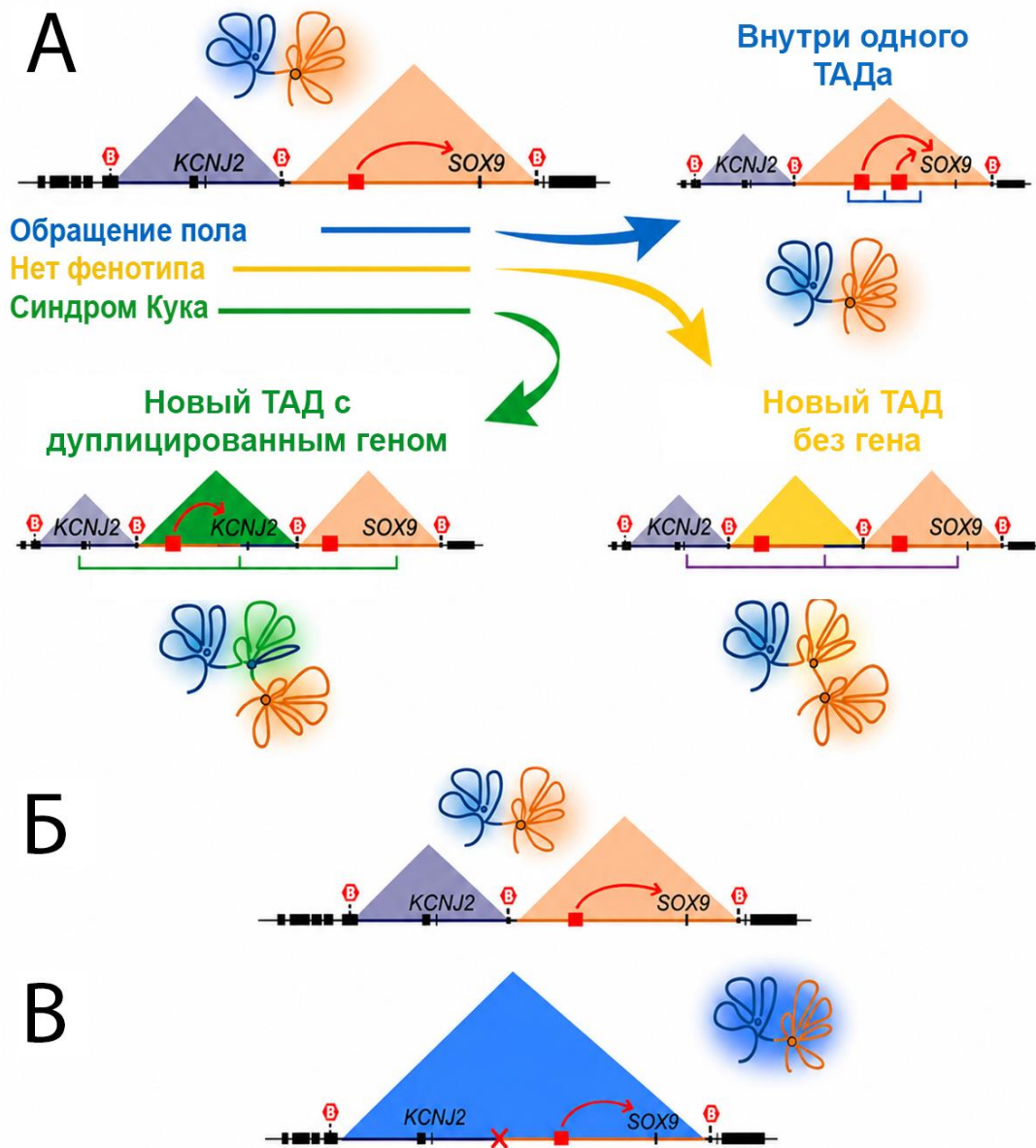


Рисунок 5. А — Карта топологических доменов и положение генов в локусе *KCNJ2/SOX9*. Показаны границы известных у человека дупликаций, приводящих к развитию синдрома Кука, обращению пола, или не имеющие фенотипического проявления. Б — Карта топологических доменов и положение генов в локусе *KCNJ2/SOX9* дикого типа. В — Карта топологических доменов и положение генов в локусе *KCNJ2/SOX9* при делеции сайтов связывания CTCF в границе ТАДов. Несмотря на слияние

топологических доменов, ландшафт промотор-энхансерных взаимодействий в конечностях не был критически нарушен. Из Franke et al., 2016 с изменениями.

Это объясняется тем, что дупликация, включающая границу ТАДа, приводит к возникновению нового ТАДа между удвоенными границами. Вторая копия энхансера *Sox9*, в коротком варианте дупликации взаимодействовавшая с промотором этого гена и увеличивавшая его активность, при попадании в новый ТАД экранируется и не ведет к aberrантной экспрессии. Но, как только в новый ТАД включается еще и ген *Kcnj2*, он попадает под влияние энхансера *Sox9*, что приводит к его ectопической экспрессии и развитию патологий конечности. Исходя из этих данных можно было бы предположить, что факта сосуществования *Kcnj2* с энхансером *Sox9* в одном ТАДе достаточно для развития патологического фенотипа. Однако делеция границы ТАДов между этими двумя элементами не только не привела к какому-либо фенотипическому проявлению, но даже перераспределение пространственных контактов оказалось почти незаметным (Рис. 5Б,В). В то время как экспрессия гена *Sox9* в эмбриональных зачатках конечностей на фоне делеции осталась неизменной, экспрессия *Kcnj2* увеличилась в 2,5 раза. Интересно, что в случае дупликации, приводящей к синдрому Кука, экспрессия *Sox9* также не изменилась, но экспрессия *Kcnj2* увеличилась всего в 3 раза.

Этот интригующий результат подводит нас к следующей работе, опубликованной той же группой три года спустя на том же локусе *Kcnj2/Sox9* (88). Исследователи выдвинули гипотезу, что устойчивость ТАДов к удалению границы, показанная в предыдущем исследовании, объясняется наличием дополнительных сайтов связывания CTCF внутри ТАДов. И действительно, при последовательном удалении сайтов связывания CTCF, количество Hi-C контактов между ТАДами постепенно увеличивалось, вплоть до практически полного их слияния. Что касается экспрессии вовлеченных

генов, то постепенное слияние ТАДов возымело на них незначительное влияние. Так, экспрессия *Sox9* в эмбриональных почках конечностей осталась неизменной на всех ступенях постепенного удаления СТСФ, экспрессия же *Kcnj2* при первом удалении СТСФ сайтов увеличилась лишь в 2,5 раза, и затем оставалась на том же уровне.

Далее, исследователями были получены мутации, меняющие направления сайтов связывания СТСФ, участвующих в образовании пространственной организации локуса. Перенос границы ТАДов в пространство между энхансером и геном *Sox9* привел к реорганизации пространственных контактов и нарушению регуляции гена *Sox9*, результатом чего является перинатальная летальность. К подобному же фенотипу привела и инверсия, затрагивающая границу ТАДов и энхансер *Sox9*. Таким образом, как только ген *Sox9* оказывается пространственно разделенным со своим энхансером, его функции драматически нарушаются. Следует отметить, что на фоне упомянутой выше инверсии экспрессия *Kcnj2* также претерпевает серьезные нарушения. Так, паттерн его экспрессии в зачатках конечностей начинает в точности повторять паттерн *Sox9*, а уровень экспрессии возрастает более чем в 5 раз. Интересно, что при практически такой же инверсии, но с сохранением границы ТАДов на нативном месте (иными словами, сближение энхансера *Sox9* и гена *Kcnj2* с сохранением инсуляции между ними), ни фенотипических проявлений, ни изменений в уровне или паттерне экспрессии обоих генов замечено не было.

Помимо этого, исследователи получили еще один тип мутации, который представлял собой комбинацию инверсии с удалением границеобразующих сайтов связывания СТСФ. При этом никакого изменения в функционировании гена *Sox9* замечено не было, зато паттерн экспрессии *Kcnj2* начинает полностью повторять таковой для *Sox9*, а уровень экспрессии возрастает в 2,5 раза, что приводит к развитию уже известного нам фенотипа – синдрома Кука.

Этот ряд статей позволяет сделать нам следующие глобальные выводы. Для развития абнормального фенотипа важен не столько уровень экспрессии, сколько ее паттерн. Для формирования паттерна решающую роль играют два фактора: положение инсуляторных границ и расстояние между энхансерами и промоторами. Кроме того, неправильное положение границы играет куда большую роль, чем ее исчезновение. Судя по всему, границы ТАДов играют лишь вспомогательную роль в организации промотор-энхансерных взаимодействий. Выбор же партнера для энхансера определяется паритетом следующих факторов: энхансер-промоторной специфичности, расстояния между регуляторными элементами, и расположения инсуляторных границ. Роль последних, судя по всему, в нормальных условиях оказывается наименьшей; это доказывает тот факт, что при исчезновении всех ТАДов и хроматиновых петель в масштабах всего генома экспрессия генов либо меняется незначительно, либо не меняется вовсе [19,20]. Однако, не следует отбрасывать вероятность того, что существует другая функция пространственной организации хроматина, которая является ведущей по отношению к организации промотор-энхансерных взаимодействий (например, разграничение репликационных доменов). Возможно даже, что в эволюционном процессе не расположение инсуляторных границ определяет пространственные взаимодействия между промотором и энхансером, а наоборот, промотор-энхансерные взаимодействия определяют расположение границ ТАДов.

В противовес этой концепции встает ряд работ, показывающих для определенных локусов первичную роль пространственной организации в регуляции генной экспрессии. Например, гомеозисный ген *Pitx1* находится под контролем энхансера *Pen*, инициирующего его экспрессию. Экспрессия *Pitx1* приводит к переключению пути развития конечностей с передних на задние. При этом энхансер *Pen* находится в активном состоянии в зачатках обеих пар конечностей, а определяющее значение для его энхансерной

активности имеет пространственный контакт между ним и промотором гена *Pitx1*. Пространственная архитектура этого локуса различается в зачатках передних и задних конечностей таким образом, что в задних *Pen* и *Pitx1* могут пространственно взаимодействовать, а в передних они разделены. Формирование различий в распределении пространственных контактов определяется активностью других генов, вследствие чего мутации, затрагивающие эти гены, а также перестройки, критически сближающие *Pen* и *Pitx1*, могут приводить к патологиям развития передних конечностей [24].

Экспансия микросателлитов, расположенных в границах ТАДов и обогащенных CpG-островками, может разрушать границу между ТАДами за счет изменения статуса метилирования локуса и утраты связанных CTCF. Такие локальные изменения в пространственной организации уже сами по себе могут приводить к нарушению регуляции генной экспрессии и развитию патологий, что продемонстрировано на гене *FMRI* [144] при синдроме ломкой X-хромосомы. Кроме того, мутации в границах ТАДов, приводящие к изменению или удалению границеобразующих сайтов связывания CTCF, а также хромосомные перестройки, приводящие к разрушению или образованию новых ТАДов, могут приводить к изменениям экспрессии генов и онкогенной трансформации клеток [145]. Анализ хромосомных перестроек, приводящих к генетическим патологиям и распространенных в человеческой популяции, показал, что до 11,8% таких случаев можно объяснить изменением пространственной структуры хроматина и, как следствие, нарушением промотор-энхансерных взаимодействий [9].

С другой стороны, большое количество данных указывает на стабильность пространственной организации по отношению к удалению границ ТАДов. Помимо уже описанного нами слияния ТАДов в локусе *Kcnj2/Sox9*, примером этого является делеция локуса *Firre* мышинового генома, содержащего ген длинной некодирующей РНК *Firre* и 15 сайтов связывания CTCF. Этот локус формирует сильную инсуляторную границу, которая

сохранилась даже при внесении крупномасштабной делеции этого локуса [146]. Таким образом, удаление границеобразующих CTCF сайтов, а также и локуса с активной транскрипцией, оказалось недостаточно для реорганизации пространственной архитектуры хроматина. Вероятно, это объясняется тем, что сайты связывания CTCF, находящиеся ранее внутри ТАДов и не принимавшие участия в образовании хроматиновых петель, при удалении граничного элемента принимают на себя функцию организации границы ТАДов.

Таким образом, в настоящее время вопрос о роли ТАДов в регуляции генной экспрессии остается нерешенным. Концепция об инструктивной и необязательной роли инсуляторных границ входит в противоборство с данными об устойчивости ТАДов к потере граничных элементов, а также о консервативности их границ в ходе эволюции. Несмотря на это, из рассмотренных случаев становится ясно, что пространственная архитектура является важным регулятором промотор-энхансерных взаимодействий, наряду с их специфичностью и геномным расстоянием между регуляторными элементами. Ключевой вопрос состоит в том, какой из этих факторов является ведущим в процессах развития и жизнеобеспечения организмов. Мы полагаем, что степень влияния ТАДов на генную экспрессию является специфичной для разных локусов, и может варьировать среди разных типов клеток, степеней их дифференцировки и от условий среды.

### **Модельные локусы *Slc29a3/Unc5b* и *Epb41l4a***

Для экспериментального исследования связи пространственной организации хроматина и регуляции экспрессии генов нами были выбраны два модельных локуса мыши, а именно *Slc29a3/Unc5b* и *Epb41l4a*. Принципиальное отличие использованной нами стратегии выбора модельных локусов от подходов, применявшихся в большинстве предыдущих исследований, заключается в том, что локус был выбран не на основании уже

известных данных о связи мутаций в данном регионе с патологией человека, а исходя из свойств самой границы ТАДов, предполагающих ее потенциальную функциональную значимость. Ключевыми критериями выбора служили высокая сила инсуляции, межвидовая консервативность границы, тканеспецифические особенности активности генов и их регуляторных элементов, а также возможностью моделирования нарушений локальной архитектуры генома *in vivo*. Более подробно стратегия выбора локусов описана в разделе Результаты.

Два выбранных локуса позволяют рассмотреть проблему с разных сторон. Локус *Slc29a3/Unc5b* представляет интерес как участок, содержащий гены с широким профилем экспрессии, разделенные границей ТАДов и потенциально зависящие от тканеспецифического регуляторного контекста. Локус *Erb41l4a*, напротив, содержит одну из наиболее сильных и консервативных границ ТАДов, разделяющую регуляторные ландшафты клинически значимых генов *Nrep* и *Apc*. Совместный анализ этих моделей позволяет оценить, насколько последствия нарушения топологической организации хроматина зависят от структуры локуса, функции генов и эпигенетического состояния ткани.

Локус *Slc29a3/Unc5b* был выбран как модель для анализа последствий нарушения CTCF-зависимой архитектуры в регионе, содержащем несколько функционально значимых генов с широким, но неодинаковым тканевым профилем экспрессии. В этом локусе граница ТАДов разделяет регуляторные ландшафты генов *Slc29a3/Cdh23* и *Unc5b*, при этом в области первого интрона *Unc5b* в ряде тканей обнаруживаются выраженные энхансерные сигнатуры, что указывает на возможность тканеспецифической регуляции этого гена.

Ближайший к выбранной границе ген, *Unc5b*, кодирующий мембранный рецептор нетрина-1 [147], участвует в руководстве роста аксонов [148], подавлении ангиогенеза [149], остеогенезе [150], пролиферации и миграции

клеток [151–153], а также регуляции апоптоза [154,155]. Он играет ключевую роль в формировании сосудистой системы эмбриональной части плаценты [156], регенерации эпителия почечных трубочек [147,153] и пролиферации клеток Шванна [152]. По транскриптомным данным, наибольшая активность этого гена отмечается в разных частях кишечника, а также в мозжечке и сердце. В этих органах обнаруживаются обширные энхансерные сигнатуры (H3K27ac и H3K4me1) в области первого интрона *Unc5b*, что может говорить о существовании суперэнхансера, контролирующего экспрессию этого гена и активного лишь в некоторых органах (Рис. 8Г).

*Slc29a3 (Ent3)*, ген переносчика нуклеозидов, отличается от своих генных родственников локализацией во внутренних мембранах клетки. Он транспортирует нуклеозиды, некоторые нуклеотиды и лекарства, используемые в лечении рака и ретровирусных инфекций [157,158]. *Slc29a3* имеет важную роль в пролиферации Т-лимфоцитов [159], поддержании гомеостаза стволовых клеток и процессе аутофагии [160,161]. Наиболее высокие уровни экспрессии *Slc29a3* регистрируются в семенниках и почках, при проведенный нами анализ экспрессии генов показал, что существует корреляция между экспрессией генов *Slc29a3* и его ближайшего соседа по ТАДу гена *Cdh23*, и обратная корреляция между *Slc29a3* и *Unc5b*, разделенными границей ТАДов (Рис. 6).

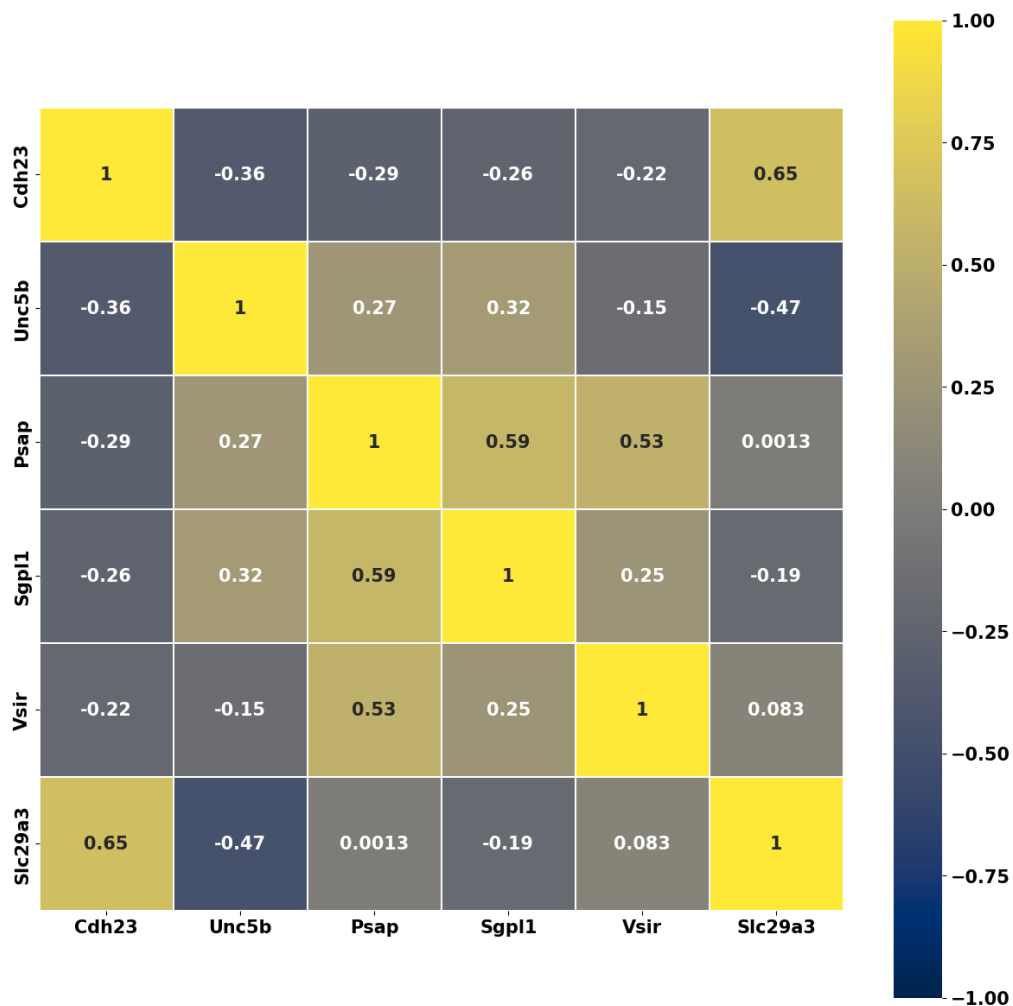


Рисунок 6. Матрица корреляции паттернов экспрессии генов локуса *Slc29a3/Unc5b*. Результаты получены на основе транскриптомных данных ENCODE для органов и тканей мыши

Анализ литературы показывает важность этих генов в медицинской генетике. Нарушение регуляции экспрессии *Unc5b* связано с некоторыми раковыми заболеваниями [162,163], в то время как мутации в *Slc29a3* связаны с наследственными заболеваниями, симптомы которых включают склеродермию, гиперпигментацию и гипертрихоз, гипертрофию внутренних органов, сердечно-сосудистые и костно-мышечные деформации, некоторые формы диабета, гистиоцитоз [164]; а также аутоиммунные и иммунодефицитные состояния [165,166].

Согласно базе OMIM, ген *Cdh23* ассоциирован с проявлениями как синдромальной, так и асиндромальной генетической глухоты [167,168], в то

время как ген *Psap* связан с синдромом дефицита сапозина, проявляющегося в виде гепатоспленомегалии и атрофии мозговых структур [169,170], и повышенным риском развития паркинсонизма [171]. Данные MGI о фенотипах нокаутных моделей следующие:

1. *Unc5b* – аномалии в развитии нейропора, арест развития эмбриона, смерть на этапе эмбриогенеза.
2. *Slc29a3* – гиперплазия селезенки, дисфункция гематопоеза, преждевременная смерть.
3. *Cdh23* – наследственная глухота, патологии в развитии Кортиева органа, атаксия, гиперактивность, вращательное поведение, невозможность плавания.
4. *Psap* – гипоактивность, атаксия, тряска головы, глухота, демиелинизация, аксональная дегенерация, недостаточная масса тела, преждевременная смерть.

Из изложенного выше следует, что гены изучаемого нами локуса критически важны для развития и функционирования организма. Мы предполагаем, что специфический характер пространственных контактов хроматина также играет существенную роль в регуляции этих генов, что объясняет их консервативность между геномами мыши и человека. Изучаемый нами локус не содержит гомеозисных генов, которые были в центре внимания большинства исследований влияния нарушений пространственной организации хроматина на экспрессию генов. Напротив, многие гены локуса имеют широкий спектр экспрессии среди тканей, и, гипотетически, изменения в механизмах цис-регуляции их экспрессии могут сказываться на поведении сразу нескольких генов локуса. Таким образом, данный локус представляет собой перспективную модель для исследования пространственной архитектуры генома.

Вторым модельным объектом был выбран локус *Epb41l4a*, в котором интерес представляет чрезвычайно сильная и консервативная граница ТАДов, расположенная в терминальной области гена *Epb41l4a* и разделяющая регуляторные ландшафты клинически значимых генов *Nrep* и *Apс*. Эта граница относится к числу наиболее сильных границ по уровню инсуляции, консервативна между человеком и мышью и формируется массивным кластером сайтов связывания CTCF.

Локус *EPB41L4A* содержит три гена, представляющих интерес (собственно *EPB41L4A*, *NREP* и *APC*). Ген *EPB41L4A*, также известный как *NBLA*, относится к суперсемейству FERM и кодирует трансмембранный белок, участвующий в регуляции формирования цитоскелета, внутриклеточного транспорта и трансдукции сигнала по пути WNT/ $\beta$ -катенина [172]. Уровни экспрессии кодирующих изоформ и длинных некодирующих РНК транскриптов *EPB41L4A*, такие как *EPB41L4A-AS1* и *EPB41L4A-AS2*, имеют прогностическую значимость в клинической онкологии и ассоциированы с повышенными уровнями пролиферативной активности и миграции раковых клеток [173–175]. Согласно данным IMPC, мыши с нокаутом этого гена жизнеспособны, но у них наблюдаются аномалии развития почек, снижение уровня циркулирующей глюкозы, увеличение веса сердца и снижение силы хватки передних лап. Однако поведенческие аномалии у мышей с нокаутом *Epb41l4a* не выявлены.

Ген *NREP* (neuronal regeneration related protein; белок, связанный с регенерацией нейронов; также известный как P311) кодирует внутриклеточный цитоплазматический белок массой 8 кДа. Этот белок высоко консервативен среди видов и обычно экспрессируется в нервной системе (144), особенно высокий уровень транскрипционной активности наблюдается в мозжечке. Несмотря на отсутствие у *NREP* характерных мотивов, указывающих на его функции, он содержит три последовательности PEST, которые являются сигналом быстрой деградации убиквитин-протеасомной

системой и свидетельствуют о коротком времени существования белка в клетке [176–178]. Известно, что *NREP* усиливает пролиферацию и миграцию клеток во время васкуляризации и заживления ран [179–181], а также участвует в развитии различных типов рака [182–184]. В нервной системе *NREP* регулирует рост нейритов посредством модуляции активности Rho-киназы. Нокаут *Nrep* у мышей не вызывает явных нарушений развития, однако при этом наблюдаются поведенческие аномалии, дефициты в обучении и памяти [185], а также нарушения восприятия боли [186]. Кроме того, интересно, что мРНК *NREP* содержит мотив связывания микроРНК miR-29b, напрямую осуществляющий деградацию этой микроРНК в клетках гранулярного слоя коры мозжечка [187].

Ген *APC*, отделённый от *NREP* инсуляторной границей между ТАДами (Рис. 10А,Д), является хорошо изученным геном-онкосупрессором, который играет ключевую роль в регуляции клеточного цикла и адгезии клеток. Основная функция *APC* заключается в контроле сигнального пути WNT: он входит в состав комплекса, разрушающего  $\beta$ -катенин, тем самым предотвращая внутриклеточную передачу сигнала WNT-пути и подавляя активацию генов, отвечающих за пролиферацию [188]. Инактивирующие мутации в этом гене могут привести к заболеваниям, таким как семейный аденоматозный полипоз, и часто ассоциированы с колоректальным раком [189–191]. У мышей гомозиготные мутации гена *Apc* имеют серьёзные фенотипические манифестации, включая аномалии гаструляции, нарушения развития мозга, отсутствие переднего и среднего мозга, отсутствие нижней челюсти, акранию [192]. Эти аномалии приводят к эмбриональной или пренатальной летальности [193].

Таким образом, граница топологических доменов в этом локусе инсулирует регуляторные ландшафты двух клинически-значимых генов, выполняющих важные функции в процессах развития и жизнеобеспечения организма. Высокий уровень инсуляции в границе, а также ее

консервативность, дают основания полагать, что мутации, нарушающие локальную пространственную организацию хроматина, могут приводить к изменениям в регуляции транскрипции генов *APC* и *NREP*, и последующим патологиям.

Таким образом, локусы *Slc29a3/Unc5b* и *Epb4114a* были выбраны как две модельные системы, позволяющие проанализировать разные типы нарушений локальной пространственной организации хроматина. В первом случае анализ был направлен на оценку последствий удаления отдельных сайтов связывания CTCF для структуры пространственных контактов и экспрессии группы близко расположенных генов в различных тканевых контекстах. Во втором случае исследовались транскрипционные эффекты крупных хромосомных перестроек, которые могут нарушать изоляцию соседних ТАДов и создавать условия для взаимного влияния разных регуляторных ландшафтов.

В совокупности эти две модели позволяют оценить функциональные последствия нарушения пространственной архитектуры хроматина в разных структурных и регуляторных контекстах: при локальном удалении сайтов связывания CTCF, участвующих в организации пространственных контактов внутри сложно организованного локуса, и при более крупных перестройках, затрагивающих границу ТАДов, разделяющую крупные регуляторные ландшафты. Такой подход дает возможность сопоставить вклад топологической организации генома в регуляцию экспрессии генов в различных по регуляторной логике генетических контекстах.

## Материалы и методы

### Материалы

#### Линия мышей

В качестве исходной линии для генно-инженерных манипуляций была взята инбредная линия мышей *Mus musculus domesticus* C57BL/6 из ЦКП «SPF Виварий» ИЦиГ. Мыши *Mus musculus castaneus* (CAST) также были предоставлены ЦКП «SPF Виварий» ИЦиГ.

#### Олигонуклеотиды

Используемые для ПЦР и получения генетических конструкций олигонуклеотиды (Приложения, Таблица 1) были синтезированы лабораторией «Biosset» (Новосибирск) либо лабораторией ООО «ДНК-Синтез» (Москва).

## **Методы**

### **Работа с мышами**

Мыши линии C57BL/6J содержались в условиях конвенционального вивария ИЦиГ и затем SPF-вивария. В конвенциональном виварии подстилка менялась еженедельно, пища (Комбикорм полнорационный для содержания мелких лабораторных грызунов SPF категории, апатогенный, автоклавируемый (ЗАО «Ассортимент-Агро», Россия)) и вода добавлялись по мере необходимости и находились в свободном доступе для животных. Клетки и бутылки с водой мылись еженедельно мыльным раствором и подвергались тепловой обработке температурой 80 °С. Ссаживание животных на размножение происходило в возрасте не менее 1,5 месяцев. Самец отсаживался после появления у самки очевидных признаков беременности. Рассаживание потомков по полу происходило на 1 месяц жизни. Во время эксперимента эвтаназия проводилась с использованием воздействия углекислого газа, а для забора тканей применялся метод цервикальной дислокации. Все методы соответствовали рекомендациям AVMA по эвтаназии животных для минимизации стресса.

Для эмбриотрансфера микроинъекцированных зигот использовали псевдобеременных самок мышей CD-1. Цитоплазматическая микроинъекция зигот была проведена с использованием стандартных техник, широко применяемых в трансгенезе. Эмбриологические работы были выполнены сотрудниками Лаборатории генетики развития ИЦиГ СО РАН Кораблевым А.Н. и Серовой И.А.

Все процедуры с животными были одобрены Комитетом по этике Института цитологии и генетики (протокол №65, выдан 09 октября 2020 г.). Экспериментальные результаты, связанные с живыми животными, были представлены в соответствии с рекомендациями ARRIVE.

## Получение и подготовка CRISPR/Cas9 хнРНК

Мы спроектировали CRISPR хнРНК для желаемых областей, используя веб-инструмент «Benchling» (<https://benchling.com>). Для экспериментов на локусе *Slc29a3/Unc5b* мы смоделировали последовательности желаемых мутаций и выбрали последовательность ssODN, одноцепочечного ДНК донора гомологии для репарации путем гомологичной рекомбинации длиной в 60 оснований, и праймеров для генотипирования (Приложения, Таблица 1). В последовательность ssODN для левого, среднего и правого локусов мы ввели узнающий сайт *HindIII* для удобства генотипирования. Шаблоны ДНК для синтеза хнРНК получали путем ПЦР, используя олигонуклеотиды, содержащие промотор T7, последовательность руководства и матрицу хнРНК. Продукты ПЦР использовались для *in vitro* транскрипции (MEGAscript™ T7 Transcription Kit, AM1354, Invitrogen, США). Полученная RNA очищалась на колонках MEGAclear™ Transcription Clean-Up Kit (AM1908, Invitrogen, США) и смешивалась с mRNA spCas9 (GeneArt™ CRISPR Nuclease mRNA, Thermo Fisher Scientific, США) в пропорциях 8,2 пмоль каждого хнРНК и 16,4 пмоль mRNA Cas9. ssODN очищали на колонках MEGAclear™ Transcription Clean-Up Kit (AM1908, Invitrogen, США) и разводили в воде mQ до конечной концентрации 250 нг/мкл. RNA и ssODN смешивали непосредственно перед микроинъекцией. Смеси для микроинъекции были переданы коллегам И.А. Серовой и А.Н. Кораблеву для проведения эмбриологических работ и микроинъекции.

## Выделение тотальной РНК

Проводилось с помощью реактива TRI-Reagent (AM9738, Sigma-Aldrich, США) по протоколу производителя. К образцу ткани добавляли 500 мкл TRI-Reagent, гомогенизировали на гомогенизаторе Доунца, инкубировали 5 минут при комнатной температуре. Добавляли 100 мкл хлороформа, перемешивали переворачиванием пробирки, инкубировали 10 минут при комнатной температуре. Центрифугировали 15 минут на 12000 g. Переносили верхнюю

фазу в новую пробирку, добавляли 150 мкл изопропанола. Тщательно перемешивали, затем центрифугировали 15 минут на 12000 g. Удаляли супернатант декантацией, промывали 75% этиловым спиртом, сушили на воздухе 10 минут. Растворяли в 50 мкл раствора DEPC в ddH<sub>2</sub>O, измеряли концентрацию РНК на спектрофотометре Nanodrop, затем при возможности доводили концентрацию до 1 мкг/мкл.

### **Обратная транскрипция**

Обратная транскрипция проводилась с помощью набора RevertAid First Strand cDNA Synthesis Kit (K1621, Thermo Fisher Scientific, США) по инструкции производителя. В случае применения обработки ДНКазой, к 1 мкг РНК добавляли 1 мкл буфера для ДНКазы I с MgCl<sub>2</sub> (100 mM Tris-HCl pH 7.5, 25 mM MgCl<sub>2</sub>, 1 mM CaCl<sub>2</sub>), 1 мкг ДНКазы I и ddH<sub>2</sub>O до объема реакции в 10 мкл. Инкубировали на 37 °C 30 минут, затем добавляли 1 мкл ЭДТА и инкубировали при 65 °C 10 минут. В случае без обработки ДНКазой разводили 1 мкг РНК водой ddH<sub>2</sub>O до объема 11 мкл. Затем добавляли 1 мкл ДНК-гексамеров случайной последовательности, 4 мкл 5X Reaction Buffer, 1 мкл Ribolock, 2 мкл dNTP и 1 мкл обратной транскриптазы RevertAid. Затем инкубировали в термоциклере 5 минут на 25 °C, 60 минут на 42 °C и 5 минут на 70 °C. Затем разбавляли раствором DEPC в ddH<sub>2</sub>O до желаемого объема. кДНК хранили при температуре -80 °C.

### **Выделение геномной ДНК**

Выделение геномной ДНК производили фенол-хлороформным методом по следующему протоколу. К образцам хвостов мышей добавляли 500 мкл лизирующего буфера (100 mM NaCl, 10 mM Tris pH 8.0, 25 mM ЭДТА pH 8.0, 0,5% SDS (SD00100050, Scharlab, Испания)), протеиназу К (20 мг/мл) (405002 BIORON GmbH, Германия) до конечной концентрации 0,5 мкг/мл. Инкубировали клетки при 55 °C в течение 2 часов до полного лизиса ткани, после чего добавляли 55 мкл 3M ацетата натрия (pH 5.0) и 500 мкл смеси

фенола и хлороформа (1:1). Аккуратно перемешивали, центрифугировали 10 мин при 14,5 тыс. об./мин. Аккуратно отбирали верхнюю фазу, содержащую ДНК, и переносили в чистую пробирку. Очищали ДНК по протоколу переосаждения изопропиловым спиртом. Добавляли изопропиловый спирт в объеме, равном 70% очищаемой смеси, перемешивали на вортексе, центрифугировали 10 мин на 14,5 об./мин, отбирали супернатант. Добавляли 3 объема исходной очищаемой смеси 70% этилового спирта, центрифугировали 5 мин при 14,5 тыс. об./мин, отбирали супернатант, добавляли 3 объема 70% этилового спирта, центрифугировали 5 минут на 14,5 тыс. об./мин, отбирали супернатант. Высушивали оставшийся спирт, растворяли ДНК в 50 мкл ddH<sub>2</sub>O. Концентрацию ДНК измеряли на спектрофотометре NanoDrop, после чего разводили ДНК до концентрации 100 нг/мкл для использования в ПЦР.

### **Полимеразная цепная реакция (ПЦР)**

Для проведения рутинной полимеразной цепной реакции в объеме 25 мкл с Taq-полимеразой замешивали следующие реагенты: 10X ПЦР буфер, 0,5 мкл 4dNTPs (dATP, dGTP, dCTP, dTTP) (10 мМ), 1 мкл каждого из соответствующих праймеров (10 мкМ), 100 нг исследуемого образца ДНК (или мазок с микробной колонии), 0,5 мкл (2,5 ед.) Taq-полимеразы, ddH<sub>2</sub>O до объема 25 мкл. Полученную реакционную смесь инкубировали в амплификаторе в следующих условиях реакции: 1 цикл – 94 °С (4 мин); 35 циклов – 94 °С (30 сек), температура отжига праймеров (35 сек), 1 цикл – 72 °С (30 сек), 1 цикл – 72 °С (3 мин); 10 °С ( $\infty$ ).

### **Цифровая ПЦР с флуоресцентными зондами**

Для проведения цифровой ПЦР с флуоресцентными зондами использовался коммерческий набор QIAcuity® Probe PCR Kit (250102, Qiagen, Германия). В соответствии с протоколом, на 1 реакцию в 12 мкл замешивали: 3 мкл реагента 4x Probe PCR Master Mix, по 1 мкл каждого из

соответствующих праймеров (10 мкМ) и зондов (10 мкМ), и 0,5 мкл образца кДНК. Реакционную смесь переносили в реакционные нанопланшеты и заклеивали специализированной пленкой. Планшеты помещались в прибор QIAcuity One (с5-канальным датчиком) (Qiagen, Германия), реакция проводилась по схеме: 1 цикл – 95 °С (2 мин); 35 циклов – 95 °С (15 сек), 62 °С (30 сек); фотодетекция флуоресценции 1; 10 циклов – 95 °С (15 сек), 62 °С (30 сек); фотодетекция флуоресценции 2.

### **Анализ данных цифровой ПЦР с флуоресцентными зондами**

Проводился в программном обеспечении QIAcuity Software Suite (Qiagen, Германия) с помощью встроенных функций, и затем с помощью собственноручно написанного скрипта Python, материалы и код доступны по ссылке [https://github.com/Somatich/QIAcuity\\_processing/](https://github.com/Somatich/QIAcuity_processing/)

### **Электрофорез в агарозном геле**

К 1-3 г агарозы (для приготовления 1-3% гелей соответственно) добавляли 100 мл ТАЕ буфера (40 мМ Tris-base, CH<sub>3</sub>COOH (до конечного рН раствора 7,6), 1 мМ ЭДТА), смесь нагревали до полного растворения агарозы и добавляли бромид этидия до концентрации 2 мкг/мл. Раствор заливали в ванночку для геля и устанавливали в ней гребенку для формирования карманов. После полимеризации гель переносили в камеру для горизонтального электрофореза. Исследуемые образцы ДНК предварительно смешивали с загрузочным буфером (0.3 мг/мл Ксилен Цианол; 60% глицерин; 20 мМ Tris-HCl рН 8,0; 60 мМ ЭДТА; 0,5% SDS), в пропорции по умолчанию 5 мкл образца к 2 мкл загрузочного буфера, тщательно перемешивали и помещали в карманы геля.

Электрофорез вели в ТАЕ-буфере при напряжении 5 В/см. Для определения длины фрагментов ДНК в дополнительный карман геля добавляли 5 мкл маркера длины 1000 п.н. (ДНК-маркеры 1 тыс. п.н. (13 фрагментов от 0,25 до 10 тыс. п.н.), СибЭнзим) Для фрагментов длиной менее

1000 п.н. использовался маркер с шагом 100 п.н. (ДНК-маркеры 100 п.н. + 1,5 тыс. п.н. (11 фрагментов от 100 до 1500 п.н.), СибЭнзим, Россия). Результаты визуализировали с помощью трансиллюминатора (ChemiDoc, BioRad, США) в проходящем УФ.

### **Секвенирование по Сэнгеру**

Для проведения реакции секвенирования набором BigDye™ Terminator v3.1 Cycle Sequencing Kit (4337455, Thermo Fisher Scientific, США) в объеме 20 мкл замешивали следующие реагенты: 5 мкл 5X Big Dye buffer, 2 мкл BigDye™ Terminators v3.1; 0,7 мкл соответствующего праймера (10 мМ); 500-800 нг исследуемого образца ДНК, milliQ H<sub>2</sub>O до 20 мкл. Полученную реакционную смесь инкубировали в амплификаторе в следующих условиях реакции: 1 цикл – 96 °С (10 мин); 35 циклов — 96 °С (10 сек), 50 °С (10 сек), 60 °С (4 мин); 10 °С (∞). Полученный продукт реакции очищали на колонках с сорбентом Sephadex G50. Для приготовления Sephadex G50 4 г сухого порошка сорбента заливали 100 мл ddH<sub>2</sub>O и оставляли перемешиваться на шейкере на 12-18 часов при комнатной температуре. Подготовленный Sephadex G50 перемешивали, аликвотой (700 мкл) суспензии сорбента заполняли пустой чистый корпус колонки, изготовленной следующим образом. Брали наконечник с фильтром на 1000 мкл для автоматической пипетки, пинцетом проталкивали фильтр к соплу до упора, сопло обрезали так, чтобы наконечник плотно закреплялся в пробирке на 2 мл, не оставалось большого зазора, в котором могла бы задержаться капля жидкости, между фильтром и краем разреза, и расстояние от дна пробирки до края разреза было достаточным для накопления в нем центрифугируемой жидкости. Полученную таким образом колонку заполняли сорбентом, помещали в колонкоприемник на 2 мл и центрифугировали 5 мин с ускорением 900 g. Воду из колонкоприемника убрали. Добавляли на колонки повторную аликвоту суспензии (доверху), повторно центрифугировали в пробирках 5 мин с ускорением 900 g. Затем промывали сорбент (для этого добавляли в каждую

колонку на столбик сорбента 200 мкл milliQ H<sub>2</sub>O и центрифугировали 5 мин с ускорением 900 g). После этого помещали подготовленные колонки в чистые, заранее подписанные пробирки на 1,5 мл. На центр столбика сорбента наносили образец реакции Сэнгера (20 мкл) и центрифугировали 5 мин с ускорением 900 g. Собранную в пробирки жидкость высушивали центрифугированием в течение 45 минут при 40 °С с ускорением 10000 g. Подготовленные таким образом продукты далее анализировали в ЦКП «Молекулярная и клеточная биология» ИМКБ СО РАН.

### **ПЦР-генотипирование мышей на делецию dM1R1**

Для определения мутационного аллеля dM1R1 с выделенного из мышечной ткани образца ДНК ставили первую реакцию вложенного ПЦР с праймерами MLF2U и RRRD (ожидаемая длина продукта – 470 п.н.). С продуктом этой реакции, разведенным в 1000 раз, ставили следующий раунд вложенного ПЦР с праймерами MLF2 и RRR (ожидаемая длина продукта – 200 п.н.). Для определения аллеля дикого типа ставили реакции ПЦР с праймерами MLF2 и MLR2 (ожидаемая длина продукта – 450 п.н.); RRF и RRR (ожидаемая длина продукта – 330 п.н.). Поиск инверсий и дупликаций у отредактированных животных проводился с помощью ПЦР со следующими парами праймеров: MLF2U и RRF; MLR2 и RRRD; MLR2 и RRF. Продукты реакции анализировались с помощью гель-электрофореза в 2-3% агарозном геле.

### **ПЦР-ПДФ-генотипирование мышей на мутации в сайтах L2, M2, R2**

Для определения мутаций в сайтах связывания CTCF L2, M2, R2 с выделенного из мышечной ткани образца ДНК ставили реакцию ПЦР с праймерами *Unc5b-L2-F* и *Unc5b-L2-R* (ожидаемая длина продукта – 461 п.н.); *Unc5b-M2-F* и *Unc5b-M2-R* (ожидаемая длина продукта – 276 п.н.); *q-mUNC5B-R2* и *Unc5b-R2-R* (ожидаемая длина продукта – 324 п.н.) соответственно. Продукты ПЦР-реакции добавляли в смесь для тестовой

реакции гидролиза с ферментами рестрикции *HindIII* или *PstI*, *SacI*, *HaeIII* соответственно. Анализ длин продуктов гидролиза осуществлялся с помощью гель-электрофореза в 2–3% агарозном геле.

### **Приготовление NGS-библиотек для генотипирования**

Для каждого животного и для каждого сайта из L2, M2 и R2 были получены ПЦР-продукты с использованием баркодированных праймеров для различения последовательностей, полученных из разных животных. Приготовление библиотек было выполнено с использованием КАРА HyperPrep Kit (КК8500, Roche, Швейцария) по протоколу производителя, без амплификации продуктов лигирования адаптеров (для избежания рекомбинации ДНК в ходе амплификации). Секвенирование было выполнено компанией BGI (Шэньчжэнь, Китай). В общей сложности получено 25522262 парных прочтения (2\*150 п.н.). Таким образом на каждый локус каждого животного приходилось в среднем по 400 тысяч прочтений.

### **Генотипирование мышей на мутации в локусе *Erb4114a***

Для генотипирования полученных генетически модифицированных мышей и их потомства мы использовали следующие пары праймеров:

Аллель с делецией: Erb\_DEL\_F, Erb\_DEL\_R

Аллель с инверсией:

Для левого края инверсии (сопровождающейся незаслуженной ~7,8 тыс. п.н. делецией): Erb\_INV\_L\_F, Erb\_INV\_L\_R

Для правого края инверсии: Erb\_INV\_R\_F, Erb\_DEL\_R

Аллель дикого типа: (Обнаружение интактных участков CRISPR)

Пара 1: DEL\_F, INV\_R\_F

Пара 2: INV\_L\_R, DEL\_R

Для каждого предсказанного нецелевого сайта внесения двуцепочечного разрыва участка в пределах блока сцепления в 50 млн. п.н. от целевой мутации мы провели секвенирование по Сэнгеру геномной ДНК мышей, выбранных для анализа РНК-секвенирования.

### **Секвенирование иммунопреципитации хроматина (ChIP-Seq)**

Свежеизвлеченные образцы печени (около одного грамма) измельчались бритвой, гомогенизировались в стеклянном гомогенизаторе Даунса в 1 мл 3% формальдегида, затем образцы переносились в пробирку объемом 50 мл, и объем 3% формальдегида доводился до 40 мл. После 10-минутной инкубации при комнатной температуре фиксатор был нейтрализован путем добавления глицина до конечной концентрации 125 мМ, образцы инкубировались 10 минут и дважды промывались ледяным PBS. Затем образцы быстро замораживали и хранили при  $-80^{\circ}\text{C}$ . Замороженные образцы размораживали и лизировали в течение 30 минут в буфере для лизиса (10 мМ Tris-HCl, 1 мМ ЭДТА, 1% Triton X-100, 0.1% деоксихолат натрия, ингибиторы протеаз, 0.5% SDS). Хроматин измельчался с использованием соникатора Bandelin Sonopulse (70% мощности, 11 циклов 30/90 сек ON/OFF). Для одного эксперимента использовалось 20-40 мкг фиксированного хроматина в пересчете на ДНК (концентрация определена и подобрана путем выделения тотальной ДНК из части образца). Перед инкубацией со специфическими антителами, образец разбавлялся в 5 раз буфером для лизиса без SDS и предварительно очищался путем инкубации с магнитными частицами Protein A (New England Biolabs, S1425S) в течение 2 часов при  $4^{\circ}\text{C}$  с медленной ротацией. В это время другая порция магнитных частиц Protein A промывалась в PBS, объединялась с 5 мкг целевых антител, и инкубировалась в течение 2 часов при  $4^{\circ}\text{C}$  с медленной ротацией. Частицы удаляли на магнитном штативе и предварительно очищенный хроматин иммунопреципитировали с комплексами антител/магнитных частиц в течение ночи при  $4^{\circ}\text{C}$  с медленной ротацией. На следующий день образцы тщательно промывались в ряде буферов (Буфер 1:

10 mM Tris-HCl, 1 mM ЭДТА, 1% Triton X-100, 0.1% SDS, 0.1% деоксихолат натрия, ингибиторы протеаз; Буфер 2: 500 mM NaCl, 10 mM Tris-HCl, 1 mM ЭДТА, 1% Triton X-100, 0.1% SDS, 0.1% деоксихолат натрия, ингибиторы протеаз; Буфер 3: 0.25 M LiCl, 10 mM Tris-HCl, 1 mM ЭДТА, 0.5% NP-40, 0.5% деоксихолат натрия; Буфер TE/Triton: 10 mM Tris-HCl, 1 mM ЭДТА, 1% Triton X-100; Буфер TE: 10 mM Tris-HCl, 1 mM ЭДТА). ДНК элюировалась в 100 мкл буфера для элюции (10 mM Tris-HCl, 1 mM ЭДТА, 1% SDS) путем инкубации при 65 °C в течение 14 часов. После обработки с помощью РНКазы А (New England Biolabs, T3018) и Протеиназы К (New England Biolabs, P8107S), магнитные частицы удалялись. ДНК очищали с использованием колонок ChIP DNA Clean & Concentrator (Zymo Research, D5205). Библиотеки ChIP-Seq подготавливались для секвенирования с использованием Кара Hyper Prep Kit (KK8504, Roche, Швейцария) и набора одиночно индексированных адаптеров КАРА (KK8701, Roche, Швейцария). Секвенирование было выполнено компанией BGI (Шэньчжэнь, Китай). Данные ChIP-Seq обрабатывались с использованием стандартного алгоритма ENCODE ([https://github.com/kundajelab/chipseq\\_pipeline](https://github.com/kundajelab/chipseq_pipeline)).

### **Статистический анализ результатов количественной ПЦР**

Относительная экспрессия генов была рассчитана с помощью алгоритма  $2^{-\Delta\Delta Ct}$  с модификацией для множественного эндогенного контроля (при расчете  $\Delta Ct$  контрольного гена бралось геометрическое среднее между  $\Delta Ct$  Eef2 и Actb). Статистический анализ проводился с помощью U-критерия Манна-Уитни.

### **Приготовление библиотек Hi-C**

Подготовка библиотеки Hi-C проводилась в соответствии с оригинальным протоколом Hi-C 2.0 с незначительными адаптациями для образцов тканей:

Образцы тканей вручную измельчали с помощью острого лезвия, затем суспендировали в 2 мл PBS. Смесь гомогенизировали с использованием

стеклянного гомогенизатора в течение 20 движений на льду. Добавляли формальдегид до конечной концентрации 2% и инкубировали смесь в течение 10 минут при комнатной температуре с вращением. Затем добавляли глицин до конечной концентрации 250 мМ и инкубировали еще 10 минут при тех же условиях. Образцы центрифугировали при 1000g в течение 5 минут, затем дважды промывали холодным PBS, после чего супернатант удаляли. Осадок быстро замораживали в жидком азоте или непосредственно ресуспендировали в холодном лизисном буфере (10 мМ Tris, 10 мМ NaCl, 0.2% Igepal CA-630 (9002-93-1, Sigma-Aldrich, Германия)). Гомогенизацию проводили с использованием шприца, а лизат фильтровали через марлю. Гомогенат инкубировали в течение 30 минут на льду, центрифугировали при 600 g в течение 5 минут, затем дважды промывали холодным лизисным буфером и один раз буфером NeBuf 3.1 (B7203, NEB, США), содержащим 0.3% SDS. Осадок ресуспендировали в NeBuf 3.1 с 0.1% SDS и инкубировали в течение 30 минут при 37°C, затем добавляли 200 мкл 1.5% Triton X-100 и продолжали инкубацию при тех же условиях. Добавляли фермент DpnII (5 мкл), и смесь переваривали в течение ночи при 37°C. После 20-минутной инкубации при 65°C смесь центрифугировали для удаления супернатанта. ДНК подготавливали для восстановления концов, ресуспендируя в 150 мкл смеси для восстановления концов ДНК (50 мМ dGTP, 50 мМ dATP, 50 мМ dTTP, 50 мМ dCTP-15bio, 1X NEBuffer 2.1, 25U PolIII Klenow fragment) и инкубировали в течение 4 часов при 23°C. Лигирование проводили, добавляя 1 мл смеси для лигирования (1X T4 Ligase Buffer, 1% Triton X-100, 5% PEG, 100 мкг/мл BSA, 1 мМ АТР, 4000U T4 Ligase) и инкубировали в течение ночи при 16°C. Далее процедура продолжалась в соответствии с установленным протоколом. Секвенирование было выполнено компанией BGI (Шэньчжэнь, Китай). Обогащение библиотек за последовательности, принадлежащие локусу *Slc29a3/Unc5b*, было проведено с помощью произведенной по индивидуальному заказу таргетной панели РНК-олигонуклеотидов myBaits

Custom DNA-Seq (Arbor Bioscience) для региона chr10:60103000-61356000 (mm10) по протоколу производителя.

### **Приготовление библиотек для UMI-опосредованного таргетного РНК-секвенирования**

Для приготовления NGS-библиотек с целью получения информации о сравнительных уровнях экспрессии аллелей мы проводили обратную транскрипцию с использованием ген-специфичных праймеров, содержащих уникальные молекулярные идентификаторы (UMI) и часть адаптера для секвенирования на 5'-конце, используя набор для обратной транскрипции RNAScribe (№R04-10 Biolabmix, Россия) в соответствии с протоколом производителя. Далее амплифицировали кДНК с использованием ген-специфичного обратного праймера и праймера, отжигающегося на части адаптера для секвенирования праймера обратной транскрипции. Эти ампликлоны индексировались на втором этапе ПЦР с использованием самодельных индексных праймеров с завершением последовательностей адаптера для секвенирования, затем очищались с помощью SPRI-частиц и секвенировались в режиме парных прочтений с использованием сервиса BGI (Шэньчжэнь, Китай) на платформе секвенирования DNBSEQ.

Данные секвенирования выравнивались по ожидаемым последовательностям транскриптов с использованием инструмента bowtie2 с параметрами по умолчанию. Последовательности UMI извлекались с помощью umi\_tools extract с параметром bc-pattern=NNNNXNNNNXNNNN, а затем дублирующиеся последовательности удалялись с помощью umi\_tools dedup. Данные о SNP собирались с помощью bcftools mpileup с параметром -a AD, затем файлы VCF анализировались с использованием самодельного скрипта Python. Для каждого животного количество SNP C57BL/6 нормализовалось делением на количество SNP CAST, что приводило к сравнительному уровню аллельной экспрессии. Значимость различий этих распределений оценивалась с помощью теста Манна-Уитни,

скорректированного с помощью поправки Бенджамини/Хохберга для множественного тестирования с порогом p-value 0.1.

### **Секвенирование РНК**

Ткани 6-месячных самцов мышей были отобраны и мгновенно заморожены. РНК была выделена с помощью реагента TRIzol (Thermo Fisher Scientific, США, 15596026) в соответствии с инструкциями производителя. Мы подготовили два биологических повтора для каждого генотипа и органа. Библиотеки РНК были приготовлены с использованием метода с сохранением информации о направлении транскрипции и секвенированы в режиме парного чтения с использованием сервиса BGI на платформе секвенирования DNBSEQ. Средняя глубина секвенирования составила 24 миллиона пар прочтений на образец с длиной прочтения 100 пар оснований. Контроль качества и предварительная обработка сырых данных секвенирования были выполнены программой fastp (<https://github.com/OpenGene/fastp>) с параметрами по умолчанию. Далее данные РНК-секвенирования были количественно оценены с помощью программы Salmon с использованием последовательностей транскриптов из GENECODE (версия M33). Дифференциально экспрессирующиеся гены были выявлены с помощью программы DESeq2 со стандартными параметрами. Мы выбрали дифференциально экспрессирующиеся гены как все гены с скорректированным значением  $p\text{-adj} < 0,05$  (значение p по тесту Вальда с использованием поправки по методу Бенджамини-Хохберга).

## Результаты

### 1. Стратегия выбора модельных локусов

Цель данной работы заключается в изучении факторов, опосредующих участие пространственной организации хроматина в регуляции генной экспрессии. Одним из ключевых факторов является взаимодействие между ДНК-связывающим белком CTCF и белковым комплексом когезина, которое играет важную роль в формировании и поддержании паттерна генной экспрессии в клетках. Литературные данные показывают, что реорганизация архитектуры хроматина может иметь разнообразные последствия, включая значительное влияние на генную экспрессию и фенотип организма, в зависимости от исследуемой модели [14,24,88]. Основная гипотеза нашей работы заключается в том, что регуляторная функция архитектурных элементов хроматина может различаться в зависимости от эпигенетического окружения и вовлеченных молекулярно-генетических механизмов.

Топология хроматина на уровне ТАДов зависит от геномного распределения транскрипционного фактора CTCF. Мутации в сайтах связывания CTCF могут вызывать патологии и изменять локальную структуру хроматина. Для изучения этих мутаций требуются функциональные исследования *in vivo*. Один из подходов – удаление кластеров сайтов CTCF в границах ТАДов с помощью технологий геномного редактирования.

Первым этапом является выбор модифицируемого локуса в геноме мыши, для чего мы сформулировали ряд критериев. Так, локус должен содержать сильную инсуляторную границу двух топологических доменов, образованную кластером сайтов связывания CTCF. Топологические домены локуса должны воспроизводимо идентифицироваться в независимых исследованиях, а также консервативно сохраняться между типами клеток и синтетическими локусами различных видов. Важно, чтобы в локусе содержались функционально значимые гены с выраженными фенотипами при нарушении экспрессии, но

при этом важно, чтобы участвующие в образовании границы ТАДов сайты связывания CTCF не пересекались с функциональными элементами этих генов (экзонами или энхансерами сплайсинга). Кроме того, для исследования генной экспрессии на организменном мультиорганном уровне, нам необходимо было выбрать конститутивно активные гены, поэтому мы сосредоточились на локусах, содержащих конститутивно активные гены. Согласно данным [82,194,195], можно оценить число локусов в геноме мыши, удовлетворяющим перечисленным критериям, в несколько сотен (Рис. 7).

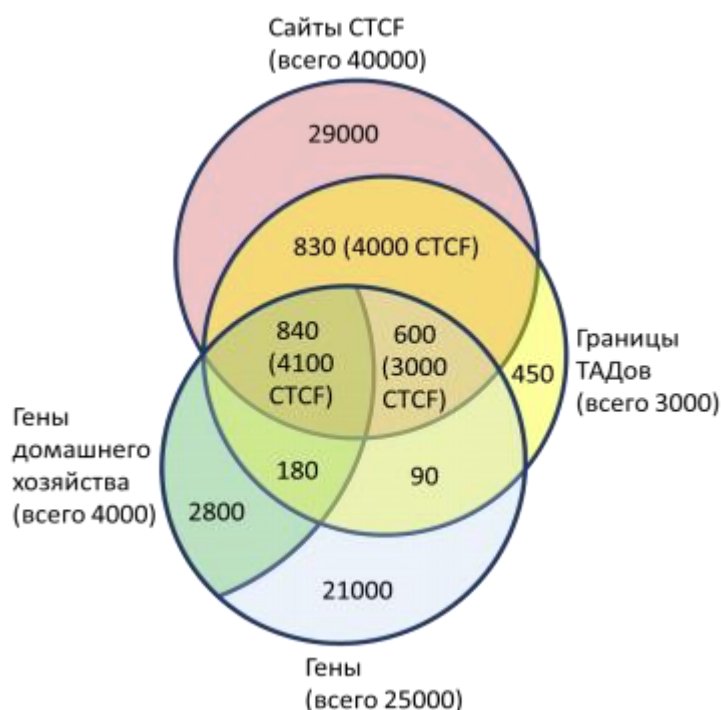


Рисунок 7. Диаграмма Венна, показывающая соотношение границ ТАДов с CTCF сайтами и генами (в том числе домашнего хозяйства), встречающимися в них.

Руководствуясь перечисленными соображениями при визуальном анализе публичных Hi-C карт мыши [196], мы выбрали две хорошо картированные и ранее не исследованные границы ТАДов. Одна из них располагается в короткой межгенной области между генами *Slc29a3* и *Unc5b*

на хромосоме 10 мыши, а вторая в 3'-терминальной области гена *Epb4114a* на хромосоме 18. В дальнейшем мы будем коротко называть эти локусы по именам генов – *Slc29a3/Unc5b* и *Epb4114a* соответственно.

Локус *Slc29a3/Unc5b* примечателен тем, что он содержит два ярко выраженных топологических домена, разделенных очень сильной границей. Граница находится между двумя генами и является одной из самых сильных по инсуляции в геноме мыши. Анализ публичных данных Hi-C показал, что эта граница эволюционно консервативна и присутствует в геномах мыши, человека, курицы и даже шпорцевой лягушки, что согласуется с ее функциональной значимостью для регуляции генов локуса (Рис. 8А-В). Мы предположили, что мутации в этой границе могут привести к эпигенетическим и фенотипическим изменениям, которые мы сможем наблюдать.

Границу ТАДов между генами *Slc29a3* и *Unc5b* образуют два кластера сайтов связывания CTCF (Рис. 8А,Г, Рис. 11А). Три центромерных сайта связывания CTCF имеют обратную ориентацию, в то время как три теломерных – прямую ориентацию относительно оси геномных координат. Такое расположение сайтов связывания CTCF предполагает возможность образования устойчивых хроматиновых петель с локусами, расположенными как выше, так и ниже по геномной координате. Действительно, на Hi-C картах ясно видно, что граница образует множество петель с другими регионами внутри прилежащих ТАДов (Рис. 8А).

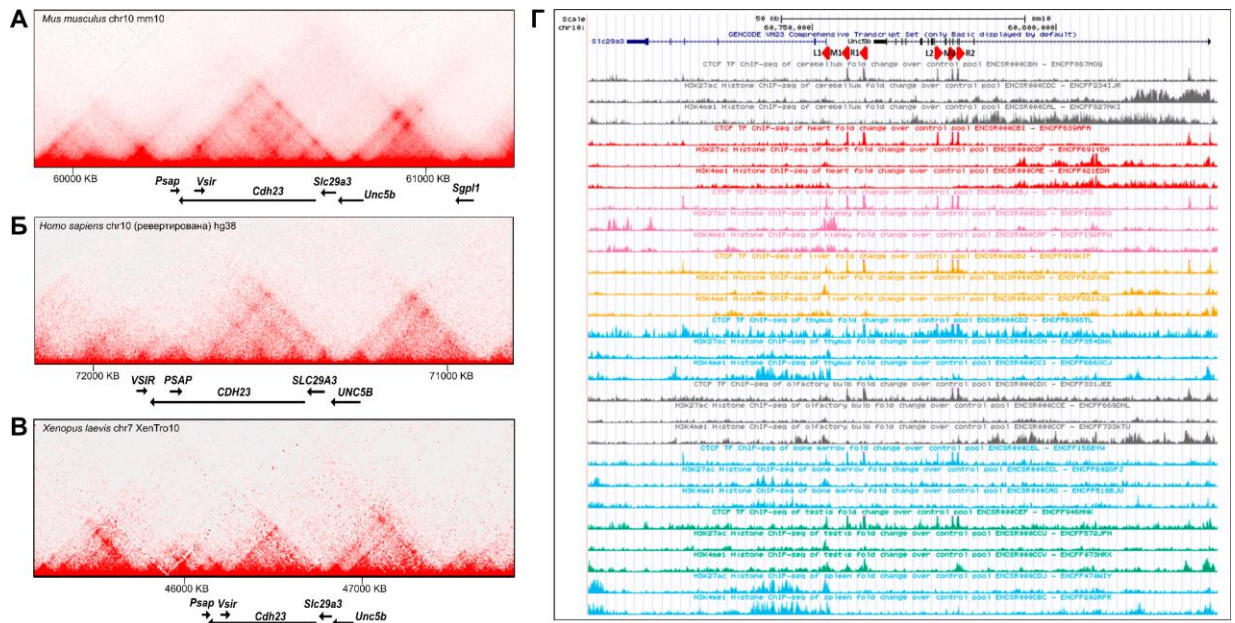


Рисунок 8. Пространственная организация локуса *Slc29a3/Unc5b* в соответствии с публичными данными. А-В – карты пространственных контактов Hi-C в ортологичных регионах геномов мыши (*Mus musculus*, хромосома 10, mm10), человека (*Homo sapiens*, хромосома 10, hg38; регион показан в обратной ориентации) и шпорцевой лягушки (*Xenopus laevis*, хромосома 7, XenTro10). Под картами показано расположение генов локуса и направление их транскрипции. Сходство структуры контактных карт у разных видов указывает на консервативность топологической организации данного региона. Г – отображение локуса *Slc29a3/Unc5b* генома мыши в геномном браузере UCSC. Показаны ChIP-Seq профили для CTCF и гистоновых меток H3K4me1 и H3K27ac, колокализация которых является маркером активных энхансеров, и расположение сайтов связывания CTCF в границе ТАДов (ориентация отмечена стрелками).

В межгенной области находятся три сайта связывания CTCF, один – в промоторной области гена *Slc29a3*, из-за чего мутации в нем могут привести к не связанному с трехмерной архитектурой хроматина изменению генной экспрессии и потому нежелательны. Удаление двух оставшихся сайтов не встречает технических трудностей. Кроме того, так как эти сайты имеют реверс-ориентацию, именно они, вероятно, формируют контакты с

вышележащими районами генома. Поэтому мы ожидаем, что уже только удаление этих двух сайтов должно привести к значительному изменению пространственной архитектуры локуса и функциональным последствиям. Еще три сайта в форвард-ориентации находятся в терминальной области гена *Unc5b*, в интронах длиной менее 1000 п.н. Удаление этих сайтов встречает трудности, так как могут быть затронуты важные регуляторные элементы гена *Unc5b* и его экзоны.

Во второй части нашего исследования мы сосредоточились на границе ТАДов между клинически-значимыми генами *Nrep* и *Apc*, расположенной в терминальной области гена *Epb41l4a*. Эта граница у человека располагается на хромосоме 5, а ортологичный участок у мыши – на синтенном участке мышиной хромосомы 18 (Рис. 10А,Д). Данная граница является одной из наиболее сильных по инсуляции в человеческом геноме. Мы провели полногеномный анализ инсуляции границ ТАДов с использованием открытых данных и определили, что эта граница входит в 99,8-й перцентиль по индексу инсуляции в геноме линии эмбриональных стволовых клеток человека H1-hESC (Рис. 9). Согласно проведенному нами межвидовому сравнению, граница консервативна как между синтенными регионами генома у человека и мыши, так и в различных типах клеток обоих видов (Рис. 10А). Формирование этой границы обусловлено массивным кластером сайтов связывания CTCF, расположенных в терминальной области гена *EPB41L4A* (Рис. 10В).

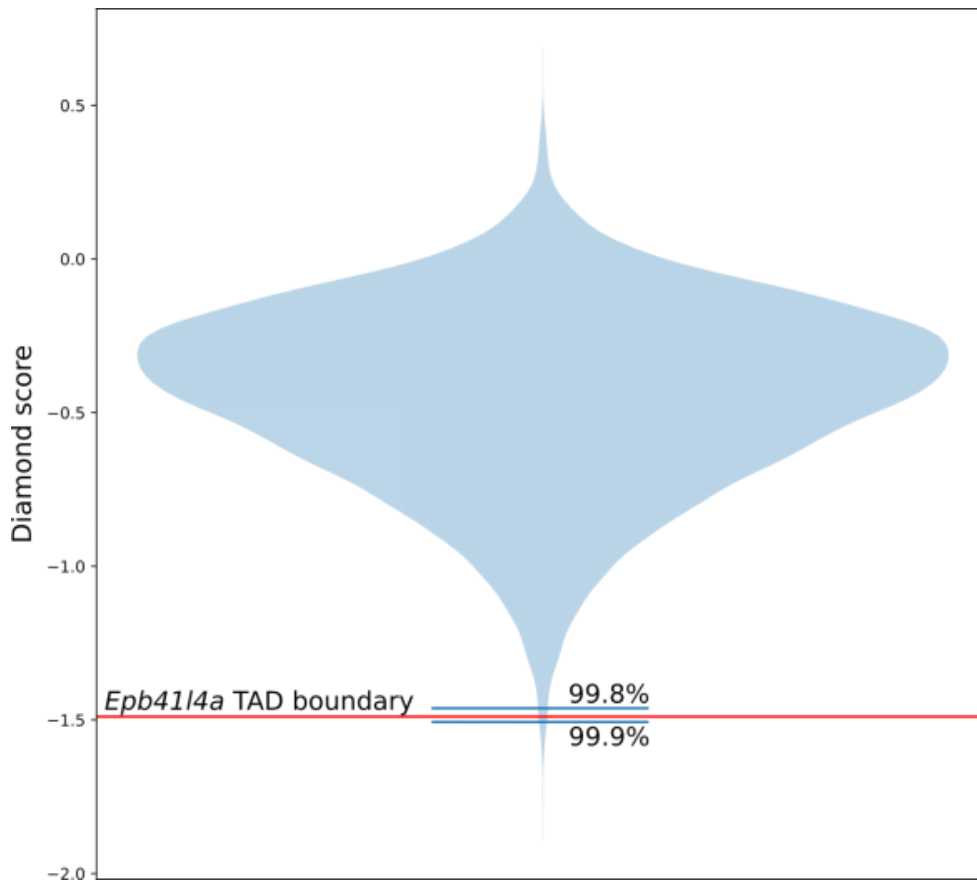


Рисунок 9. График распределения индекса инсуляции для всех границ ТАДов, выделяемых в геноме линии H1ESC (4D Nucleome project accession number 4DNES21D8SP8). Красной линией отмечен уровень инсуляции в границе ТАДов локуса *EPB41L4A*.

В базе данных DECIPHER мы идентифицировали четыре пациента с CNV, затрагивающими границу ТАДов в локусе *EPB41L4A* (Рис. 10Б). У двух из них были обнаружены делеции (пациенты 258489 и 4691), а у других двух — дупликации (340372 и 276088). Описанные у этих пациентов CNV также затрагивают тело гена *EPB41L4A*. У трёх из этих пациентов наблюдаются

неврологические аномалии, такие как дислексия, умственная отсталость, поведенческие отклонения и задержка развития речи, наряду с другими дефектами развития. Мы также выявили несколько меньших по размеру CNV, расположенных в теле гена *EPB41L4A* и ассоциированных с аналогичными неврологическими фенотипами (пациенты 287946, 289658, 326853, 339727 и 412086). Устойчивая и эволюционно консервативная пространственная архитектура этого локуса, а также его близость к генам, участвующим в развитии мозга, указывают на потенциальную патогенность этих CNV. Для изучения последствий нарушения границы TAD *EPB41L4A* мы решили воспроизвести описанные генотипы на мышинных моделях.

Поскольку выбранные локусы различались по характеру генетической интервенции, дальнейшая экспериментальная стратегия для них также отличалась. Для локуса *Slc29a3/Unc5b*, где предполагалось удалить отдельные сайты связывания CTCF, было необходимо напрямую проверить, приводят ли такие локальные мутации к изменению пространственной архитектуры хроматина. Поэтому для этого локуса был выполнен анализ трехмерных контактов методом capture Hi-C, позволивший охарактеризовать изменения топологической организации локуса на фоне полученных мутаций и сопоставить их с транскрипционным ответом в различных тканях. Для локуса *Epb41l4a*, напротив, были смоделированы крупные структурные перестройки, затрагивающие консервативную границу TADов, гарантированно приводящие к нарушению ее функции. В связи с этим основной акцент был сделан на оценке транскрипционных последствий этих перестроек, тогда как дополнительное картирование пространственной архитектуры мутантного локуса не являлось необходимым для проверки основной гипотезы данной части работы.

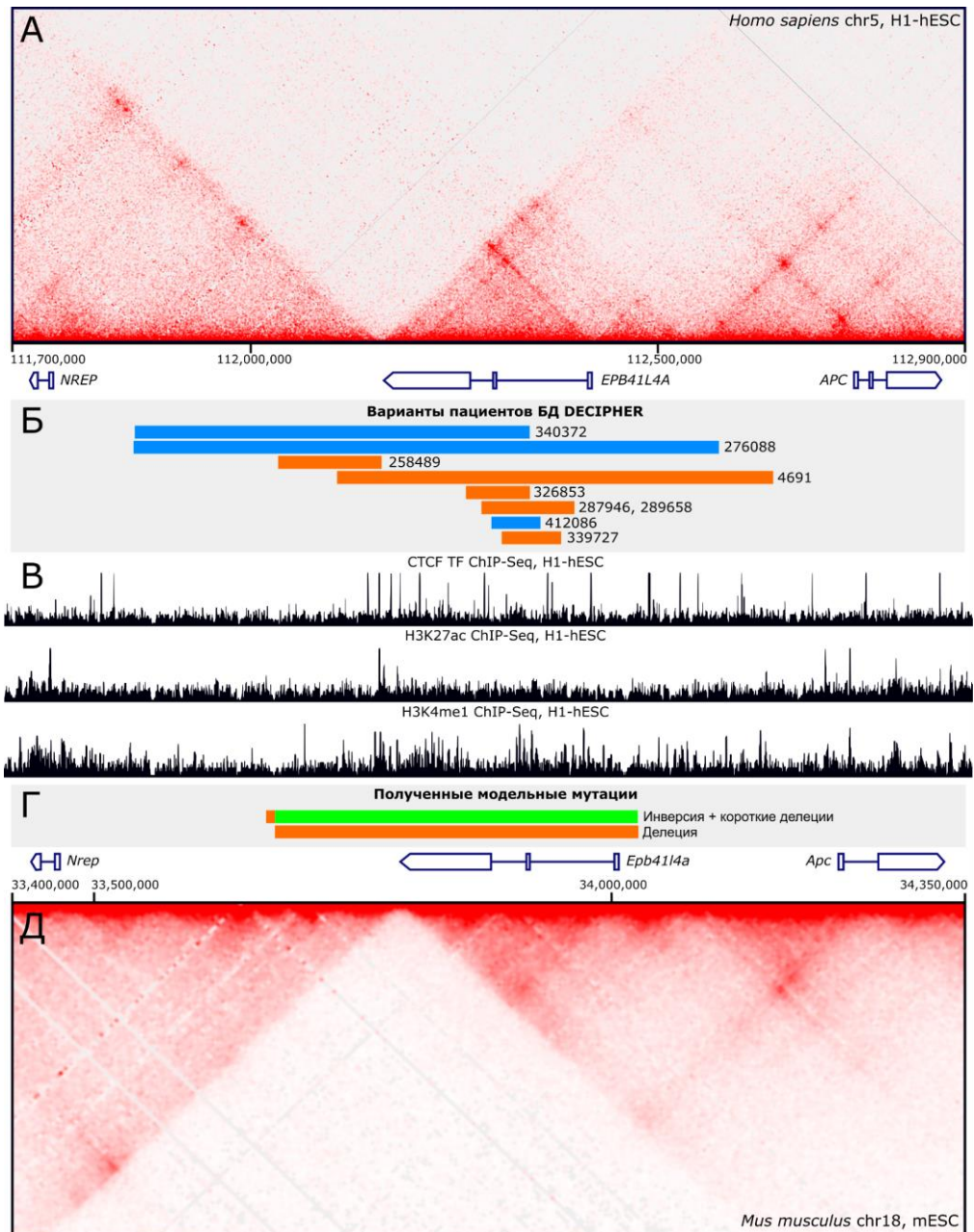


Рисунок 10. А — карта Hi-C локуса человеческого гена *EPB41L4A* и местоположения кодирующих генов. Б — координаты мутаций у пациента из базы данных DECIPHER. Синий — дупликации, оранжевый — делеции. В — профили ChIP-Seq для распределения связывания CTCF и гистоновых меток H3K27ac и H3K4me1 в локусе человеческого гена *EPB41L4A*. Г — координаты полученных мутаций у мышиных линий. Красный — делеция фрагмента ДНК, зеленый — инверсия. Д — карта Hi-C локуса мышинового гена *Erb41l4a* и

местоположения кодирующих генов.

## 2. Локус *Slc29a3/Unc5b*

### Получение мышей, несущих делеции сайтов связывания CTCF в локусе *Slc29a3/Unc5b*

Для изучения роли пространственной архитектуры генома в регуляции генной экспрессии в контексте тканевой и органной специфичности мы решили получить ряд генетически модифицированных линий мышей с локальными нарушениями структуры ТАДов. Первым модельным локусом, исследованным нами, стал локус *Slc29a3/Unc5b* генома мыши, имеющий выраженную консервативную геномную архитектуру и высокий уровень инсуляции между соседними ТАДами. Мы решили нарушить связывание CTCF в этих сайтах с помощью внесения небольших мутаций, не нарушающих другие регуляторные элементы. Вызванное этим снижение инсуляции между ТАДами, по нашей гипотезе, должно привести к взаимовлиянию регуляторных ландшафтов соседних ТАДов и нарушению регуляции экспрессии генов локуса.

Для получения мышей, несущих делеции сайтов связывания CTCF в локусе *Slc29a3/Unc5b*, мы применили метод геномного редактирования, основанный на системе CRISPR/Cas9. Мы удалили два межгенных сайта, индуцировав делецию около 5,5 тысяч п.н. (chr10: 60755585-60761088, mm10). Мы сконструировали две хнРНК (химерные направляющие РНК), которые таргетируют систему CRISPR/Cas9 на участки, фланкирующие удаляемые сайты, и дополнили их синтетической одноцепочечной ДНК-матрицей (ssODN), служащей донором для репарации по пути гомологичной рекомбинации, приводящей к возникновению делеции (Рис. 11Б). Конструкции вместе с белком Cas9 вводили в зиготы мышей C57BL/6J. Мы получили 21 животное, 19 из которых, как показало ПЦР-генотипирование, имели различные мутации в целевом локусе. Из них было идентифицировано

16 делеций (Приложения, Рис. 1), 4 инверсии, 1 дупликация (Приложения, Рис. 2) и 6 случаев INDEL достаточного размера для обнаружения с помощью электрофореза в агарозном геле (Приложения, Рис. 3 и 4). По крайней мере у четырех особей была выявлена генетическая мозаичность (при генотипировании мы обнаружили более двух разных аллелей). Мы выбрали животное, несущее целевую делецию, для возвратного скрещивания, и из полученного потомства вывели линию гомозиготных животных. Генотипирование с помощью ПЦР и секвенирование по Сэнгеру подтвердили, что делеция произошла между сайтами внесения DSB и, вероятно, представляет собой продукт репарации по пути HDR, поскольку полностью совпадает с последовательностью ssODN (Рис. 11B).

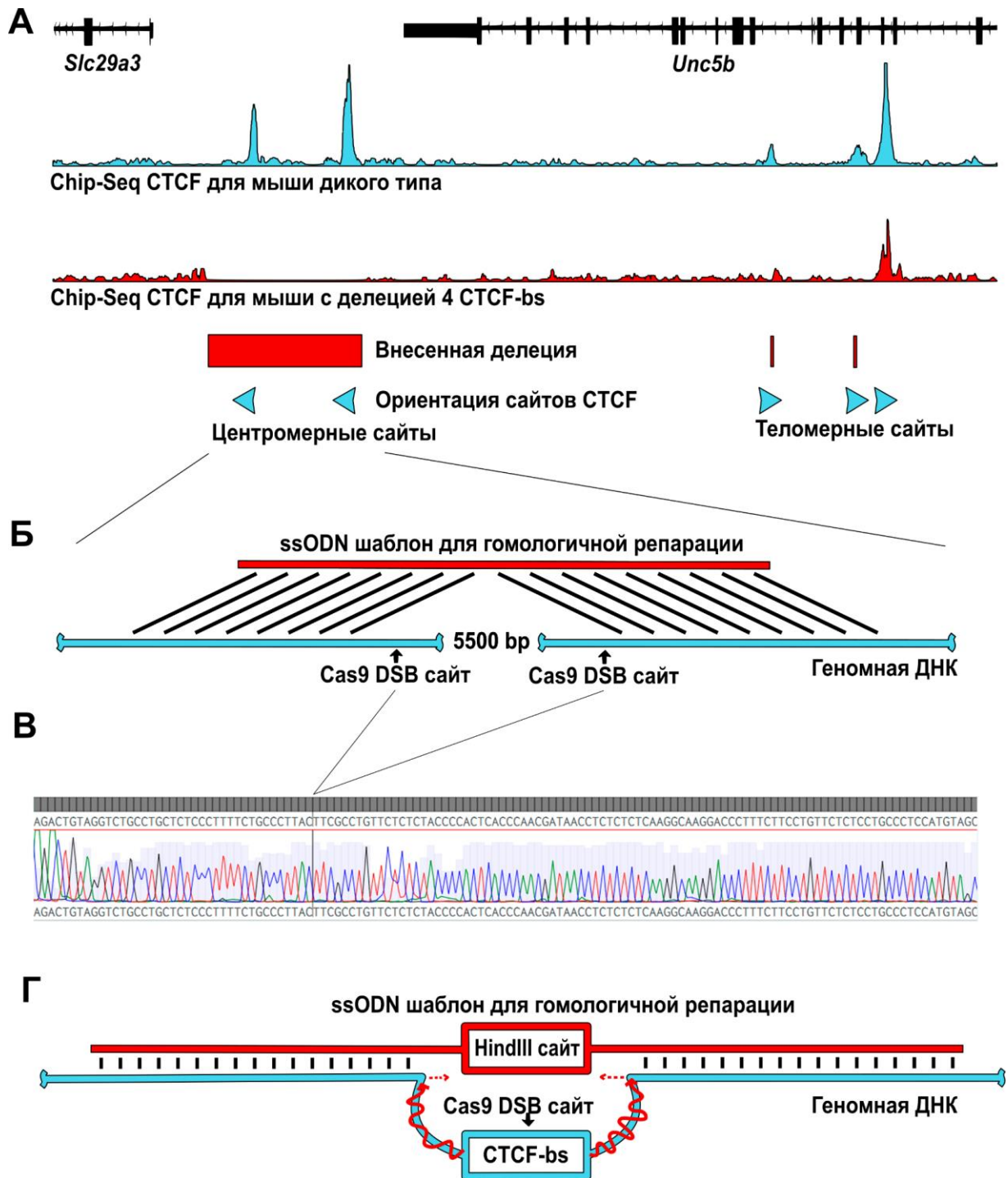


Рисунок 11. А — кластер сайтов связывания CTCF в мышинном локусе *Slc29a3/Unc5b* и полученные нами треки CTCF ChIP-Seq из ткани печени мышей, гомозиготных по аллелю dM1R1L2M2, и мышей дикого типа. Координаты мутаций показаны красными прямоугольниками, ориентация сайтов связывания CTCF показана красными треугольниками. Б — дизайн внесения делеции геномной области, содержащей кластер сайтов связывания CTCF. В — секвенирование по Сэнгеру полученной делеции. Полученная

последовательность выровнена на смоделированную ожидаемую последовательность делеции. Г — дизайн эксперимента по точечной модификации сайтов связывания CTCF. После внесения CRISPR/Cas9 разрыва в ДНК для направленного изменения последовательности с использованием шаблона ssODN необходима резекция свободных концов разрыва, негомологичных фланкирующим областям ssODN. В результате репарации по гомологичному механизму в целевой участок вносится мутация, нарушающая сайт связывания CTCF, и одновременно вносящая сайт узнавания фермента рестрикции *HindIII*, используемый для последующего генотипирования.

Сайты связывания CTCF часто обнаруживаются вблизи других регуляторных элементов, таких как промоторы и энхансеры, или находятся в небольших интронах генов, что делает невозможным удаление всего кластера сайтов связывания CTCF путем внесения протяженной делеции без нарушения других функциональных геномных элементов. Альтернативной стратегией модификации отдельных сайтов связывания CTCF является таргетирование их индивидуальными хнРНК, что приводит к индукции INDEL-мутаций и нарушению связывания CTCF с мутантным сайтом. В случае модификации кластера близко расположенных сайтов связывания CTCF этот подход теоретически может привести к возникновению хромосомных aberrаций из-за возможности перекрестного влияния соседних DSB. Таким образом, в настоящее время исследователи выполняют модификации отдельных сайтов связывания CTCF в кластере в режиме последовательного внесения модификаций в череде экспериментов [88], что является как трудо-, так и времязатратным. В этой работе мы впервые совершили попытку одновременно модифицировать несколько близко расположенных сайтов связывания CTCF, путем таргетирования их набором хнРНК и ssODN, направляющих репарацию по HDR, в зиготах мышей.

Такой подход мы решили применить для нарушения связывания с CTCF трех сайтов связывания CTCF, расположенных в интронах *Unc5b*. Для удобства назовем эти сайты связывания «левым», «средним» и «правым» в соответствии с их расположением по хромосомной оси от центромеры до теломеры длинного плеча (Рис. 8Г, 11А). Близость к другим регуляторным элементам гена *Unc5b* исключает использование простых схем редактирования, таких как удаление всех трех сайтов связывания CTCF путем внесения одной большой делеции. Мы сконструировали три хнРНК для CRISPR/Cas9, которые обеспечивали внесение DSB в пределах предсказанных мотивов CTCF. Кроме того, мы сконструировали ssODN, содержащие последовательность, комплементарную фланкирующим областям сайтов связывания CTCF, но заменяющую 15-30 п.н. мотива CTCF на сайт узнавания фермента рестрикции *HindIII* (A/AGCTT). Таким образом, репарация ДНК в присутствии ssODN должна приводить к замене корового мотива CTCF сайтом *HindIII*, что позволяет как нарушить связывание CTCF, так и проводить простое генотипирование полученных животных с помощью метода ПЦР-ПДРФ. Здесь уместно отметить, что в нашем дизайне репарация ДНК с использованием ssODN требует резекции до 30 нуклеотидов в месте внесения двуцепочечного разрыва CRISPR/Cas9 (Рис. 11Г).

Три хнРНК, три репарационных матрицы ssODN и белок Cas9, были инъецированы в зиготы, полученные от самок дикого типа C57BL/6J, оплодотворенных гетерозиготным самцом, несущим ранее полученную делецию двух межгенных сайтов CTCF. В этом эксперименте было получено и генотипировано 19 отредактированных животных.

Исходя из дизайна эксперимента, мы выбрали метод ПЦР-ПДРФ в качестве стратегии генотипирования. Для каждой мыши мы амплифицировали короткие фрагменты ДНК, содержащие целевые сайты связывания CTCF и обработали их ферментом *HindIII*. Поскольку мы ввели сайт узнавания *HindIII* в последовательность ssODN, положительный результат генотипирования

указывает на получение HDR-опосредованного мутантного аллеля. В результате оказалось, что такой мутационный вариант возник только у трех мышей (№ 1, 2 и 9) из 19 и только в среднем сайте связывания CTCF (Приложения, Рис. 5 и 6). Кроме того, мы наблюдали продукты ПЦР, отличающиеся по молекулярному весу от ожидаемых, для левого сайта CTCF мыши № 4, среднего сайта мыши № 10 и правого сайта мыши № 16. Эти продукты могли репрезентировать INDEL-мутации, произошедшие вследствие репарации двуцепочечных разрывов альтернативными путями.

Исходя из полученных данных, мы провели новый эксперимент по генотипированию ПЦР-ПДРФ, который позволил нам идентифицировать аллели, содержащие мутации, опосредованные альтернативными путями репарации. Для этого были выбраны ферменты рестрикции, сайты узнавания которых располагались на расстоянии порядка 5 п.н. от точки инициации DSB Cas9. Этот анализ показал, что мутации, отличные от продуктов HDR, гораздо чаще встречались у полученных отредактированных мышей. В частности, такие мутации наблюдались у 11/19 мышей для левого сайта связывания CTCF (Приложения, Рис. 7), у 12/19 для среднего (Приложения, Рис. 8) и у 1/19 для правого сайта (Приложения, Рис. 9). Более низкая эффективность редактирования правого сайта может указывать на относительно низкую *on-target* активность соответствующей хнРНК. Примечательно, что попытки обнаружить инверсии и делеции между мишенями CRISPR/Cas9 с помощью ПЦР с парами праймеров для ПЦР-ПДРФ в соответствующих сочетаниях показали отсутствие таких мутаций, несмотря на небольшое расстояние между сайтами-мишенями (3044 и 1097 п.н.). Мы предполагаем, что это может быть объяснено либо влиянием ssODN на путь репарации, либо летальностью таких делеций/инверсий из-за нарушения *Unc5b*.

Затем мы проанализировали сегрегацию полученных мутантных аллелей путем аналитического скрещивания с мышами C57BL/6J дикого типа. Всего было исследовано восемь отредактированных животных первого поколения,

несущих мутации левого и правого сайтов связывания CTCF, отобранных нами по результатам генотипирования. Мутантные аллели были унаследованы только потомством шести из восьми мышей, и у пяти из этих животных мы наблюдали проявления мозаицизма зародышевой линии (продукцию более чем двух типов зигот). Только три мыши (№ 8, 9 и 16) несли мутации в левом и среднем сайтах сайтов связывания CTCF в цис-положении, а мышь №16 также несла этот компаунд-аллель в сочетании с ранее полученной делецией двух межгенных сайтов связывания CTCF.

Поскольку чувствительность ПЦР-ПДРФ ограничена частотой мутантных аллелей и выбором ферментов рестрикции, мы решили использовать метод генотипирования, основанный на высокопроизводительном секвенировании NGS, чтобы лучше оценить разнообразие мутантных аллелей. Для этого были сконструированы пары праймеров, расположенные на расстоянии 105–120 п.н. от каждой из точек внесения DSB, что позволяет получать последовательность мутации при использовании секвенирования с длинами прочтений в 150 п.н. Анализ NGS позволил нам идентифицировать варианты мутаций, не обнаруживаемые с помощью ПЦР-ПДРФ (Рис. 12А-В). Мы не нашли никаких новых продуктов HDR в дополнение к тем, что были обнаружены с помощью ПЦР-ПДРФ; однако мы идентифицировали ряд различных делеционных аллелей, происхождение которых было, по предположению, опосредовано альтернативными путями репарации. Основываясь на наличии коротких гомологичных мотивов на границах, мы предположили, что многие из наблюдаемых делеций, скорее всего, опосредованы путем репарации MMEJ (microhomology-mediated end-joining). Мы наблюдали 24 случая MMEJ-опосредованных мутаций (считая каждый вариант у каждого животного), 23 случая NHEJ-опосредованных мутаций и только три случая HDR-опосредованных мутаций. Наблюдаемые делеции и их частота хорошо согласуются с предсказаниями программного обеспечения inDelphi [197].

Мы произвели предсказание аффинности мотивов CTCF для всех полученных последовательностей с помощью Web-инструмента, предоставляемого базой данных CTCFBSDB 2.0 [198,199]. Мы обнаружили, что большинство полученных последовательностей ДНК с делециями, опосредованных NHEJ и MMEJ, имели низкие оценки связывания CTCF, сравнимые с таковыми для HDR-опосредованных мутаций, которые заменяют основной мотив CTCF сайтом узнавания *HindIII*.

Мы выбрали животное № 16, несущее компаунд-аллель делеции двух межгенных сайтов связывания CTCF с мутациями левого и среднего сайтов связывания CTCF из внутригенного кластера, имеющих низкую оценку аффинности к CTCF, для получения гомозиготной по этому аллелю линии. На полученных гомозиготных животных мы провели анализ ChIP-Seq и подтвердили отсутствие связывания CTCF для мутантных сайтов (Рис. 11А). В результате мы получили гомозиготную линию мышей, несущих делеции [mm10] chr10:60755585-60761088 (два межгенных сайтов связывания CTCF, 5504 п.н.), chr10:60775689-60775698 (делеция 10 п.н. левого сайта внутригенного кластера), chr10:60778725-60778738 (делеция 14 п.н. со вставкой динуклеотида -AA- в центральном сайте кластера). На этой линии, а также на линии, несущей только делецию двух межгенных сайтов связывания CTCF, были выполнены все последующие эксперименты данной работы. В дальнейшем для краткости изложения мы будем называть полученные аллели dM1R1 и dM1R1L2M2.

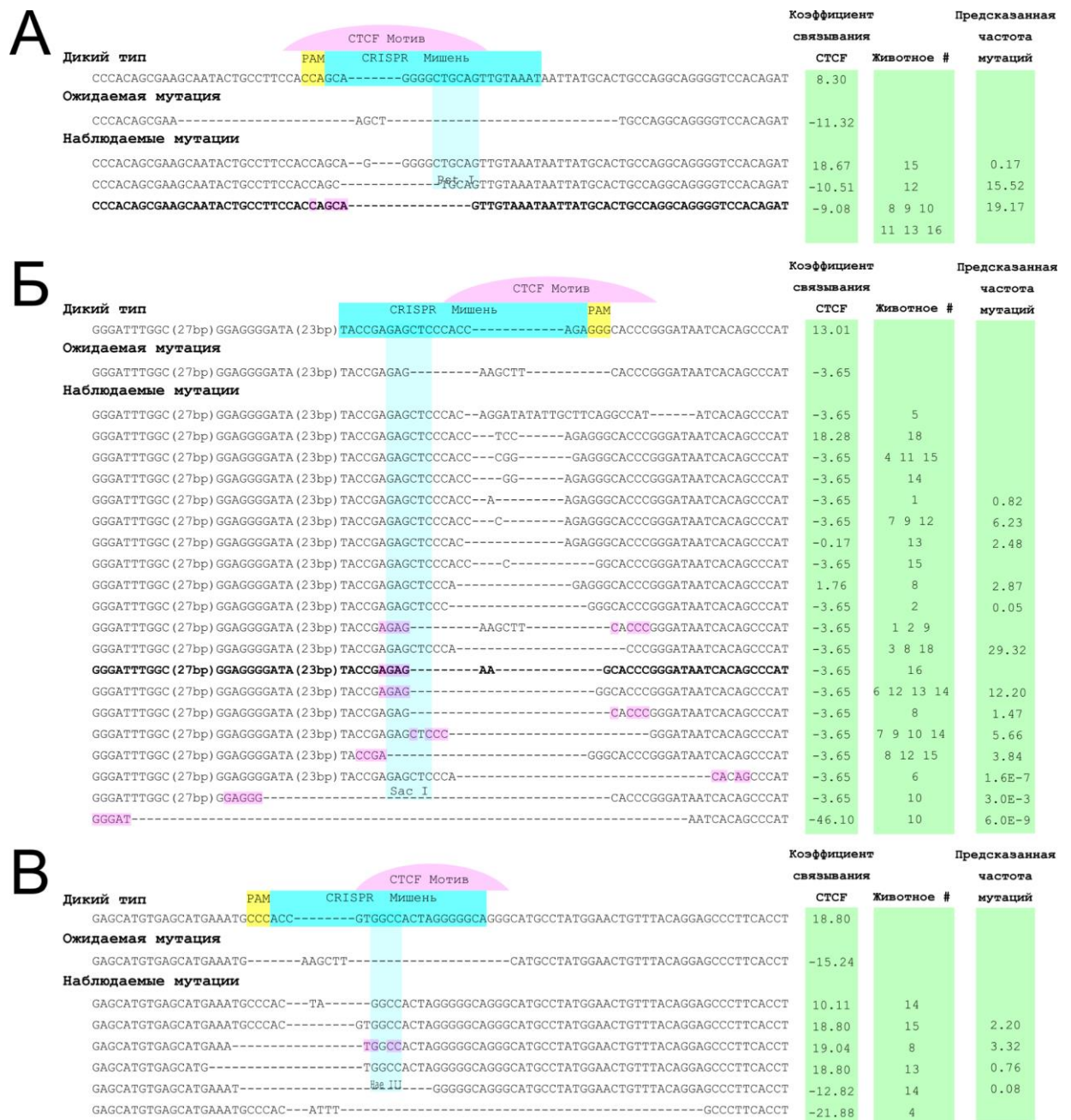


Рисунок 12. Результаты NGS-генотипирования для левого (А), среднего (Б) и правого (В) сайтов. ref. — последовательность аллеля дикого типа, exp. — ожидаемая последовательность опосредованной HDR мутации (идентичная последовательности ssODN). Мотивы микрогомологии на границах делеции выделены пурпурным цветом. Варианты, выбранные для получения гомозиготной линии (мышь № 16), выделены жирным шрифтом.

### **Изменения пространственной архитектуры локуса *Slc29a3/Unc5b* у мышей, гомозиготных по аллелю dM1R1L2M2**

Чтобы исследовать последствия индуцированных мутаций на пространственную архитектуру локуса *Slc29a3/Unc5b*, мы подготовили библиотеки Hi-C с обогащением с помощью панели гибридизационных олигонуклеотидов (capture Hi-C, cHi-C) для мозжечка, почек и печени, используя мышей, гомозиготных по делеции двух (dM1R1/dM1R1) и четырех (dM1R1L2M2/dM1R1L2M2) сайтов связывания CTCF и мышей дикого типа.

Пространственная архитектура локуса у мышей дикого типа почти идентична во всех анализируемых тканях (Рис. 13В,Г). ТАД *Slc29a3* содержит сложную систему вложенных петель. Три центромерных основания петель включают: дистальную границу ТАДа (расположенную между терминальной областью гена *Cdh23* и геном *Psap*) и два основания петель в теле гена *Cdh23*. Эти три сайта формируют петли с нижележащими основаниями, которые находятся в промоторе *Cdh23* и на границе ТАД, расположенной между генами *Slc29a3* и *Unc5b*. В ТАДе гена *Unc5b* петли формируются между его границей и промотором *Unc5b*, а также между двумя границами ТАДа. Дистальная граница расположена рядом с кластером генов длинных некодирующих РНК между генами *Unc5b* и *Sgpl1*. Заметно, что почти все описанные петли формируются сайтами связывания CTCF, имеющими конвергентную направленность, и, весьма вероятно, образуются и поддерживаются благодаря механизму протягивания петли когезином (Рис. 13А-Г).

У животных с делециями двух и четырёх сайтов CTCF карты контактов хроматина демонстрировали сходные изменения на суб-ТАДовом уровне во всех исследованных тканях. Сходство паттернов указывает на то, что ослабление границы за счёт удаления двух межгенных сайтов CTCF достаточно для появления детектируемых изменений контактов, тогда как

удаление дополнительной пары не влечет дальнейших заметных изменений трехмерной организации локуса (Рис. 13В,Г). Для обоих генотипов мы подтвердили, что петли, образованные удалённым кластером сайтов связывания CTCF и направленные в сторону ТАДа *Slc29a3*, нарушены во всех тканях (Рис. 13В,Г, черные заполненные стрелки). Напомним, что у животных дикого типа эти петли соединяют промоторную область *Slc29a3* (расположенную в непосредственной близости с границей ТАДов) с основаниями петель в теле гена *Cdh23* и геном *Psap*. В то же время петли, направленные в противоположную сторону, сохранились.

При этом, область генома от границы ТАДов до промотора *Cdh23*, содержащая ген *Slc29a3*, приобрела интенсивные пространственные взаимодействия со всем ТАДом *Unc5b* (Рис. 13В-Д). Эффект усиливается от печени к мозжечку, коррелируя с общей активностью локуса в ткани (Рис. 13Ж). Таким образом, можно сказать, что граница ТАДов сместилась к промотору *Cdh23*, переместив ген *Slc29a3* из пространственного окружения генов *Cdh23* и *Vsir* к пространственной близости с геном *Unc5b*.

Помимо нарушения внутренних петлевых структур ТАДов, мы отметили формирование новых, эктопических взаимодействий (Рис. 13В,Г, белые стрелки). Эти новые петли возникают между внутренними сайтами связывания CTCF гена *Cdh23* и дистальной границей ТАДа *Unc5b*, содержащей в непосредственной близости кластер генов некодирующих РНК, как правило не экспрессирующихся в исследованных тканях. Примечательно, что эти взаимодействия пересекают новую границу ТАДов, образовавшуюся в промоторной области гена *Cdh23*. Это позволило нам предположить, что их поддержание может включать дополнительные факторы, помимо механизма экструзии петли, например, сегрегация хроматиновых состояний. Еще более интересно то, что узор этих новых петель значительно различается между проанализированными тканями, позволяя предположить, что процессы, приводящие к возникновению этих контактов, также тканеспецифичны.

Таким образом, проведенный анализ показывает, что внесенные нами делеции сайтов связывания CTCF в локусе *Slc29a3/Unc5b* вызывают изменения в локальной архитектуре хроматина, нарушение функции границы между ТАДами и их внутренней петлевой организации. При этом наблюдается появление новых эктопических взаимодействий, формирование которых в норме подавлялось инсуляторной функцией кластера CTCF. Наблюдаемые изменения дают возможность ожидать реорганизацию взаимодействий между промоторами генов и регуляторными последовательностями, что может привести к изменениям в уровнях транскрипции генов локуса.

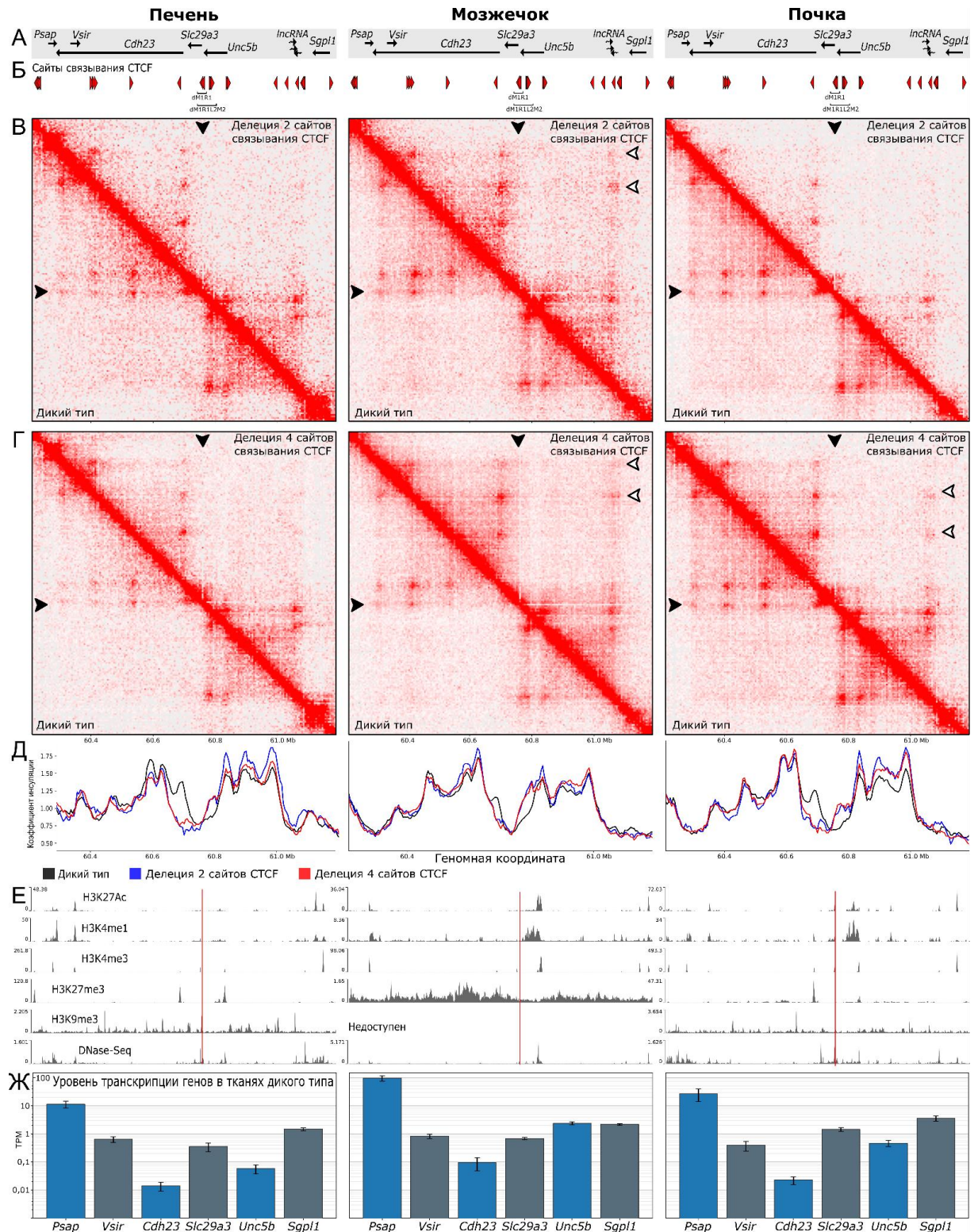


Рисунок 13. А — Положение и направление транскрипции генов локуса. Б — Положение и ориентация сайтов связывания CTCF, показаны удаленные сайты для аллелей dM1R1 и dM1R1L2M2. В, Г — Карты пространственных контактов локуса *Slc29a3/Unc5b* для печени, мозжечка и

почки в координатах chr10:60103000-61356000, mm10. Над главной диагональю представлены карты локуса для мышей, гомозиготных по аллелю dM1R1 (B) или dM1R1L2M2 (Г), под диагональю — для мышей дикого типа. Черные стрелки отмечают исчезнувшие петли модифицированной границы ТАДов, белые стрелки — новые петли, преодолевающие границу ТАДов. Д — Индекс инсуляции для мышей, гомозиготных по аллелю dM1R1 (синий), dM1R1L2M2 (красный), или мышей дикого типа (черный). Е — профили ChIP-Seq локуса *Slc29a3/Unc5b* по данным ENCODE. Вертикальная линия показывает внесенную протяженную 5 тыс. п.н. делецию межгенных сайтов связывания CTCF. Ж — Уровни транскрипции генов локуса для печени, мозжечка и почки мышей дикого типа согласно публичным данным ENCODE.

**Изменения экспрессии генов *Psap*, *Vsir*, *Cdh23*, *Slc29a3*, *Unc5b* и *Sgpl1* у мышей, гомозиготных по аллелю dM1R1L2M2, оцененные методом цифровой ПЦР**

Исходя из результатов предыдущих экспериментов [15,88], мы ожидали, что изменения экспрессии генов будут не более чем двукратными по сравнению с исходным уровнем. Поэтому нам было необходимо применить высокоточный метод для детекции уровней транскрипции РНК исследуемых генов. Конвенциональный дизайн количественной ПЦР в реальном времени (ПЦР-рв) подвержен множеству посторонних факторов, таких как влияние флуктуаций выбранных в качестве эндогенного контроля генов домашнего хозяйства, транс-эффектов, вызванных различиями в условиях содержания и индивидуальных особенностях мышей, неравномерного состава образцов по тканевому составу и прочее. Для устранения таких ошибок, а также увеличения чувствительности измерений, мы разработали новый дизайн эксперимента. На гибридах между полученными нами отредактированными животными (или мышами дикого типа, в качестве контроля) на фоне генотипа

C57BL/6 и линии мышей с генотипом подвида *Mus musculus castaneus* (CAST), с помощью метода цифровой ПЦР, мы провели моноаллельное измерение экспрессии генов исследуемого локуса. Таким образом, дискриминируя с помощью межлинейных SNP аллели, находящиеся в цис-положении к мутантной границе ТАДов локуса, или же к границе дикого типа, мы имели возможность нормировать уровень экспрессии аллеля, подверженного изменениям вследствие реорганизации локальной топологии хроматина, на аллель того же гена, находящегося в нормальном топологическом окружении.

Мы исследовали экспрессию 6 генов локуса (*Psap*, *Vsir*, *Cdh23*, *Slc29a3*, *Unc5b*, и *Sgpl1*) в следующем ряду органов: почка, печень, мозжечок, обонятельная луковица, мочевой пузырь, легкие, сердце, семенники, селезенка и тонкая кишка (Рис. 14). Мы обнаружили статистически значимые отличия для гена *Slc29a3* в виде повышения его экспрессии на 10-20% в мозжечке и печени, и снижение его экспрессии на 30% в почках. Экспрессия гена *Vsir* повысилась на 10% в мочевом пузыре и снизилась на 10% в обонятельных луковицах. Самым сильным обнаруженным изменением стало двухкратное снижение экспрессии гена *Cdh23* в мозжечке. Интересным является разнонаправленность изменений для одного и того же гена в различных тканях, что подтверждает нашу гипотезу о тканеспецифичности связи пространственной организации и генной экспрессии. На основании этих данных мы выбрали пять органов, в которых наблюдались статистически значимые изменения (почка, печень, мозжечок, обонятельная луковица и мочевой пузырь), для проведения дальнейших экспериментов по анализу транскрипционных эффектов. Также на основании данных результатов были выбраны три органа, в которых мы обнаружили разнонаправленные изменения гена *Slc29a3* для оценки изменений в них пространственной архитектуры хроматина, результаты чего представлены в предыдущей главе (Рис. 13В,Г).

Следует отметить, что данные, полученные с использованием цифровой ПЦР, требуют дополнительного подтверждения в силу низкой экспрессии некоторых генов в исследуемых органах, и в целом неоптимального дизайна системы, использующей в мультиплексной реакции пару конкурирующих за связывание с переменным сайтом ДНК-матрицы флуоресцентных зондов. Этот дизайн первоначально предложен для генотипирования диплоидных организмов. Результат этой операции дискретен и может быть легко интерпретирован, в отличие от континуального соотношения аллель-специфичных продуктов транскрипции, точность оценки количества которых весьма затруднительна, в том числе из-за недостаточного разделения результатов индивидуальных микрореакций на основании уровней флуоресценции по причине континуальной природы измеряемого параметра. Более точным методом оценки количества ДНК-матрицы в образце может быть метод секвенирования РНК нового поколения с применением уникальных молекулярных идентификаторов, который мы в дальнейшем применили для нашей системы.

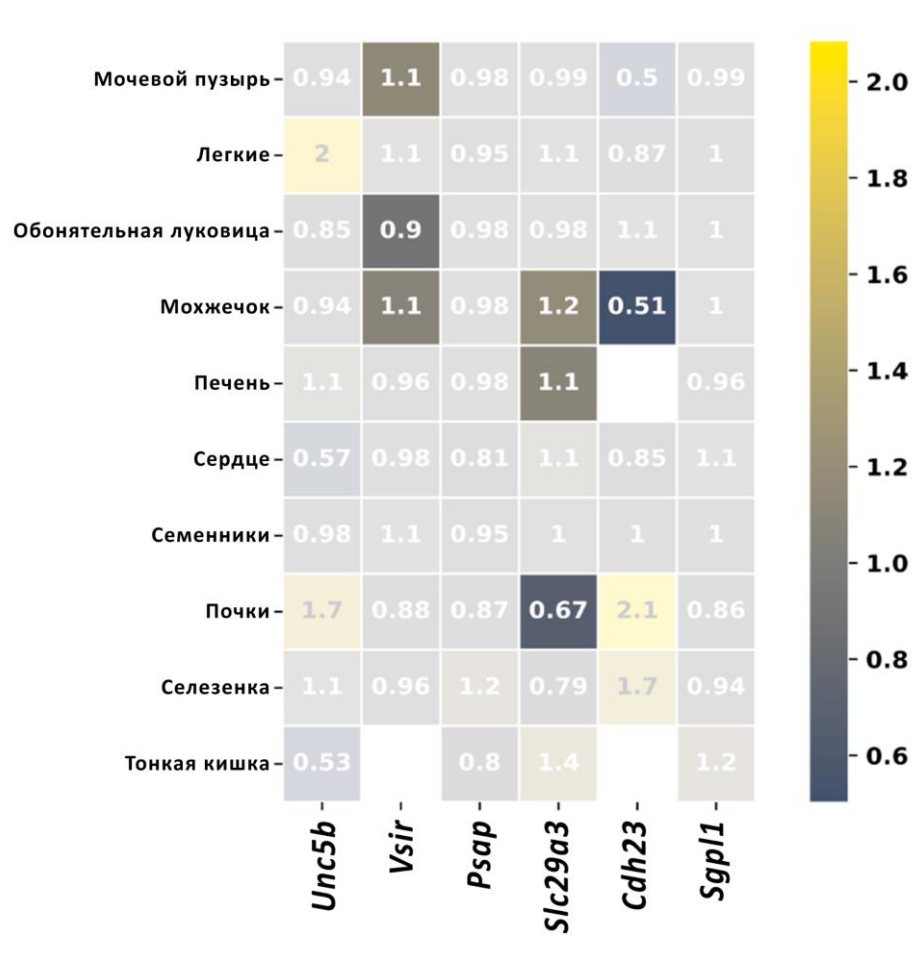


Рисунок 14. Результаты измерения экспрессии генов локуса *Slc29a3/Unc5b* у мышей, несущих аллель dM1R1L2M2, методом цифровой ПЦР. Представлены значения изменений экспрессии аллелей генов на фоне делеции сайтов связывания CTCF относительно аллелей с цис-окружением дикого типа в количестве раз. Ярким цветом отмечены результаты, проходящие статистическую проверку ( $FDR < 0.1$ )

**Изменения экспрессии генов *Psap*, *Vsir*, *Cdh23*, *Slc29a3*, *Unc5b* и *Sgpl1* у мышей, гомозиготных по аллелю dM1R1 и dM1R1L2M2, оцененные УМІ-опосредованным методом секвенирования целевой РНК**

Для подтверждения и дополнения картины транскрипционных изменений, полученных методом цифровой ПЦР, мы разработали инновационный подход для оценки количества ДНК-матрицы в образце на основе технологии секвенирования нового поколения с применением

уникальных молекулярных идентификаторов. Для количественного определения транскриптов различных аллелей в образцах гибридных животных мы провели обратную транскрипцию РНК с использованием UMI-меченных (12 вырожденных оснований), геноспецифичных праймеров. Последующее массовое параллельное секвенирование и анализ полученных данных позволяют дифференцировать аллельные продукты и количественно оценить их соотношение (Рис. 15).

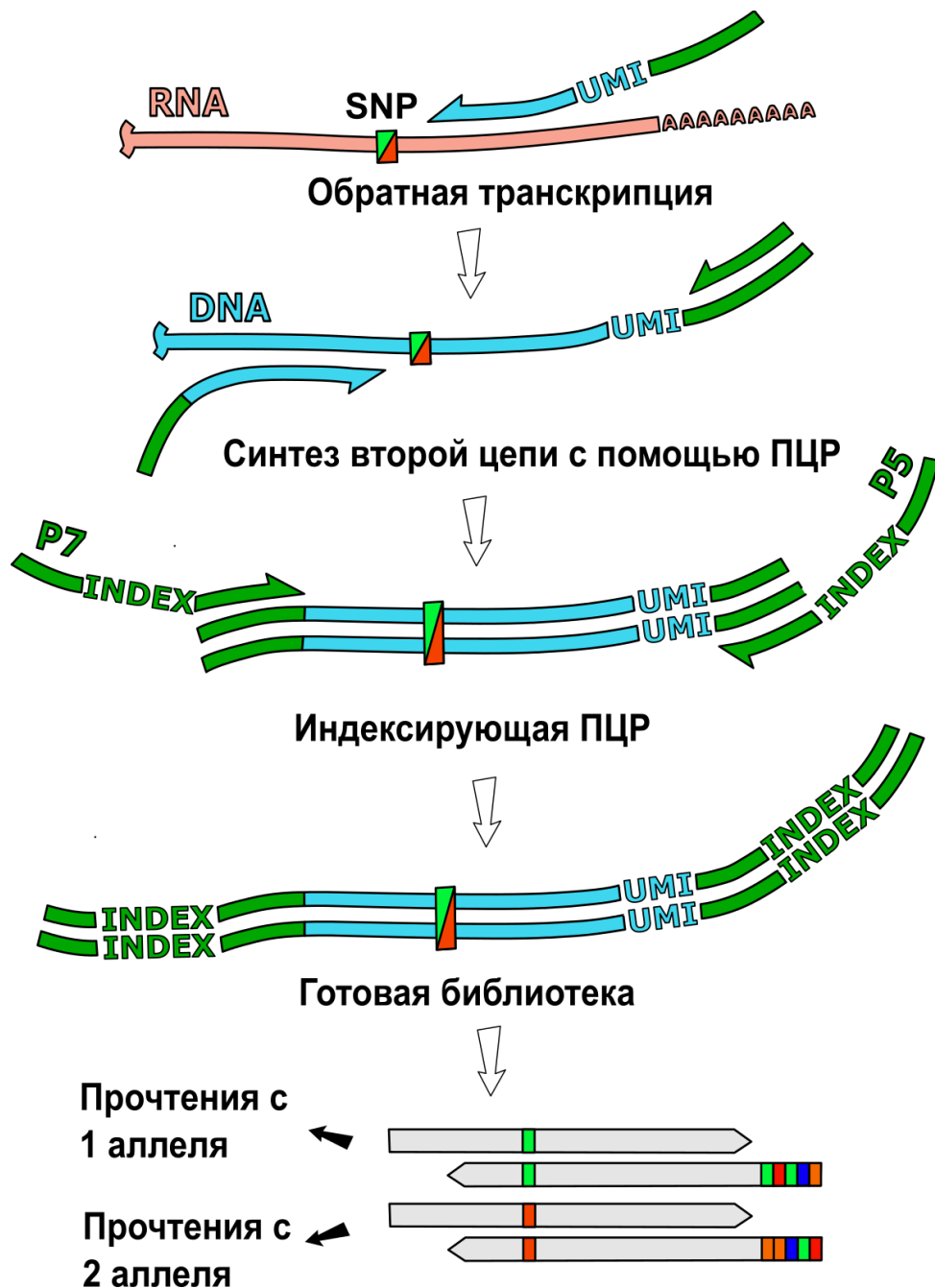


Рисунок 15. Графическая схема, иллюстрирующая метод оценки аллельного неравновесия у гибридов с помощью UMI-опосредованного таргетного РНК-секвенирования.

Хотя сравнение картин пространственных контактов в линиях с двумя и четырьмя удалёнными сайтами связывания CTCF не выявило заметных различий между мутантными генотипами, мы тем не менее проанализировали транскрипционные изменения для них как между собой, так и по сравнению с диким типом, поскольку ранее было показано, что по мере постепенного удаления сайтов связывания архитектурных белков на границах ТАДов происходит градиентное изменение межгенных регуляторных взаимодействий (88). Мы провели следующий эксперимент с использованием трех групп гибридов F1, полученных путем скрещивания самок CAST с 1) самцами дикого типа C57BL/6, и 2) гомозиготными мышами C57BL/6 с удаленными двумя (dM1R1/dM1R1) или 3) четырьмя (dM1R1L2M2/dM1R1L2M2) сайтами связывания CTCF в локусе *Slc29a3/Unc5b*, что позволило нам оценить эффект прогрессирующего ослабления функций границы ТАДов.

С помощью UMI-опосредованного таргетного РНК-секвенирования мы количественно оценили экспрессию шести генов в целевом локусе (*Psap*, *Vsir*, *Cdh23*, *Slc29a3*, *Unc5b* и *Sgpl1*) в пяти органах (почке, печени, мозжечке, обонятельной луковице и мочевом пузыре). Мы выявили 20 случаев значимых изменений экспрессии генов, вызванных делециями сайтов связывания CTCF среди всего 60 проанализированных случаев (Рис. 16А,Б). Как мы и ожидали из предыдущих исследований, наблюдаемые изменения в транскрипции генов были умеренными, не превышая 50%. Также подтвердился характер паттерна изменений, варьирующий не только в масштабе, но и в направлении изменений уровня экспрессии генов.

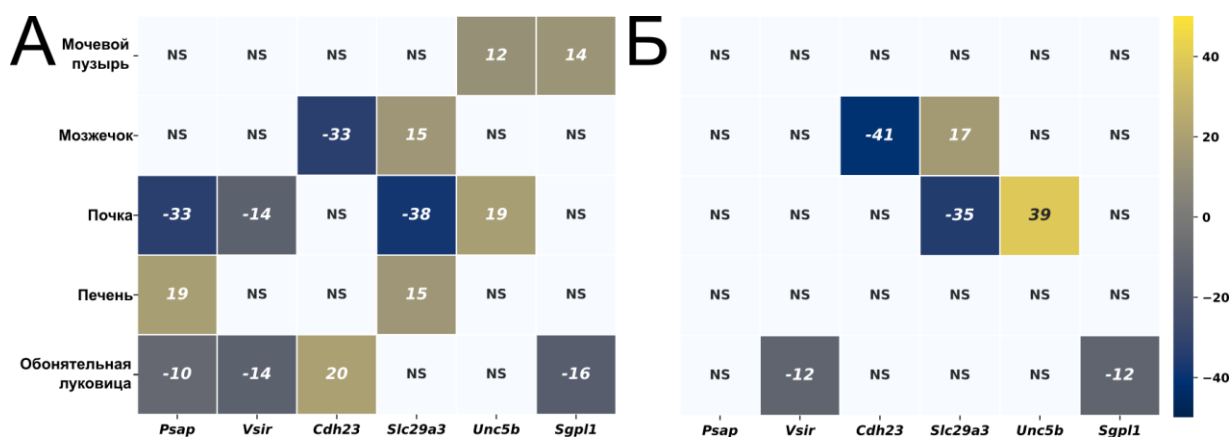


Рисунок 16. Результаты измерения экспрессии генов локуса *Slc29a3/Unc5b* у мышей, несущих аллели dM1R1 (А) или dM1R1L2M2 (Б), полученные с помощью УМІ-опосредованного таргетного РНК-секвенирования. Представлены значения изменений экспрессии аллелей генов на фоне делеции сайтов связывания CTCF относительно аллелей с цис-окружением дикого типа в процентах. Отмечены только результаты, проходящие статистическую проверку ( $FDR < 0.1$ ).

Нарушение CTCF-зависимой архитектуры локуса *Slc29a3/Unc5b* сопровождалось тканеспецифическими изменениями экспрессии исследуемых генов. В почке наблюдалось снижение экспрессии *Slc29a3* на 40% и повышение экспрессии *Unc5b*, которое варьируется от 20% у животных с двумя удаленными сайтами связывания CTCF до 40% у животных с четырьмя (Рис. 16А,Б). В мозжечке было выявлено увеличение экспрессии *Slc29a3* на 15% в мозжечке, тогда как снижения экспрессии *Unc5b* обнаружено не было. В печени, где оба гена характеризуются низким уровнем экспрессии (Рис. 13Ж), наблюдалось слабое повышение экспрессии *Slc29a3*. Кроме того, в мозжечке было подтверждено снижение экспрессии *Cdh23* на 30-40% в зависимости от генотипа, причем данные Hi-C указывали на появление эктопических дальних взаимодействий между телом *Cdh23* и дистальной границей ТАДа *Unc5b*. Ген *Psap* демонстрировал различия между аллелями с двумя и четырьмя удаленными сайтами CTCF: его экспрессия заметно менялась у животных с двумя удаленными сайтами, но не изменялась у

животных с четырьмя удаленными сайтами. Наконец, мы наблюдали небольшое по амплитуде, хотя и статистически значимое, увеличение экспрессии генов *Unc5b* и *Sgpl1* в мочевом пузыре, увеличение *Cdh23* и снижение *Vsir* и *Sgpl1* в обонятельной луковице.

### 3. Локус *Erb41l4a*

#### Получение мышей, несущих хромосомные перестройки, затрагивающие границу ТАДов в локусе *Erb41l4a*

Во второй части нашего исследования мы сосредоточились на консервативной между человеком и мышью границе ТАДов между клинически значимыми генами *NREP* и *APC*, расположенной в терминальной области гена *EPB41L4A*.

Для этого локуса описано несколько структурных аномалий хромосом человека, затрагивающих эту границу, часто сопутствующих неврологическим нарушениям (Рис. 10Б). Мы задались целью смоделировать на мышах перестройки, приводящие к нарушению функции границы ТАДов в этом локусе, и проанализировать их влияние на транскрипцию генов локуса.

Как было показано нами в предыдущих экспериментах, а также и другими группами исследователей [88,146], границы ТАДов обладают определённой устойчивостью к делециям сайтов связывания CTCF, их формирующих. Таким образом, наша задача заключалась в устранении всех сайтов связывания CTCF, потенциально участвующих в организации пространственной архитектуры хроматина в этом локусе, или способных компенсаторно её выполнять после внесения мутации. Используя публичные данные ENCODE, мы идентифицировали массивный кластер сайтов связывания CTCF, находящийся в пределах 100 тысяч п.н. от 3'-терминальной области гена *Erb41l4a*, и множество диффузно расположенных сайтов в теле этого гена (Рис. 10В). Мы решили индуцировать делецию, приводящую к удалению всей этой области и, предположительно, потере инсуляции между топологическими доменами, содержащими гены *Nrep* и *Apc*, при этом избегая нарушений в прилегающих геномных регионах, потенциально содержащих регуляторные цис-элементы этих генов. Для этого мы использовали систему CRISPR/Cas9, нацелив её на геномные сайты chr18:33673963-33673983 и

chr18:34025700-34025720 (mm10) для внесения двуцепочечных разрывов и индукции перестройки.

Мы ввели молекулярные компоненты системы CRISPR/Cas9 (в частности, мРНК Cas9 и две направляющие РНК, нацеленные на фланкирующие межгенные области *Epb4114a*) в зиготы мышей линии C57BL/6J с помощью пронуклеарной микроинъекции. В результате эксперимента мы получили четыре животных F0, три из которых, согласно данным ПЦР-генотипирования, несли целевые мутации в локусе *Epb4114a*. В результате генотипирования и последующего возвратного скрещивания было показано, что животное №1 (самка) несет делецию с ожидаемыми координатами, которая присутствует и в клетках зародышевой линии и способна передаваться потомству. От скрещивания самки №1 с самцом дикого типа было получено жизнеспособное и фертильное потомство, среди которого были гетерозиготные носители делеции, скрещиванием которых друг с другом была получена гомозиготная линия носителей делеции. Животное №2, самка с выявленными тремя типами перестроек (делецией, инверсией и дупликацией), что свидетельствует о мозаичности этого животного, не дала жизнеспособного потомства и погибла. Животное №3, самец без обнаруженных перестроек, также не передал мутаций своему потомству, в результате чего скрещивания были прекращены и линия была исключена из исследования. Наконец, животное №4, самец с инверсией, дал плодовитое потомство, среди которых были получены гетерозиготные носители инверсии, и в последующих скрещиваниях была получена гомозиготная линия. Генетически модифицированные гомозиготные мыши, несущие делецию либо инверсию в локусе *Epb4114a*, не проявляли заметных аномалий в развитии и нормально размножались.

Точность геномного редактирования была подтверждена методом секвенирования по Сэнгеру последовательностей стыков двуцепочечных разрывов полученных мутаций. Интересно, что одна из точек стыков инверсии

не детектировалась при ПЦР-генотипировании. После смещения сайта связывания праймера и последующего секвенирования мы выявили делецию длиной около 7,8 тыс. п.н. вблизи 3'-конца точки разрыва инверсии и делецию 27 п.н. на 5'-конце, что привело к удалению сайта гибридизации одного из праймеров ПЦР-системы генотипирования. Утраченный генетический материал находился в межгенной области между генами *Epb4114a* и *Nrep* и, по доступным аннотациям, не включал известных функциональных элементов, поэтому маловероятно, что данные дополнительные изменения могли значительно модифицировать транскрипционный и эпигенетический эффект внесенных мутаций. Мы полагаем, что эти последовательности были утрачены в процессе репарации ДНК, что согласуется с ранее наблюдавшимися побочными эффектами при использовании CRISPR/Cas9, приводящими к нежелательным мутациям в зиготах мышей (170).

Кроме того, для выявления нежелательных мутаций, вызванных нецелевой активностью системы CRISPR/Cas9, мы провели секвенирование по Сэнгеру всех потенциальных нецелевых сайтов внесения двуцепочечных разрывов на расстоянии 50 сантиМорган (50 млн. п.н.) от региона внесения мутации. Данный анализ был проведен на животных полученных гомозиготных линий (и непосредственно на всех животных, использованных в экспериментах по оценке экспрессии) и показал отсутствие нецелевых мутаций, могущих внести неучтенное влияние на транскрипционную активность генов в эксперименте.

В итоге мы получили две модельные линии мышей на генетическом фоне C57BL/6J. Одна линия несёт делецию длиной около 350 тыс. п.н. с координатами (chr18:33673967-34025702, mm10), а другая — инверсию того же региона (chr18:33673967-34025702, mm10), сопровождаемую делецией длиной около 7,8 тыс. п.н. (chr18:33666190-33673967, mm10) и 27 п.н. (chr18:34025676-34025702, mm10) (Рис. 10Г).

Дальнейшая экспериментальная логика анализа этого локуса отличалась от подхода, использованного для локуса *Slc29a3/Unc5b*. В первом случае индуцированные мутации представляли собой локальные делеции отдельных сайтов связывания CTCF, поэтому их последствия для пространственной архитектуры нельзя было предсказать только на основании структуры аллеля. Это потребовало прямой оценки трехмерных контактов хроматина методом capture Hi-C. В случае локуса *Epb41l4a* объектом моделирования были крупные хромосомные перестройки, затрагивающие область сильной и консервативной границы ТАДов и гарантированно нарушающие ее функцию. Положение этой границы и ее сохранность между человеком и мышью были предварительно охарактеризованы на основании публичных Hi-C-данных, поэтому основным вопросом этой части работы стал анализ транскрипционных последствий нарушения границы, а картирование пространственной архитектуры мутантного локуса не проводилось, поскольку оно не было необходимо для проверки основной гипотезы этой части исследования.

### **Транскрипционный эффект хромосомных перестроек, затрагивающих границу ТАДов в локусе *Epb41l4a***

Для оценки изменений в экспрессии генов локуса *Epb41l4a* мы провели эксперимент RNA-seq на самцах мышей, несущих полученные делеции и инверсии гена *Epb41l4a* в гомозиготном состоянии, используя самцов C57BL/6J дикого типа в качестве контроля. Мы выбрали ткани, где гены *Epb41l4a* и *Nrep* мышей демонстрируют контрастные паттерны экспрессии. В частности, мозжечок был выбран как орган с наивысшим уровнем экспрессии гена *Nrep*, а обонятельная луковица — как ткань с наивысшей экспрессией гена *Epb41l4a*, согласно данным ENCODE.

В первую очередь мы сосредоточились на анализе изменений уровня транскрипции генов, расположенных вблизи целевого региона. В

соответствии с нашими ожиданиями, у мышей с делецией экспрессия *Epb4114a* не детектировалась (т.к. делеция захватывает ген целиком) (Рис. 10Г, 17Б). При инверсии же мы обнаружили, что экспрессия *Epb4114a* в мозжечке увеличилась более чем в четыре раза (Рис. 17Б). Этот феномен хорошо объясняется механизмом кражи энхансера: инверсия перемещает промотор гена *Epb4114a* из ТАДа, содержащего ген *Apc*, в ТАД, содержащий ген *Nrep* (Рис. 17А). В итоге, новое регуляторное цис-окружение активно экспрессирующегося в мозжечке гена *Nrep* активирует обычно репрессированный в этом органе ген *Epb4114a*. Механизм кражи энхансера ранее был описан в большом количестве экспериментов, хотя и не является универсальным правилом для взаимодействия промоторов и энхансеров после реорганизации топологических доменов, так как некоторые промоторы могут сопротивляться эктопическому регуляторному влиянию регуляторных элементов [14,16,200–202].

Мы также отметили, что экспрессия *Nrep* в мозжечке снизилась на 70% как в случае инверсии, так и в случае делеции, и аналогично снизилась на 30% в обонятельной луковице у мышей с инверсией (Рис. 17В). Возможно, это снижение связано с введением новых регуляторных элементов в цис-окружение гена *Nrep*, которые входят в конкурентные взаимодействия с его исходным регуляторным ландшафтом. В мозжечке эти элементы могут быть репрессорами, в норме подавляющими экспрессию гена *Epb4114a*, что после инверсии может приводить к снижению экспрессии *Nrep*. Данные для обонятельной луковицы сложнее интерпретировать; мы можем предложить гипотезу, состоящую в том, что активный промотор *Epb4114a* после инверсии взаимодействует с энхансерами *Nrep*, перераспределяя контактную частоту между промоторами. Гипотеза о конкуренции промоторов за взаимодействие с энхансерами и ограниченности активирующей активности энхансера остаётся спорной, однако полученные нами модельные линии могут стать ценным инструментом для изучения этого вопроса.

Интересно, что экспрессия гена *Apc*, в норме инсулированного от *Nrep* границей ТАДов в терминальной области гена *Erb41l4a* (Рис. 17Г), не показала значимых изменений в уровнях транскрипции ни при делеции, ни при инверсии.

В целом, в ходе нашего анализа данных РНК-секвенирования мы выявили 23 дифференциально экспрессируемых гена (ДЭГ) в мозжечке и 30 ДЭГ в обонятельной луковице у мышей с делецией. Модель с инверсией показала 55 ДЭГ в мозжечке и 78 ДЭГ в обонятельной луковице, со строгими параметрами фильтрации – изменением уровня экспрессии более чем в два раза и скорректированным  $p$ -значением  $p\text{-adj} < 0,001$ . Эти результаты указывают на возможное наличие транс-эффектов, вызванных изменением экспрессии *Nrep* и *Erb41l4a* из-за нарушения локальной пространственной организации. Модель с инверсией демонстрирует более выраженные эффекты, что согласуется и с масштабом наблюдаемых цис-эффектов, вызванных внесением соответствующих мутаций. Мы также обнаружили изменения экспрессии генов, находящихся на той же хромосоме на расстоянии более 1,5 млн. п.н. от региона внесения мутаций.

Такие эффекты теоретически могут быть опосредованными изменениями локальной пространственной организации хроматина, или же опосредованными транс-эффектами, включающими изменения белок-белковых взаимодействий в каскадах генных взаимодействий, включающих цис-измененные гены. Дискриминация между двумя этими возможностями для столь удалённых генов остается сложной задачей, требующей специального дизайна эксперимента, наподобие эксперимента с гибридами и оценкой аллельного дисбаланса, описанного нами для локуса *Slc29a3/Unc5b*. Максимальное расстояние, зафиксированное для распространения эффекта цис-регуляторных взаимодействий при нарушении структуры ТАДов, составляет 1,45 млн. п.н., что делает гипотезу о более дальних эффектах менее вероятной [37].

В заключение, наши данные свидетельствуют о том, что мутации, изменяющие границу ТАДов в локусе *Epb4114a*, могут привести к значительной дисрегуляции транскрипции гена *Nrep*. Хотя в рамках этого исследования поведенческий фенотип мутантных мышей не оценивался, исходя из того, что у мыши *Nrep* ассоциирован с неврологическими и поведенческими фенотипами, мы предполагаем, что снижение его экспрессии может способствовать развитию патологии у человека. По данным DECIPHER, перестройки, перекрывающие границу ТАДов в локусе *Epb4114a*, часто сопровождаются неврологическими симптомами у пациентов. Наш эксперимент показывает, что такие перестройки потенциально могут приводить к снижению уровня экспрессии гена *Nrep*, опосредованному нарушениями локальной архитектуры хроматина. Это позволяет выдвинуть гипотезу о взаимосвязи таких перестроек и неврологических нарушений развития, однако для проверки этой гипотезы требуются дальнейшие масштабные эксперименты.

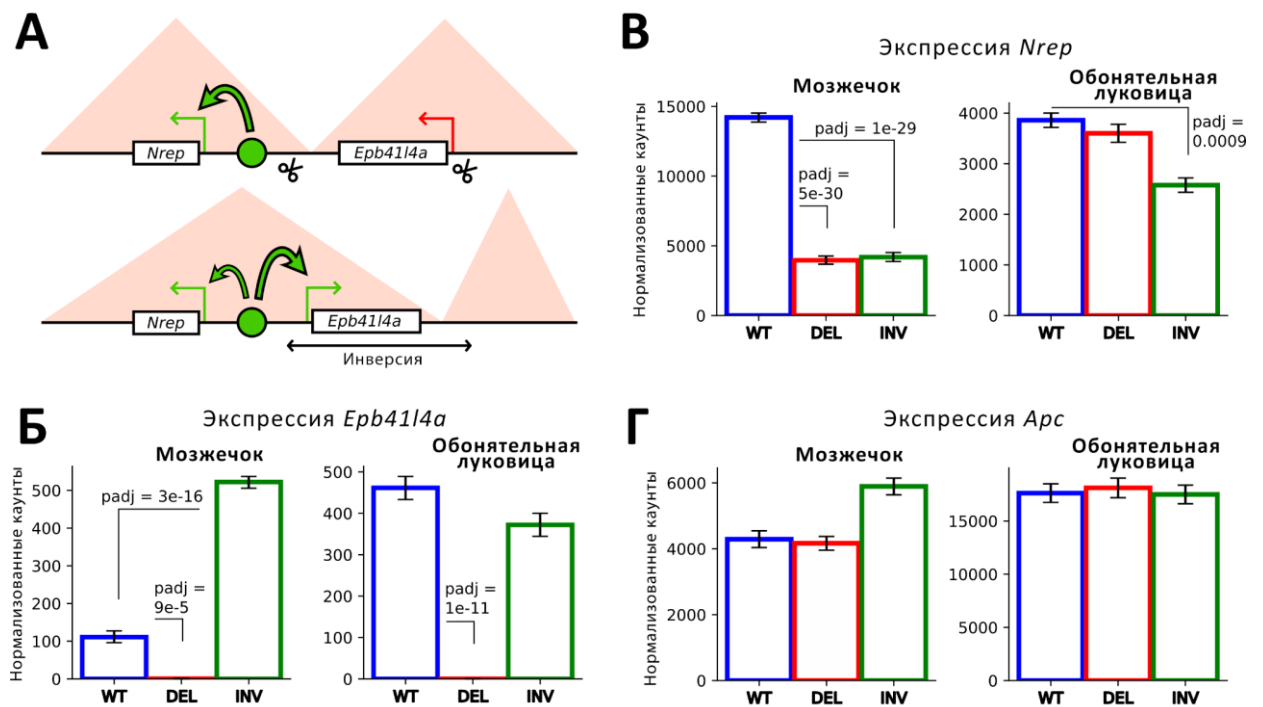


Рисунок 17. А — Схема механизма кражи энхансера на модели инверсии в локусе *Epb4114a*. Зеленый круг обозначает энхансер *Nrep*, символами

ножниц отмечены мишени CRISPR/Cas9. Зеленым или красным цветом промоторов отмечена их высокая или низкая транскрипционная активность, соответственно. Б — уровни экспрессии *Epb41l4a* в мозжечке и обонятельной луковицы для всех экспериментальных групп. В — уровни экспрессии *Nrep* в мозжечке и обонятельной луковицы для всех экспериментальных групп. Г — уровни экспрессии *Arc* в мозжечке и обонятельной луковицы для всех экспериментальных групп. WT — уровень экспрессии у мышей дикого типа, INV — у мышей с инверсией, DEL — с делецией целевого участка. Ось Y на гистограммах экспрессии генов представляет нормализованные значения по DeSeq2.

## Обсуждение

Данная работа посвящена вопросу о тканеспецифичности эффекта генетических мутаций некодирующих регионов генома, приводящих к нарушению пространственной архитектуры хроматина. В отличие от ранних исследований, выполненных на генах развития, имеющих узкоспецифичный пространственно-временной паттерн экспрессии [14,15,24,195], мы сфокусировались на локусах, содержащих гены, широко экспрессирующиеся во многих тканях взрослого организма. С помощью технологий редактирования генома нами получены четыре гомозиготные линии генетически модифицированных мышей, несущих мутации, приводящие к нарушению функции границы ТАДов. На этих модельных животных мы показали, что эффекты этих мутаций значительно различаются между органами, как с точки зрения изменений в пространственной организации, так и транскрипционной активности близлежащих генов, и зависят от эпигенетического ландшафта локуса, устанавливающегося в процессе развития и дифференцировки тканей.

В первой части данной работы мы получили две линии мышей, несущих делеции двух или четырёх сайтов связывания CTCF между генами *Slc29a3* и *Unc5b*. Мы продемонстрировали, что нарушение функции границы между двумя ТАДами, образованной этими сайтами связывания CTCF, привело к сложной реорганизации пространственных контактов локуса, изменению характера пространственных взаимодействий региона гена *Slc29a3* и образованию новых эктопических петлевых структур, механизмы формирования которых пока неизвестны. Хотя эта реорганизация не столь радикально отличается между проанализированными органами, она вызывает тканеспецифические и разнонаправленные изменения экспрессии генов локуса.

В процессе создания линий генетически модифицированных мышей, помимо стандартного дизайна внесения протяженных делеций, мы использовали мультиплексную модификацию соседних сайтов связывания CTCF с помощью генетического редактирования системой CRISPR/Cas9. Наша стратегия предполагала использование ssODN в качестве матрицы для гомологичной рекомбинации, имеющей преимущество перед другими методами, заключающееся в предсказуемости результатов мутагенеза [196]. Однако эта стратегия оказалась малоэффективной, так как подавляющее количество событий репарации произошло по пути микрогомологичного соединения концов (ММЕJ) [197]. Тем не менее, нам удалось получить аллель с ММЕJ-опосредованными мутациями, которые нарушают одновременно два сайта связывания CTCF. Таким образом, мы считаем, что индукция коротких делеций с участием ssODN-матриц гомологии и необходимостью резекции концов двуцепочечного разрыва (Рис. 11Г) оказалась малоэффективной, по крайней мере в условиях применения на мышинных зиготах, где путь гомологичной рекомбинации оказывается подавлен с преимущественным выбором механизма ММЕJ. Несмотря на это, для данной задачи может быть применен дизайн мутагенеза с использованием пути ММЕJ, показывающий хорошо предсказуемый спектр генерируемых мутаций [197]. С определенными оговорками, этот подход может быть применен даже в мультиплексном режиме.

Мы продемонстрировали, что полученные нами делеции нарушают границу между топологическими доменами и изменяют внутреннюю петлевую структуру ТАДа *Slc29a3*. Хотя в проанализированных органах слияния доменов не произошло, граница сместилась к внутреннему сайту связывания CTCF в промоторе гена *Cdh23*. Этот сдвиг приводит к перемещению региона гена *Slc29a3*, в новое цис-окружение ТАДа *Unc5b*. Кроме того, мы обнаружили появление новых крупномасштабных пространственных контактов, отражающих образование хроматиновых

петель, преодолевающих новую границу между ТАДами. Происхождение этих архитектурных элементов не может быть легко объяснено с помощью каких-либо известных механизмов на данный момент, и ранее такие элементы не были описаны как ожидаемый эффект делеций границ топологических доменов. Кроме того, в отличие от предыдущих исследований, наблюдавших градиентный эффект при удалении разного количества формирующих границу сайтов [88], мы не зафиксировали заметного различия в топологической структуре локуса у мышей, несущих делеции двух и четырех сайтов связывания CTCF.

В этом исследовании мы разработали и применили основанный на высокопроизводительном секвенировании метод, позволяющий точно количественно оценить изменения экспрессии генов, вызванные реорганизацией структуры ТАДов (и, в целом, других цис-элементов). Мы показали, что предложенный подход позволяет оценивать небольшие изменения генной экспрессии и может работать с генами с достаточно низким уровнем экспрессии. Данный метод открывает возможности для анализа небольших изменений, вызванных мутациями цис-регуляторных элементов, а также может быть применен для дискриминации цис- и транс-эффектов таких мутаций. В совокупности с гибридологическим методом, или с помощью мечения аллелей однонуклеотидными мутациями (например, с помощью технологий редактирования оснований [125]), такой подход уникальную возможность проследить влияние цис-регуляторных вариантов на экспрессию и функциональную активность генов на моноаллельном уровне, позволяя оценить цис-регуляторные эффекты при нивелировании любых транс-влияний. Таким образом, разработанный подход имеет большую ценность для исследований регуляторных механизмов генома в рамках прикладных и фундаментальных исследований.

Разработанный метод позволил количественно оценить экспрессию шести генов локуса *Unc5b/Slc29a3* (а именно, *Psap*, *Vsir*, *Cdh23*, *Slc29a3*, *Unc5b*

и *Sgpl1*) в пяти органах (почке, печени, мозжечке, обонятельной луковице и мочевом пузыре), и выявить 20 случаев значимых изменений экспрессии генов, вызванных делециями сайтов связывания CTCF (Рис. 16А,Б). Попробуем соотнести наблюдаемые различия в экспрессии генов с известными сценариями взаимодействий между промоторами и регуляторными элементами в контексте реорганизации пространственной архитектуры генома. Поскольку мы наблюдали зависимость изменений экспрессии генов от клеточного типа, логично предположить, что промотор-энхансерные взаимодействия в этом локусе зависят от местного эпигенетического контекста и, вероятно, определяются активностью транскрипции гена в конкретной ткани (Рис. 13В-Е).

Случай I. Экспрессия генов, разделенных границей ТАДов, становится более сходной после удаления сайтов связывания CTCF. Наблюдаемые нами изменения экспрессии близко расположенных генов *Slc29a3* и *Unc5b* дают возможность предполагать, что их цис-регуляторные ландшафты могут взаимодействовать и конкурировать друг с другом после нарушения границы ТАД. В почке мы обнаружили почти 40% снижение экспрессии *Slc29a3*. В той же ткани мы наблюдаем повышение экспрессии *Unc5b*. Увеличение экспрессии *Unc5b* варьируется от ~20% у животных с двумя удаленными сайтами связывания CTCF до ~40% у животных с четырьмя (Рис. 16А,Б). Примечательно, что в этом органе *Slc29a3* в норме обладает высоким уровнем транскрипционной активности, в то время как экспрессия *Unc5b* подавлена (Рис. 13Ж). Таким образом, мы предполагаем, что когда инсуляторная граница между этими генами нарушается, они подвергаются влиянию регуляторных элементов друг друга. В результате экспрессия этих генов «усредняется».

Случай II. Кража энхансера (англ. enhancer hijacking) вызывает повышение экспрессии низко экспрессирующихся генов. Мы обнаружили приблизительное увеличение экспрессии *Slc29a3* на 15% в мозжечке. В этом органе экспрессия гена *Unc5b* значительно выше по сравнению со средним

уровнем по всем тканям мыши, что предполагает воздействие на него активных цис-факторов (Рис. 13Ж). Хорошим кандидатом на эту роль является суперэнхансер первого интрона гена *Unc5b*, выделяемый по специфическим сигнатурам, согласно базе данных ENCODE (Рис. 8Г, 13Е). Ген *Slc29a3*, вступая в пространственное взаимодействие с геном *Unc5b*, также взаимодействует с его регуляторной средой, что объясняет увеличение активности *Slc29a3*. В отличие от случая совместных изменений этих генов в почке, уменьшение экспрессии активного гена *Unc5b* не обнаруживается, что можно приписать особенностям действия его энхансера или недетектируемым масштабам этого изменения.

Помимо двух описанных выше случаев, существует множество других примеров изменений экспрессии генов, которые плохо соответствуют нашим ожиданиям о механизмах дисрегуляции генной экспрессии при реорганизации структуры ТАДов.

В печени оба гена *Unc5b* и *Slc29a3* подавлены; однако, мы наблюдаем слабое увеличение экспрессии *Slc29a3* при нарушении границы. Это может быть объяснено взаимодействием гена *Slc29a3* с регуляторными элементами из пространственного окружения *Unc5b*. Однако нам не удалось обнаружить кандидатных элементов на основе доступных данных ChIP-Seq, и неясно, почему, если эти элементы существуют в среде *Unc5b*, не вызывают его активной экспрессии.

В мозжечке мы зафиксировали значительное снижение экспрессии гена *Cdh23*. Существует два известных механизма подавления генов, вызванных нарушением границы топологических доменов: разобщение промотора с его регуляторным элементом и распространение гетерохроматина [203]. По доступным данным ChIP-Seq/ATAC-seq мы не нашли очевидного кандидатного элемента с энхансерными сигнатурами, который изменял бы

частоту контактов с промотором *Cdh23* при нарушении функции границы ТАДов локуса.

В то же время, как показали эксперименты Hi-C, нарушение границы ТАДов привело к появлению эктопических дальних взаимодействий между телом гена *Cdh23* и дистальной границей *Unc5b*-содержащего ТАДа. Интересно, что эти петли, похоже, наблюдаются только для тканей, где экспрессируется *Cdh23* (для мозжечка и почки, но не для печени, Рис. 13Ж). Мы рассматривали гипотезу о том, что эти новые взаимодействия связаны с распространением репрессивных хроматиновых меток, ответственных за снижение экспрессии гена *Cdh23* в мозжечке. Однако в публичных данных мы не нашли никаких признаков присутствия репрессивных элементов, расположенных в основаниях возникших *de novo* архитектурных элементов.

В нашем эксперименте ген *Psap* продемонстрировал нетривиальную динамику. Несмотря на то что он находится дальше всех остальных генов от мутированной границы ТАДов, его экспрессия подавляется в почке животных с двумя удаленными сайтами связывания CTCF. Напротив, у животных с четырьмя удаленными сайтами связывания CTCF он экспрессируется на уровне нормы. При этом интересно, что паттерн изменений этого гена у животных с двумя удаленными сайтами связывания CTCF близко похож на паттерн гена *Slc29a3* между тканями.

Гены *Psap* и *Slc29a3* обычно сближены в основании петли, оформляющей ТАД *Slc29a3*. Таким образом, параллельные изменения, наблюдаемые в этих генах, можно объяснить их взаимовлиянием через пространственный контакт. Однако проведённый анализ Hi-C не выявил различий в архитектуре локуса при делеции двух и четырёх сайтов связывания CTCF, которые могли бы объяснить наблюдаемые различия в транскрипционном эффекте. Стандартный Hi-C обеспечивает разрешение порядка 5 тыс. п.н., чего недостаточно для детального анализа контактов промоторов и энхансеров на

околонуклеосомном уровне. Таким образом, нельзя исключить, что функционально значимые для транскрипции различия в архитектуре хроматина существуют, однако примененных методов недостаточно для их выявления. Перспективными подходами для такого анализа являются Micro-C [77] и Hi-TrAC [204].

Наконец, наблюдаемое увеличение экспрессии генов *Unc5b* и *Sgpl1* в мочевом пузыре и *Cdh23* в обонятельной луковице трудно объяснимы даже на теоретическом уровне. Следует учитывать, что изменения экспрессии в 10-20% являются нижним порогом детекции существующих сегодня методов. Поэтому, хотя все приведенные результаты и являются статистически значимыми, по статистическим причинам какая-то часть из них может являться артефактами. Кроме того, такие низкие по амплитуде изменения могут не иметь важного биологического и эволюционного влияния на функционирование организмов, поэтому слабо объясняют консервативность пространственной архитектуры хроматина локуса *Slc29a3/Unc5b*.

Таким образом, мы выявили широкий диапазон изменений экспрессии генов, несмотря на чрезвычайно низкую величину индуцированного транскрипционного эффекта. Такое незначительное воздействие на молекулярный фенотип, подтверждаемое и другими исследованиями (15,88,177), входит в противоречие с фактом высокой консервативности сайтов связывания CTCF и структуры ТАДов. Это несоответствие по-прежнему остается серьезным испытанием для всей теории о функциональной роли пространственной архитектуры в регуляции генной экспрессии. Возможно, такой умеренный эффект уже является значимой с точки зрения эволюции и позволяет стабилизирующему отбору сохранять положение границ топологических доменов и другие элементы пространственной архитектуры хроматина. Другое предположение состоит в том, что биологически значимые регуляторные нарушения обнаруживаются в очень узком клеточном типе и/или на конкретной стадии развития, и приводят к

фенотипам, которые сложно обнаружить, и поэтому такие важные молекулярные и фенотипические эффекты, находящиеся под давлением отбора, от нас ускользают.

Существует и другое предположение. Регуляторный контекст промотор-энхансерных взаимодействий задается не только расположением границ ТАДов, но еще и расстоянием между регуляторными элементами, и их взаимоспецифичностью [202,205]. Давление эволюционных сил должно приводить к тому, что эти механизмы оказываются взаимосогласованы в регуляции уровня транскрипции генов локуса. В таком случае, исчезновение границ ТАДов не ведет к тотальной реорганизации пространственных взаимодействий цис-элементов, так как компенсируется другими механизмами. Стоит заметить, что тотальная дезорганизация структуры пространственной архитектуры хроматина является не столь редким и чрезвычайным явлением, и имеет место в каждом цикле клеточного деления, а именно при формировании митотических хромосом [206], что не приводит к нарушению транскрипционных программ и гибели клеток. Кроме прочего, в подтверждение этой гипотезы выступает тот факт, что зачастую инверсии границ ТАДов, или их эктопическое появление, несут более выраженный эффект на регуляцию близлежащих генов, чем делеции оных [14,15,88].

Выявленный нами паттерн изменений транскрипционной активности генов локуса *Slc29a3/Unc5b* проявляет тканеспецифичный характер не только в плане амплитуды изменений в зависимости от конкретного гена и ткани, но и в направлении этих изменений. Хотя кажется очевидным, что изменения, вызванные нарушением границы ТАДов, обусловлены эпигенетическим состоянием локуса, их все еще трудно предсказать и объяснить. В настоящее время данный вопрос имеет высокую практическую ценность, поскольку играет важную роль в интерпретации мутаций в некодирующих областях генома, что актуально для задач медицинской генетики.

Наше видение основано на предположении, что активно транскрибируемые гены склонны терять уровень экспрессии, а репрессированные гены, наоборот, приобретать новые взаимодействия с энхансерами и повышать уровень транскрипции. Мы смогли наблюдать этот эффект в нашем эксперименте в случае взаимодействия генов *Slc29a3* и *Unc5b*, которые стремятся к усреднению уровней их экспрессии после потери инсуляторной границы между ними. Проявлением этого принципа, возможно, является снижение активности гена *Cdh23*, в норме высоко экспрессирующегося в мозжечке, при реорганизации структуры ТАДов. Можно ожидать, что нарушение функции границы ТАДов будет иметь тканеспецифичный эффект в зависимости от эпигенетического состояния локуса и истории дифференцировки клетки.

В нашем эксперименте последовательное уменьшение количества формирующих границу сайтов CTCF не привело к пропорциональному усилению эффекта как с точки зрения пространственной архитектуры локуса, так и регуляции транскрипции генов. Интересно, что для *Psap* и *Vsir* частичное нарушение границы оказало более значительный транскрипционный эффект, чем ее полное исчезновение. Это может быть свидетельством того, что делеции двух и четырех сайтов связывания CTCF, могут затрагивать различные функции границы ТАДов, а не просто масштабируют один и тот же эффект. Отсюда следует гипотеза о том, что границы ТАДов могут одновременно выполнять разные функции в регуляции экспрессии генов, и эти функции могут быть разделены (например, могут выполняться различными сайтами связывания CTCF). Это согласуется с предыдущими работами, показывающими, что отдельные сайты CTCF вносят неодинаковый вклад в инсуляторную функцию границы [207].

Здесь мы продемонстрировали, что воздействие нарушения границы ТАДов на регуляцию транскрипции генов имеет тканеспецифичный характер, показывая различные эффекты в разных тканях с труднопредсказуемыми

паттернами изменений. Это демонстрирует сложный баланс между архитектурными процессами ядра и эпигенетической регуляцией экспрессии генов в процессе дифференциации клеток. Кроме того, многие наблюдаемые нами эффекты сложно объяснимы с привлечением известных механизмов, что поднимает новые вопросы для дальнейших исследований.

Вторым локусом, исследованным в этой работе, стали окрестности гена *Epb4114a*, включающие клинически значимые гены *Nrep* и *Apc*, чьи регуляторные ландшафты разделены границей ТАДов. Данная граница сформирована крупным кластером сайтов связывания CTCF и обладает одним из самых высоких уровней инсуляции среди всех границ ТАДов человеческого генома. В публичных данных мы обнаружили хромосомные aberrации, приводящие к нарушению функции этой границы у пациентов с неврологическими заболеваниями, и воспроизвели подобные мутации на моделях животных. Мы показали, что внесенные нами делеции и инверсии приводят к снижению уровня экспрессии гена *Nrep* в некоторых частях мозга, что может свидетельствовать о возможной клинической значимости мутаций, нарушающих границу ТАДов в локусе *Epb4114a* у человека.

В настоящей работе не проводилось прямое картирование пространственной архитектуры мутантных аллелей в локусе *Epb4114a* методом Hi-C. Это ограничивает возможность количественно оценить степень изменения контактов хроматина после делеции или инверсии. Однако выбор данного локуса был основан на предварительно охарактеризованной структуре сильной консервативной границы ТАДов, и, в соответствии с данными литературы, протяженные делеции таких регионов почти гарантированно приводят к потере инсуляции между соседними регуляторными ландшафтами [16,88]. Таким образом, основной экспериментальный вопрос заключался в том, приводят ли перестройки, затрагивающие эту границу, к нарушениям экспрессии генов локуса *Epb4114a*.

Сценарий дисрегуляции *Nrep* во всех исследованных случаях трудно поддается интерпретации. Хотя как в случае делеции, так и в случае инверсии, *Nrep*, предположительно, вступает во взаимодействие с регуляторными элементами соседнего ТАДа, для разных типов мутаций происходит слияние ТАДа *Nrep* с разными фрагментами ТАДа *Apc*, разделенными точкой внесения двуцепочечного разрыва ДНК. Для делеции — это межгенная область между генами *Epb4114a* и *Apc*, для инверсии — промотор и тело гена *Epb4114a*. Маловероятно, что эктопические взаимодействия с различным набором регуляторных цис-элементов могли реализоваться в практически одинаковое (в обоих случаях трехкратное) снижение экспрессии *Nrep* в мозжечке. Более того, хотя мы только начинаем понимать механизмы функционирования репрессивных цис-элементов [208–210] и особенности этого процесса в нервной ткани [211], кажется маловероятной в нашем контексте возможность влияния репрессивных элементов на расстояние более 300 тыс. п.н., разделяющих промотор гена *Nrep* с границей внесенных мутаций, особенно учитывая, что соседний ТАД в этом органе не является глубоко репрессированным, и не обнаруживает гетерохроматиновых модификаций. Еще более странным кажется снижение экспрессии *Nrep* при инверсии в обонятельной луковице, где соседний ТАД, наоборот, демонстрирует более активный эпигенетический статус по сравнению с другими органами. Наша исходная гипотеза для этого случая заключалась в том, что при возникновении контакта между активным эпигенетическим ландшафтом гена *Epb4114a* и малоактивным промотором гена *Nrep* экспрессия последнего должна увеличиться. Это ожидание подкрепляется и подтвержденным случаем их взаимовлияния в мозжечке. Наблюдаемая же картина прямо противоположная, что говорит о более сложном механизме цис-взаимодействий этих генов, приводящем к диаметральному сценарию в разных органах.

Еще одним интересным предположением является гипотеза о возможности активирования энхансеров находящимися рядом активными промоторами. Такая картина могла бы реализовываться на фоне инверсии в обонятельной луковице, где активный ген *Epb4114a* попадает в пространственную доступность с молчащими энхансерами гена *Nrep*. В нашем случае эта гипотеза также не находит подтверждения в силу падения экспрессии гена *Nrep*, которое противоположно ожидаемой картине при активации его энхансеров.

Для объяснения снижения экспрессии гена *Nrep* следует рассмотреть другие факторы, помимо взаимодействия этого гена с новыми регуляторными элементами соседнего ТАДа. Мы полагаем, что нарушение контакта *Nrep* с самой областью границы доменов могло привести к дисрегуляции его экспрессии. Наиболее вероятным сценарием, на наш взгляд, является существование в границе ТАДов неких цис-регуляторных элементов, либо имеющих значительное влияние на промотор гена *Nrep*, либо выполняющих некую пионерскую функцию на одной из стадий развития нервной ткани. Также существует вероятность, что разделение энхансерных и топологических функций в данном случае невозможно, и тогда следует говорить об одном цис-факторе, одновременно формирующем границу ТАДов и модулирующем экспрессию генов локуса. Удаление этого элемента вместе с регионом границы, в таком случае, и является причиной дисрегуляции *Nrep*.

Альтернативная гипотеза заключается в том, что граница сама по себе не несет энхансерной функции, однако участвует в пространственной координации энхансеров, которая необходима для осуществления их функций. Такой сценарий показан для нескольких LCR, регионов контроля локуса, которые представляют собой, как правило, молекулярный хаб из нескольких энхансеров, способных взаимодействовать с несколькими промоторами локуса. Еще до начала осуществления ими энхансерной

функции, и до вступления в пространственный контакт с каким-либо промотором, эти элементы вступают в пространственное взаимодействие друг с другом, что является необходимым этапом их «созревания» и затем правильной регуляции экспрессии генов локуса. Кажется, что такая ситуация является довольно редкой, и встречается в особенных с точки зрения эволюции локусах с большим количеством генов-паралогов, однако нельзя отрицать возможность такого механизма в некоторых менее выраженных формах и для других случаев.

Таким образом, нарушение пространственной организации хроматина само по себе может приводить к нарушению регуляции генов затронутого локуса. Особенно этого следует ожидать для генов, выполняющих важные функции в развитии и жизнеподдержании организма, имеющих узкий пространственно-временной паттерн экспрессии и сложную систему регуляции транскрипции. Такими генами в нашем исследовании являются *Nrep* и *Cdh23*, демонстрирующими труднообъяснимое снижение экспрессии при нарушении топологии хроматина в органах, где наблюдается их повышенная экспрессия.

Резюмируя, в данной работе мы показали, что нарушение пространственной структуры хроматина в некодирующих регионах генома приводит к сложным и разнообразным изменениям экспрессии генов, которые зависят от тканеспецифического эпигенетического контекста. В то время как некоторые гены, такие как *Slc29a3* и *Unc5b*, продемонстрировали относительно предсказуемые изменения экспрессии, другие, как *Nrep* и *Cdh23*, оказались подвержены эффектам, которые трудно объяснить с точки зрения известных механизмов регуляции. Это подтверждает предположение о том, что роль границ ТАДов в регуляции транскрипции генов может зависеть от конкретного эпигенетического ландшафта локуса в определенной ткани или органе.

Полученные результаты открывают дорогу для дальнейшего изучения тканеспецифичности эффектов нарушений хроматиновой архитектуры, особенно в контексте генов с широким профилем активности. Наши данные также показывают, что функции границ ТАДов в регуляции генной экспрессии могут иметь более сложный характер, чем предполагалось ранее. Более того, в настоящий момент мы зачастую оказываемся неспособны предсказать или хотя бы интерпретировать транскрипционные последствия их нарушения. Таким образом, нам необходимо прикладывать усилия к развитию новых инструментов для исследования и предсказания механизмов пространственной регуляции генов и их значимости в патогенезе заболеваний.

## Выводы

1. На примере локуса *Slc29a3/Unc5b* мыши показано, что наиболее частым результатом репарации двуцепочечных разрывов в близкорасположенных сайтах связывания белка CTCF, одновременно таргетированных системой CRISPR/Cas9, является возникновение независимых коротких инсерций и делеций в каждом из сайтов. Такой подход позволяет получать генетически модифицированных животных с модифицированными кластерами сайтов CTCF, не затрагивая лежащие между ними функциональные геномные элементы.

2. Делеции сайтов связывания CTCF в локусе *Slc29a3/Unc5b* приводят к изменению локальной структуры трехмерных контактов хроматина, включая ослабление инсуляторной функции границы ТАДов и формирование эктопических петлевых взаимодействий в пределах локуса.

3. Нарушение CTCF-зависимой архитектуры локуса *Slc29a3/Unc5b* сопровождается разнонаправленными, тканеспецифическими изменениями экспрессии генов *Psap*, *Vsir*, *Cdh23*, *Slc29a3*, *Unc5b* и *Sgpl1*. Направление и величина изменений в экспрессии генов отличаются между тканями, что свидетельствует о тканеспецифичности влияния границы ТАДов на транскрипцию генов.

4. На модели генетически модифицированных мышей с хромосомными перестройками, затрагивающими регион консервативной границы ТАДов в локусе *Epb41l4a*, показано, что хромосомные перестройки, затрагивающие границу ТАДов в локусе *Epb41l4a*, приводят к нарушению регуляции транскрипции гена *Nrep* в головном мозге. Кроме того, на модели инверсии границы ТАДов показано, что взаимовлияние регуляторных ландшафтов *Nrep* и *Epb41l4a* асимметрично и зависит от тканевого контекста.

## Список литературы

1. de Wit E., de Laat W. A decade of 3C technologies: insights into nuclear organization: 1 // *Genes Dev.* 2012. Т. 26, № 1. С. 11–24.
2. Lieberman-Aiden E. и др. Comprehensive mapping of long range interactions reveals folding principles of the human genome: 5950 // *Science.* 2009. Т. 326, № 5950. С. 289–293.
3. Rao S. S. P. и др. A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping: 3 // *Cell.* Elsevier BV, 2015. Т. 162, № 3. С. 687–688.
4. Furlong E. E. M., Levine M. Developmental enhancers and chromosome topology // *Science.* New York, N.Y., 2018. Т. 361, № 6409. С. 1341–1345.
5. de Laat W., Duboule D. Topology of mammalian developmental enhancers and their regulatory landscapes // *Nature.* 2013. Т. 502, № 7472. С. 499–506.
6. Merckenschlager M., Nora E. P. CTCF and Cohesin in Genome Folding and Transcriptional Gene Regulation: 1 // *Annual Review of Genomics and Human Genetics.* Annual Reviews, 2016. Т. 17, № 1. С. 17–43.
7. Bell A. C., West A. G., Felsenfeld G. The Protein CTCF Is Required for the Enhancer Blocking Activity of Vertebrate Insulators: 3 // *Cell.* Elsevier BV, 1999. Т. 98, № 3. С. 387–396.
8. Hnisz D., Day D. S., Young R. A. Insulated neighborhoods: structural and functional units of mammalian gene control: 5 // *Cell.* 2016. Т. 167, № 5. С. 1188–1200.
9. Ibn-Salem J. и др. Deletions of chromosomal regulatory boundaries are associated with congenital disease: 9 // *Genome Biology.* 2014. Т. 15, № 9. С. 423.
10. Akdemir K. C. и др. Disruption of chromatin folding domains by somatic genomic rearrangements in human cancer // *Nat Genet.* Nature Publishing Group, 2020. Т. 52, № 3. С. 294–305.
11. Northcott P. A. и др. Enhancer hijacking activates GFI1 family oncogenes in medulloblastoma // *Nature.* Nature Publishing Group, 2014. Т. 511, № 7510. С. 428–434.
12. Hnisz D. и др. Activation of proto-oncogenes by disruption of chromosome neighborhoods // *Science.* 2016. Т. 351, № 6280. С. 1454–1458.
13. Della Chiara G. и др. Enhancers dysfunction in the 3D genome of cancer cells // *Front Cell Dev Biol.* 2023. Т. 11. С. 1303862.
14. Lupiáñez D. G. и др. Disruptions of Topological Chromatin Domains Cause Pathogenic Rewiring of Gene-Enhancer Interactions: 5 // *Cell.* Elsevier BV, 2015. Т. 161, № 5. С. 1012–1025.
15. Franke M. и др. Formation of new chromatin domains determines pathogenicity of genomic duplications: 7624 // *Nature.* Springer Science and Business Media LLC, 2016. Т. 538, № 7624. С. 265–269.

16. Kabirova E. и др. TAD border deletion at the Kit locus causes tissue-specific ectopic activation of a neighboring gene // *Nat Commun.* Nature Publishing Group, 2024. Т. 15, № 1. С. 4521.
17. Anania C. и др. In vivo dissection of a clustered-CTCF domain boundary reveals developmental principles of regulatory insulation // *Nat Genet.* Nature Publishing Group, 2022. Т. 54, № 7. С. 1026–1036.
18. Rhodes J. D. P. и др. Cohesin Disrupts Polycomb-Dependent Chromosome Interactions in Embryonic Stem Cells // *Cell Rep.* 2020. Т. 30, № 3. С. 820-835.e10.
19. Schwarzer W. и др. Two independent modes of chromatin organization revealed by cohesin removal // *Nature.* 2017. Т. 551, № 7678. С. 51–56.
20. Rao S. S. P. и др. Cohesin Loss Eliminates All Loop Domains // *Cell.* 2017. Т. 171, № 2. С. 305-320.e24.
21. Nora E. P. и др. Targeted degradation of CTCF decouples local insulation of chromosome domains from genomic compartmentalization // *Cell.* 2017. Т. 169, № 5. С. 930-944.e22.
22. Hyle J. и др. Acute depletion of CTCF directly affects MYC regulation through loss of enhancer–promoter looping // *Nucleic Acids Res.* 2019. Т. 47, № 13. С. 6699–6713.
23. Cova G. и др. Combinatorial effects on gene expression at the Lbx1/Fgf8 locus resolve split-hand/foot malformation type 3: 1 // *Nat Commun.* Nature Publishing Group, 2023. Т. 14, № 1. С. 1475.
24. Kragestein B. K. и др. Dynamic 3D chromatin architecture contributes to enhancer specificity and limb morphogenesis: 10 // *Nat Genet.* 2018. Т. 50, № 10. С. 1463–1473.
25. Long H. S. и др. Making sense of the linear genome, gene function and TADs: 1 // *Epigenetics & Chromatin.* 2022. Т. 15, № 1. С. 4.
26. Bulger M., Groudine M. Functional and Mechanistic Diversity of Distal Transcription Enhancers: 5 // *Cell.* Elsevier BV, 2011. Т. 144, № 5. С. 825.
27. Doni Jayavelu N. и др. Candidate silencer elements for the human and mouse genomes: 1 // *Nat Commun.* Nature Publishing Group, 2020. Т. 11, № 1. С. 1061.
28. Gaszner M., Felsenfeld G. Insulators: exploiting transcriptional and epigenetic mechanisms: 9 // *Nature Reviews Genetics.* Springer Science and Business Media LLC, 2006. Т. 7, № 9. С. 703–713.
29. Maeda R. K., Karch F. Gene expression in time and space: additive vs hierarchical organization of cis-regulatory regions: 2 // *Current Opinion in Genetics & Development.* Elsevier BV, 2011. Т. 21, № 2. С. 187–193.
30. Braccioli L., de Wit E. CTCF: a Swiss-army knife for genome organization and transcription regulation // *Essays Biochem.* 2019. Т. 63, № 1. С. 157–165.
31. Klenova E. M. и др. Functional Phosphorylation Sites in the C-Terminal Region of the Multivalent Multifunctional Transcriptional Factor CTCF: 6 // *Molecular and Cellular Biology.* Informa UK Limited, 2001. Т. 21, № 6. С. 2221–2234.

32. Haberle V. и др. Transcriptional cofactors display specificity for distinct types of core promoters // *Nature*. 2019. Т. 570, № 7759. С. 122–126.
33. Hampsey M. Molecular genetics of the RNA polymerase II general transcriptional machinery // *Microbiol Mol Biol Rev*. 1998. Т. 62, № 2. С. 465–503.
34. Deaton A. M., Bird A. CpG islands and the regulation of transcription: 10 // *Genes & Development*. Cold Spring Harbor Laboratory, 2011. Т. 25, № 10. С. 1010–1022.
35. Saxonov S., Berg P., Brutlag D. L. A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters // *Proc Natl Acad Sci U S A*. 2006. Т. 103, № 5. С. 1412–1417.
36. Haberle V., Stark A. Eukaryotic core promoters and the functional basis of transcription initiation // *Nat Rev Mol Cell Biol*. 2018. Т. 19, № 10. С. 621–637.
37. Long H. K. и др. Loss of Extreme Long-Range Enhancers in Human Neural Crest Drives a Craniofacial Disorder: 5 // *Cell Stem Cell*. 2020. Т. 27, № 5. С. 765-783.e14.
38. Arnold C. D. и др. Genome-Wide Quantitative Enhancer Activity Maps Identified by STARR-seq: 6123 // *Science*. American Association for the Advancement of Science, 2013. Т. 339, № 6123. С. 1074–1077.
39. Kvon E. Z. Using transgenic reporter assays to functionally characterize enhancers in animals // *Genomics*. 2015. Т. 106, № 3. С. 185–192.
40. Pennacchio L. A. и др. In vivo enhancer analysis of human conserved non-coding sequences // *Nature*. 2006. Т. 444, № 7118. С. 499–502.
41. Visel A. и др. ChIP-seq accurately predicts tissue-specific activity of enhancers // *Nature*. 2009. Т. 457, № 7231. С. 854–858.
42. Maeso I. и др. Deep conservation of cis-regulatory elements in metazoans: 1632 // *Philosophical Transactions of the Royal Society B: Biological Sciences*. The Royal Society, 2013. Т. 368, № 1632. С. 20130020.
43. Vavouri T., Lehner B. Conserved noncoding elements and the evolution of animal body plans // *Bioessays*. 2009. Т. 31, № 7. С. 727–735.
44. Ruf S. и др. Large-scale analysis of the regulatory architecture of the mouse genome with a transposon-associated sensor: 4 // *Nature Genetics*. Springer Science and Business Media LLC, 2011. Т. 43, № 4. С. 379–386.
45. Symmons O. и др. Functional and topological characteristics of mammalian regulatory domains: 3 // *Genome Research*. Cold Spring Harbor Laboratory, 2014. Т. 24, № 3. С. 390–400.
46. Hnisz D. и др. Super-Enhancers in the Control of Cell Identity and Disease: 4 // *Cell*. Elsevier BV, 2013. Т. 155, № 4. С. 934–947.
47. Will A. J. и др. Composition and dosage of a multipartite enhancer cluster control developmental expression of *Ihh* (Indian hedgehog): 10 // *Nature Genetics*. Springer Science and Business Media LLC, 2017. Т. 49, № 10. С. 1539–1545.

48. Lettice L. A. и др. A long-range Shh enhancer regulates expression in the developing limb and fin and is associated with preaxial polydactyly // *Hum Mol Genet.* 2003. Т. 12, № 14. С. 1725–1735.
49. Kagey M. H. и др. Mediator and cohesin connect gene expression and chromatin architecture // *Nature.* 2010. Т. 467, № 7314. С. 430–435.
50. Hnisz D. и др. A Phase Separation Model for Transcriptional Control // *Cell.* 2017. Т. 169, № 1. С. 13–23.
51. Cho W.-K. и др. Mediator and RNA polymerase II clusters associate in transcription-dependent condensates // *Science.* New York, N.Y., 2018. Т. 361, № 6400. С. 412–415.
52. Sabari B. R. и др. Coactivator condensation at super-enhancers links phase separation and gene control // *Science.* New York, N.Y., 2018. Т. 361, № 6400. С. eaar3958.
53. Boija A. и др. Transcription Factors Activate Genes through the Phase-Separation Capacity of Their Activation Domains // *Cell.* 2018. Т. 175, № 7. С. 1842-1855.e16.
54. Basu S. и др. Unblending of Transcriptional Condensates in Human Repeat Expansion Disease: 5 // *Cell.* Elsevier BV, 2020. Т. 181, № 5. С. 1062-1079.e30.
55. Chen H. и др. Dynamic interplay between enhancer–promoter topology and gene activity: 9 // *Nature Genetics.* Springer Science and Business Media LLC, 2018. Т. 50, № 9. С. 1296–1303.
56. Williamson I. и др. Shh and ZRS enhancer colocalisation is specific to the zone of polarising activity // *Development.* Cambridge, England, 2016. Т. 143, № 16. С. 2994–3001.
57. Cremer T., Cremer C. Chromosome territories, nuclear architecture and gene regulation in mammalian cells // *Nat Rev Genet.* 2001. Т. 2, № 4. С. 292–301.
58. Bolzer A. и др. Three-Dimensional Maps of All Chromosomes in Human Male Fibroblast Nuclei and Prometaphase Rosettes: 5 // *PLoS Biology* / под ред. Misteli T. Public Library of Science (PLoS), 2005. Т. 3, № 5. С. e157.
59. Branco M. R., Pombo A. Intermingling of Chromosome Territories in Interphase Suggests Role in Translocations and Transcription-Dependent Associations: 5 // *PLoS Biology* / под ред. Becker P. Public Library of Science (PLoS), 2006. Т. 4, № 5. С. e138.
60. Shopland L. S. и др. Folding and organization of a contiguous chromosome region according to the gene distribution pattern in primary genomic sequence: 1 // *The Journal of Cell Biology.* Rockefeller University Press, 2006. Т. 174, № 1. С. 27–38.
61. Chambeyron S., Bickmore W. A. Chromatin decondensation and nuclear reorganization of the HoxB locus upon induction of transcription: 10 // *Genes & Development.* Cold Spring Harbor Laboratory, 2004. Т. 18, № 10. С. 1119–1130.

62. Ferrai C. и др. Poised Transcription Factories Prime Silent uPA Gene Prior to Activation: 1 // PLoS Biology / под ред. Misteli T. Public Library of Science (PLoS), 2010. Т. 8, № 1. С. e1000270.
63. Kosak S. T. и др. Subnuclear compartmentalization of immunoglobulin loci during lymphocyte development // Science. New York, N.Y., 2002. Т. 296, № 5565. С. 158–162.
64. Osborne C. S. и др. Active genes dynamically colocalize to shared sites of ongoing transcription // Nat Genet. 2004. Т. 36, № 10. С. 1065–1071.
65. Tsukamoto T. и др. Visualization of gene activity in living cells // Nat Cell Biol. 2000. Т. 2, № 12. С. 871–878.
66. Peric-Hupkes D. и др. Molecular Maps of the Reorganization of Genome-Nuclear Lamina Interactions during Differentiation: 4 // Molecular Cell. Elsevier BV, 2010. Т. 38, № 4. С. 603–613.
67. Finlan L. E. и др. Recruitment to the Nuclear Periphery Can Alter Expression of Genes in Human Cells: 3 // PLoS Genetics / под ред. Reik W. Public Library of Science (PLoS), 2008. Т. 4, № 3. С. e1000039.
68. Dekker J. и др. Capturing chromosome conformation: 5558 // Science. 2002. Т. 295, № 5558. С. 1306–1311.
69. Palstra R.-J. и др. The beta-globin nuclear compartment in development and erythroid differentiation: 2 // Nat Genet. 2003. Т. 35, № 2. С. 190–194.
70. Murrell A., Heeson S., Reik W. Interaction between differentially methylated regions partitions the imprinted genes Igf2 and H19 into parent-specific chromatin loops: 8 // Nat Genet. 2004. Т. 36, № 8. С. 889–893.
71. Spilianakis C. G. и др. Interchromosomal associations between alternatively expressed loci: 7042 // Nature. 2005. Т. 435, № 7042. С. 637–645.
72. Vernimmen D. и др. Long-range chromosomal interactions regulate the timing of the transition between poised and active gene expression: 8 // EMBO J. 2007. Т. 26, № 8. С. 2041–2051.
73. Yaffe E., Tanay A. Probabilistic modeling of Hi-C contact maps eliminates systematic biases to characterize global chromosomal architecture: 11 // Nat Genet. 2011. Т. 43, № 11. С. 1059–1065.
74. Fishman V. S., Salnikov P. A., Battulin N. R. Interpreting Chromosomal Rearrangements in the Context of 3-Dimensional Genome Organization: A Practical Guide for Medical Genetics: 4 // Biochemistry (Mosc). 2018. Т. 83, № 4. С. 393–401.
75. Schoenfelder S. и др. The pluripotent regulatory circuitry connecting promoters to their long-range interacting elements: 4 // Genome Res. 2015. Т. 25, № 4. С. 582–597.
76. Downes D. J. и др. Capture-C: a modular and flexible approach for high-resolution chromosome conformation capture: 2 // Nat Protoc. Nature Publishing Group, 2022. Т. 17, № 2. С. 445–475.
77. Hamley J. C. и др. Determining chromatin architecture with Micro Capture-C // Nat Protoc. Nature Publishing Group, 2023. Т. 18, № 6. С. 1687–1711.

78. Hughes J. R. и др. Analysis of hundreds of cis-regulatory landscapes at high resolution in a single, high-throughput experiment: 2 // *Nat Genet.* Nature Publishing Group, 2014. Т. 46, № 2. С. 205–212.
79. Imakaev M. и др. Iterative correction of Hi-C data reveals hallmarks of chromosome organization // *Nat Methods.* 2012. Т. 9, № 10. С. 999–1003.
80. Dixon J. R. и др. Chromatin architecture reorganization during stem cell differentiation // *Nature.* 2015. Т. 518, № 7539. С. 331–336.
81. Dekker J., Heard E. Structural and functional diversity of Topologically Associating Domains // *FEBS Lett.* 2015. Т. 589, № 20 Pt A. С. 2877–2884.
82. Dixon J. R. и др. Topological Domains in Mammalian Genomes Identified by Analysis of Chromatin Interactions: 7398 // *Nature.* 2012. Т. 485, № 7398. С. 376–380.
83. Pope B. D. и др. Topologically associating domains are stable units of replication-timing regulation: 7527 // *Nature.* Nature Publishing Group, 2014. Т. 515, № 7527. С. 402–405.
84. Ulianov S. V. и др. Active chromatin and transcription play a key role in chromosome partitioning into topologically associating domains // *Genome Res.* 2016. Т. 26, № 1. С. 70–84.
85. Tang Z. и др. CTCF-Mediated Human 3D Genome Architecture Reveals Chromatin Topology for Transcription: 7 // *Cell.* Elsevier BV, 2015. Т. 163, № 7. С. 1611–1627.
86. Holwerda S., Laat W. de. Chromatin loops, gene positioning, and gene expression // *Frontiers in Genetics.* Frontiers Media SA, 2012. Т. 3.
87. Wutz G. и др. Topologically associating domains and chromatin loops depend on cohesin and are regulated by CTCF, WAPL, and PDS5 proteins: 24 // *The EMBO Journal.* Springer Science and Business Media LLC, 2017. Т. 36, № 24. С. 3573–3599.
88. Despang A. и др. Functional dissection of the Sox9–Kcnj2 locus identifies nonessential and instructive roles of TAD architecture: 8 // *Nature Genetics.* Springer Science and Business Media LLC, 2019. Т. 51, № 8. С. 1263–1271.
89. Sanborn A. L. и др. Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes: 47 // *Proc Natl Acad Sci U S A.* 2015. Т. 112, № 47. С. E6456–6465.
90. Cobbe N., Heck M. M. S. The Evolution of SMC Proteins: Phylogenetic Analysis and Structural Implications: 2 // *Molecular Biology and Evolution.* 2004. Т. 21, № 2. С. 332–347.
91. Cheng H., Zhang N., Pati D. Cohesin subunit RAD21: From biology to disease // *Gene.* 2020. Т. 758. С. 144966.
92. Hirano T. Condensins: universal organizers of chromosomes with diverse functions: 15 // *Genes Dev.* Cold Spring Harbor Lab, 2012. Т. 26, № 15. С. 1659–1678.
93. Serrano D. и др. The Smc5/6 Core Complex Is a Structure-Specific DNA Binding and Compacting Machine: 6 // *Molecular Cell.* Elsevier, 2020. Т. 80, № 6. С. 1025–1038.e5.

94. Hirano M., Hirano T. Hinge-mediated dimerization of SMC protein is essential for its dynamic interaction with DNA: 21 // *The EMBO Journal*. John Wiley & Sons, Ltd, 2002. Т. 21, № 21. С. 5733–5744.
95. Wells J. N. и др. Evolution of condensin and cohesin complexes driven by replacement of Kite by Hawk proteins: 1 // *Current Biology*. Elsevier, 2017. Т. 27, № 1. С. R17–R18.
96. Kikuchi S. и др. Crystal structure of the cohesin loader Scc2 and insight into cohesinopathy: 44 // *Proceedings of the National Academy of Sciences*. Proceedings of the National Academy of Sciences, 2016. Т. 113, № 44. С. 12444–12449.
97. Shi Z. и др. Cryo-EM structure of the human cohesin-NIPBL-DNA complex: 6498 // *Science*. American Association for the Advancement of Science, 2020. Т. 368, № 6498. С. 1454–1459.
98. Pradhan B. и др. SMC complexes can traverse physical roadblocks bigger than their ring size: 3 // *Cell Reports*. Elsevier, 2022. Т. 41, № 3.
99. Kim T. H. и др. Analysis of the Vertebrate Insulator Protein CTCF-Binding Sites in the Human Genome: 6 // *Cell*. Elsevier, 2007. Т. 128, № 6. С. 1231–1245.
100. Li Y. и др. The structural basis for cohesin–CTCF-anchored loops: 7795 // *Nature*. Nature Publishing Group, 2020. Т. 578, № 7795. С. 472–476.
101. Nora E. P. и др. Molecular basis of CTCF binding polarity in genome folding: 1 // *Nat Commun*. Nature Publishing Group, 2020. Т. 11, № 1. С. 5612.
102. Banigan E. J. и др. Chromosome organization by one-sided and two-sided loop extrusion // *eLife* / под ред. Struhl K. eLife Sciences Publications, Ltd, 2020. Т. 9. С. e53558.
103. Ganji M. и др. Real-time imaging of DNA loop extrusion by condensin: 6384 // *Science*. American Association for the Advancement of Science, 2018. Т. 360, № 6384. С. 102–105.
104. Kim J. и др. Condensin DC loads and spreads from recruitment sites to create loop-anchored TADs in *C. elegans* // *eLife* / под ред. Tyler J. K. eLife Sciences Publications, Ltd, 2022. Т. 11. С. e68745.
105. Kong M. и др. Human Condensin I and II Drive Extensive ATP-Dependent Compaction of Nucleosome-Bound DNA: 1 // *Molecular Cell*. Elsevier, 2020. Т. 79, № 1. С. 99–114.e9.
106. Nuebler J. и др. Chromatin organization by an interplay of loop extrusion and compartmental segregation // *Proc Natl Acad Sci U S A*. 2018. Т. 115, № 29. С. E6697–E6706.
107. Banigan E. J. и др. Transcription shapes 3D chromatin organization by interacting with loop extrusion // *Proc Natl Acad Sci U S A*. 2023. Т. 120, № 11. С. e2210480120.
108. Dequeker B. J. H. и др. MCM complexes are barriers that restrict cohesin-mediated loop extrusion: 7912 // *Nature*. Nature Publishing Group, 2022. Т. 606, № 7912. С. 197–203.

109. Gaj T., Gersbach C. A., Barbas C. F. ZFN, TALEN and CRISPR/Cas-based methods for genome engineering // *Trends Biotechnol.* 2013. Т. 31, № 7. С. 397–405.
110. Mianné J. и др. Analysing the outcome of CRISPR-aided genome editing in embryos: Screening, genotyping and quality control // *Methods.* San Diego, Calif., 2017. Т. 121–122. С. 68–76.
111. Korablev A. и др. On-Target CRISPR/Cas9 Activity Can Cause Undesigned Large Deletion in Mouse Zygotes: 10 // *Int J Mol Sci.* 2020. Т. 21, № 10. С. 3604.
112. Brandl C. и др. Creation of targeted genomic deletions using TALEN or CRISPR/Cas nuclease pairs in one-cell mouse embryos // *FEBS Open Bio.* 2015. Т. 5. С. 26–35.
113. Mali P. и др. RNA-guided human genome engineering via Cas9 // *Science.* New York, N.Y., 2013. Т. 339, № 6121. С. 823–826.
114. Lesmana H. и др. Casgevy (exagamglogene autotemcel) and Lyfgenia (lovotibeglogene autotemcel) for individuals 12 years and older with sickle cell disease (SCD) and recurrent vaso-occlusive crises (VOC): A therapeutics bulletin of the American College of Medical Genetics and Genomics (ACMG) // *Genet Med Open.* 2024. Т. 2. С. 101875.
115. Singh A. и др. Revolutionary breakthrough: FDA approves CASGEVY, the first CRISPR/Cas9 gene therapy for sickle cell disease // *Ann Med Surg (Lond).* 2012, 2024. Т. 86, № 8. С. 4555–4559.
116. Frangoul H. и др. Exagamglogene Autotemcel for Severe Sickle Cell Disease // *N Engl J Med.* 2024. Т. 390, № 18. С. 1649–1662.
117. «Any idiot can do it.» Genome editor CRISPR could put mutant mice in everyone's reach [Электронный ресурс]. URL: <https://www.science.org/content/article/any-idiot-can-do-it-genome-editor-crispr-could-put-mutant-mice-everyones-reach> (дата обращения: 27.05.2026).
118. Makarova K. S. и др. An updated evolutionary classification of CRISPR–Cas systems including rare variants // *Nat Microbiol.* Nature Publishing Group, 2025. Т. 10, № 12. С. 3346–3361.
119. Smirnov A. V. и др. CRISPR/Cas9, a universal tool for genomic engineering: 4 // *Russian Journal of Genetics: Applied Research.* Pleiades Publishing Ltd, 2017. Т. 7, № 4. С. 440–458.
120. Haeussler M. и др. Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR: 1 // *Genome Biology.* Springer Science and Business Media LLC, 2016. Т. 17, № 1.
121. Xu X., Duan D., Chen S.-J. CRISPR-Cas9 cleavage efficiency correlates strongly with target-sgRNA folding stability: from physical mechanism to off-target assessment // *Sci Rep.* 2017. Т. 7, № 1. С. 143.
122. Slaymaker I. M. и др. Rationally engineered Cas9 nucleases with improved specificity // *Science.* New York, N.Y., 2016. Т. 351, № 6268. С. 84–88.
123. Boroviak K. и др. Revealing hidden complexities of genomic rearrangements generated with Cas9 // *Sci Rep.* 2017. Т. 7, № 1. С. 12867.

124. Aussel C., Cathomen T., Fuster-García C. The hidden risks of CRISPR/Cas: structural variations and genome integrity // *Nat Commun. Nature Publishing Group*, 2025. Т. 16, № 1. С. 7208.
125. Yan A. P. и др. Base editing in the AUTS2 gene and high-throughput NGS genotyping of clones: a strategy for generating a cellular model // *Vavilovskii Zhurnal Genet Seleksii*. 2026. Т. 30, № 1. С. 72–84.
126. Chang H. H. Y. и др. Non-homologous DNA end joining and alternative pathways to double-strand break repair // *Nat Rev Mol Cell Biol*. 2017. Т. 18, № 8. С. 495–506.
127. McVey M., Lee S. E. MMEJ repair of double-strand breaks (director's cut): deleted sequences and alternative endings // *Trends Genet*. 2008. Т. 24, № 11. С. 529–538.
128. Liao H. и др. CRISPR-Cas9-mediated homology-directed repair for precise gene editing // *Mol Ther Nucleic Acids*. 2024. Т. 35, № 4. С. 102344.
129. Xue C., Greene E. C. DNA Repair Pathway Choices in CRISPR-Cas9-Mediated Genome Editing // *Trends Genet*. 2021. Т. 37, № 7. С. 639–656.
130. Yu Y. и др. Cytosine base editors with minimized unguided DNA and RNA off-target events and high on-target activity // *Nat Commun*. 2020. Т. 11, № 1. С. 2052.
131. Lu Z., Huang X. Base editors: a powerful tool for generating animal models of human diseases // *Cell Stress*. 2018. Т. 2, № 10. С. 242–245.
132. Geurts M. H. и др. One-step generation of tumor models by base editor multiplexing in adult stem cell-derived organoids // *Nat Commun*. 2023. Т. 14, № 1. С. 4998.
133. Billon P. и др. CRISPR-Mediated Base Editing Enables Efficient Disruption of Eukaryotic Genes through Induction of STOP Codons // *Mol Cell*. 2017. Т. 67, № 6. С. 1068-1079.e4.
134. Liang Y. и др. Base editors: development and applications in biomedicine // *Front Med*. 2023. Т. 17, № 3. С. 359–387.
135. Singh P., Schimenti J. C., Bolcun-Filas E. A mouse geneticist's practical guide to CRISPR applications // *Genetics*. 2015. Т. 199, № 1. С. 1–15.
136. Cheng A. W. и др. Multiplexed activation of endogenous genes by CRISPR-on, an RNA-guided transcriptional activator system // *Cell Res*. 2013. Т. 23, № 10. С. 1163–1171.
137. Rees H. A., Liu D. R. Base editing: precision chemistry on the genome and transcriptome of living cells // *Nat Rev Genet*. 2018. Т. 19, № 12. С. 770–788.
138. Gaudelli N. M. и др. Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage // *Nature*. 2017. Т. 551, № 7681. С. 464–471.
139. Grünewald J. и др. Transcriptome-wide off-target RNA editing induced by CRISPR-guided DNA base editors // *Nature*. 2019. Т. 569, № 7756. С. 433–437.
140. Anzalone A. V. и др. Search-and-replace genome editing without double-strand breaks or donor DNA // *Nature*. 2019. Т. 576, № 7785. С. 149–157.

141. Choi J. и др. Precise genomic deletions using paired prime editing // *Nat Biotechnol.* 2022. Т. 40, № 2. С. 218–226.
142. Jiang T. и др. Deletion and replacement of long genomic sequences using prime editing // *Nat Biotechnol.* 2022. Т. 40, № 2. С. 227–234.
143. Gori J. L. и др. Prime Editing for p47phox-Deficient Chronic Granulomatous Disease // *N Engl J Med.* 2026. Т. 394, № 12. С. 1195–1203.
144. He Q., Ge W. The tandem Agenet domain of fragile X mental retardation protein interacts with FUS: 1 // *Sci Rep.* 2017. Т. 7, № 1. С. 962.
145. Valton A.-L., Dekker J. TAD disruption as oncogenic driver // *Curr Opin Genet Dev.* 2016. Т. 36. С. 34–40.
146. Barutcu A. R. и др. A TAD boundary is preserved upon deletion of the CTCF-rich Firre locus: 1 // *Nature Communications.* Springer Science and Business Media LLC, 2018. Т. 9, № 1.
147. Tadagavadi R. K., Wang W., Ramesh G. Netrin-1 regulates Th1/Th2/Th17 cytokine production and inflammation through UNC5B receptor and protects kidney against ischemia-reperfusion injury: 6 // *J Immunol.* 2010. Т. 185, № 6. С. 3750–3758.
148. Kaur S. и др. Expression Pattern for unc5b, an Axon Guidance Gene in Embryonic Zebrafish Development: 6 // *Gene Expr.* 2018. Т. 13, № 6. С. 321–327.
149. Pradella D. и др. A ligand-insensitive UNC5B splicing isoform regulates angiogenesis by promoting apoptosis // *Nat Commun.* 2021. Т. 12. С. 4872.
150. Hu X. и др. UNC-5 netrin receptor B mediates osteogenic differentiation by modulating bone morphogenetic protein signaling in human adipose-derived stem cells: 1 // *Biochem Biophys Res Commun.* 2018. Т. 495, № 1. С. 1167–1174.
151. Kong C. и др. Overexpression of UNC5B in bladder cancer cells inhibits proliferation and reduces the volume of transplantation tumors in nude mice: 1 // *BMC Cancer.* 2016. Т. 16, № 1. С. 892.
152. Lee H. K. и др. Netrin-1 induces proliferation of Schwann cells through Unc5b receptor: 4 // *Biochem Biophys Res Commun.* 2007. Т. 362, № 4. С. 1057–1062.
153. Wang W., Reeves W. B., Ramesh G. Netrin-1 increases proliferation and migration of renal proximal tubular epithelial cells via the UNC5B receptor: 4 // *Am J Physiol Renal Physiol.* 2009. Т. 296, № 4. С. F723-729.
154. Ahn E. H. и др. Netrin1 deficiency activates MST1 via UNC5B receptor, promoting dopaminergic apoptosis in Parkinson's disease: 39 // *Proc Natl Acad Sci U S A.* 2020. Т. 117, № 39. С. 24503–24513.
155. Huang L. и др. Netrin-1 induces the anti-apoptotic and pro-survival effects of B-ALL cells through the Unc5b-MAPK axis: 1 // *Cell Commun Signal.* 2022. Т. 20, № 1. С. 122.
156. Prieto C. P. и др. Downregulation of the Netrin-1 Receptor UNC5b Underlies Increased Placental Angiogenesis in Human Gestational Diabetes Mellitus: 6 // *Int J Mol Sci.* 2019. Т. 20, № 6. С. 1408.

157. Baldwin S. A. и др. Functional characterization of novel human and mouse equilibrative nucleoside transporters (hENT3 and mENT3) located in intracellular membranes: 16 // *J Biol Chem*. 2005. Т. 280, № 16. С. 15880–15887.
158. Elwi A. N. и др. Renal nucleoside transporters: physiological and clinical implications: 6 // *Biochem Cell Biol*. 2006. Т. 84, № 6. С. 844–858.
159. Wei C.-W. и др. Equilibrative Nucleoside Transporter 3 Regulates T Cell Homeostasis by Coordinating Lysosomal Function with Nucleoside Availability: 8 // *Cell reports*. *Cell Rep*, 2018. Т. 23, № 8.
160. Hsu C.-L. и др. Equilibrative Nucleoside Transporter 3 Deficiency Perturbs Lysosome Function and Macrophage Homeostasis // *Science*. American Association for the Advancement of Science, 2012.
161. Nair S. и др. Adult stem cell deficits drive Slc29a3 disorders in mice: 1 // *Nat Commun*. Nature Publishing Group, 2019. Т. 10, № 1. С. 2943.
162. Wu S. и др. High expression of UNC5B enhances tumor proliferation, increases metastasis, and worsens prognosis in breast cancer: 17 // *Aging (Albany NY)*. 2020. Т. 12, № 17. С. 17079–17098.
163. Zeng Z. и др. Oleanolic Acid (OA) Targeting UNC5B Inhibits Proliferation and EMT of Ovarian Cancer Cell and Increases Chemotherapy Sensitivity of Niraparib // *J Oncol*. 2022. Т. 2022. С. 5887671.
164. Kang N. и др. Human Equilibrative Nucleoside Transporter-3 (hENT3) Spectrum Disorder Mutations Impair Nucleoside Transport, Protein Localization, and Stability: 36 // *J Biol Chem*. 2010. Т. 285, № 36. С. 28343–28352.
165. Melki I. и др. Mutation in the SLC29A3 Gene: A New Cause of a Monogenic, Autoinflammatory Condition: 4 // *Pediatrics*. 2013. Т. 131, № 4. С. e1308–e1313.
166. Cagdas D. и др. Autoinflammation in addition to combined immunodeficiency: SLC29A3 gene defect // *Molecular immunology*. 2020. Т. 121.
167. Bolz H. и др. Mutation of CDH23, encoding a new member of the cadherin gene family, causes Usher syndrome type 1D: 1 // *Nat Genet*. 2001. Т. 27, № 1. С. 108–112.
168. Bork J. M. и др. Usher syndrome 1D and nonsyndromic autosomal recessive deafness DFNB12 are caused by allelic mutations of the novel cadherin-like gene CDH23: 1 // *Am J Hum Genet*. 2001. Т. 68, № 1. С. 26–37.
169. Hulková H. и др. A novel mutation in the coding region of the prosaposin gene leads to a complete deficiency of prosaposin and saposins, and is associated with a complex sphingolipidosis dominated by lactosylceramide accumulation: 9 // *Human Molecular Genetics*. 2001. Т. 10, № 9. С. 927–940.
170. O'Brien J. S. и др. Coding of two sphingolipid activator proteins (SAP-1 and SAP-2) by same genetic locus: 4869 // *Science*. 1988. Т. 241, № 4869. С. 1098–1101.
171. Oji Y. и др. Variants in saposin D domain of prosaposin gene linked to Parkinson's disease: 4 // *Brain*. 2020. Т. 143, № 4. С. 1190–1205.

172. Ishiguro H. и др. Isolation and characterization of human NBL4, a gene involved in the beta-catenin/tcf signaling pathway: 6 // *Jpn J Cancer Res.* 2000. Т. 91, № 6. С. 597–603.
173. Du M. и др. LncRNA EPB41L4A-AS2 represses Nasopharyngeal Carcinoma Metastasis by binding to YBX1 in the Nucleus and Sponging MiR-107 in the Cytoplasm: 8 // *Int J Biol Sci.* 2021. Т. 17, № 8. С. 1963–1978.
174. Liao M. и др. LncRNA EPB41L4A-AS1 regulates glycolysis and glutaminolysis by mediating nucleolar translocation of HDAC2 // *EBioMedicine.* 2019. Т. 41. С. 200–213.
175. Yang F., Lv S. LncRNA EPB41L4A-AS1 Regulates Cell Proliferation, Apoptosis and Metastasis in Breast Cancer: 1 // *Ann Clin Lab Sci.* 2022. Т. 52, № 1. С. 3–11.
176. Wang S. и др. Reconstruction and Functional Annotation of P311 Protein–Protein Interaction Network Reveals Its New Functions // *Front Genet.* 2019. Т. 10. С. 109.
177. Paliwal S. и др. P311 binds to the latency associated protein and downregulates the expression of TGF-beta1 and TGF-beta2: 4 // *Biochemical and biophysical research communications.* Biochem Biophys Res Commun, 2004. Т. 315, № 4.
178. Taylor G. A. и др. Regulation of P311 Expression by Met-Hepatocyte Growth Factor/Scatter Factor and the Ubiquitin/Proteasome System\*: 6 // *Journal of Biological Chemistry.* 2000. Т. 275, № 6. С. 4215–4219.
179. Wang S. и др. P311 Deficiency Leads to Attenuated Angiogenesis in Cutaneous Wound Healing // *Front Physiol.* 2017. Т. 8. С. 1004.
180. Chen C. и др. P311 Promotes IL-4 Receptor–Mediated M2 Polarization of Macrophages to Enhance Angiogenesis for Efficient Skin Wound Healing: 4 // *J Invest Dermatol.* 2023. Т. 143, № 4. С. 648-660.e6.
181. Tan J. и др. Investigating the role of P311 in the hypertrophic scar: 4 // *PLoS One.* 2010. Т. 5, № 4. С. e9995.
182. Yao Z. и др. P311 Accelerates Skin Wound Reepithelialization by Promoting Epidermal Stem Cell Migration Through RhoA and Rac1 Activation: 6 // *Stem Cells Dev.* 2017. Т. 26, № 6. С. 451–460.
183. Alkhateeb A. и др. Transcriptomics Signature from Next-Generation Sequencing Data Reveals New Transcriptomic Biomarkers Related to Prostate Cancer // *Cancer Inform.* 2019. Т. 18. С. 1176935119835522.
184. Li Q. и др. NREP is a Diagnostic and Prognostic Biomarker, and Promotes Gastric Cancer Cell Proliferation and Angiogenesis: 2 // *Biochem Genet.* 2023. Т. 61, № 2. С. 669–686.
185. Taylor G. A. и др. Behavioral characterization of P311 knockout mice: 7 // *Genes Brain Behav.* 2008. Т. 7, № 7. С. 786–795.
186. Sun Y.-G. и др. Involvement of P311 in the affective, but not in the sensory component of pain // *Mol Pain.* 2008. Т. 4. С. 23.
187. Bitetti A. и др. MicroRNA degradation by a conserved target RNA regulates animal behavior: 3 // *Nat Struct Mol Biol.* Nature Publishing Group, 2018. Т. 25, № 3. С. 244–251.

188. Choi J. и др. Adenomatous polyposis coli is down-regulated by the ubiquitin-proteasome pathway in a process facilitated by Axin: 47 // *J Biol Chem*. 2004. Т. 279, № 47. С. 49188–49198.
189. Homma M. K. и др. Association and regulation of casein kinase 2 activity by adenomatous polyposis coli protein: 9 // *Proceedings of the National Academy of Sciences*. *Proceedings of the National Academy of Sciences*, 2002. Т. 99, № 9. С. 5959–5964.
190. Horii A. и др. Multiple forms of the APC gene transcripts and their tissue-specific expression: 3 // *Human Molecular Genetics*. 1993. Т. 2, № 3. С. 283–287.
191. Rowan A. J. и др. APC mutations in sporadic colorectal tumors: A mutational “hotspot” and interdependence of the “two hits”: 7 // *Proceedings of the National Academy of Sciences*. *Proceedings of the National Academy of Sciences*, 2000. Т. 97, № 7. С. 3352–3357.
192. Buchert M. и др. Genetic dissection of differential signaling threshold requirements for the Wnt/beta-catenin pathway in vivo: 1 // *PLoS Genet*. 2010. Т. 6, № 1. С. e1000816.
193. Moser A. R. и др. Homozygosity for the Min allele of Apc results in disruption of mouse development prior to gastrulation: 4 // *Dev Dyn*. 1995. Т. 203, № 4. С. 422–433.
194. Kentepozidou E. и др. Clustered CTCF binding is an evolutionary mechanism to maintain topologically associating domains // *Genome Biol*. 2020. Т. 21, № 1. С. 5.
195. Kaiser V. B., Semple C. A. When TADs go bad: chromatin structure and nuclear organisation in human disease // *F1000Res*. 2017. Т. 6. С. F1000 Faculty Rev-314.
196. Bonev B. и др. Multiscale 3D Genome Rewiring during Mouse Neural Development: 3 // *Cell*. 2017. Т. 171, № 3. С. 557–572.e24.
197. Shen M. W. и др. Predictable and precise template-free CRISPR editing of pathogenic variants: 7733 // *Nature*. 2018. Т. 563, № 7733. С. 646–651.
198. Bao L., Zhou M., Cui Y. CTCFBSDB: a CTCF-binding site database for characterization of vertebrate genomic insulators: Database issue // *Nucleic Acids Res*. 2008. Т. 36, № Database issue. С. D83–D87.
199. Ziebarth J. D., Bhattacharya A., Cui Y. CTCFBSDB 2.0: a database for CTCF-binding sites and genome organization: Database issue // *Nucleic Acids Res*. 2013. Т. 41, № Database issue. С. D188–194.
200. Ringel A. R. и др. Repression and 3D-restructuring resolves regulatory conflicts in evolutionarily rearranged genomes: 20 // *Cell*. 2022. Т. 185, № 20. С. 3689–3704.e21.
201. Pachano T. и др. Orphan CpG islands amplify poised enhancer regulatory activity and determine target gene responsiveness: 7 // *Nat Genet*. Nature Publishing Group, 2021. Т. 53, № 7. С. 1036–1049.

202. Pachano T., Haro E., Rada-Iglesias A. Enhancer-gene specificity in development and disease: 11 // *Development*. 2022. T. 149, № 11. С. dev186536.
203. Kabirova E. и др. Function and Evolution of the Loop Extrusion Machinery in Animals: 5 // *Int J Mol Sci*. 2023. T. 24, № 5. С. 5017.
204. Liu S. и др. Hi-TrAC reveals division of labor of transcription factors in organizing chromatin loops // *Nat Commun*. Nature Publishing Group, 2022. T. 13, № 1. С. 6679.
205. Whalen S., Truty R. M., Pollard K. S. Enhancer–promoter interactions are encoded by complex genomic signatures on looping chromatin: 5 // *Nature Genetics*. Springer Science and Business Media LLC, 2016. T. 48, № 5. С. 488–496.
206. Naumova N. и др. Organization of the Mitotic Chromosome // *Science*. American Association for the Advancement of Science, 2013.
207. Amândio A. R. и др. Sequential in cis mutagenesis in vivo reveals various functions for CTCF sites at the mouse HoxD cluster // *Genes Dev*. Cold Spring Harbor Lab, 2021. T. 35, № 21–22. С. 1490–1509.
208. Kraft K. и др. Polycomb-mediated genome architecture enables long-range spreading of H3K27 methylation // *Proc Natl Acad Sci U S A*. 2022. T. 119, № 22. С. e2201883119.
209. Pang B. и др. Identification of non-coding silencer elements and their regulation of gene expression: 6 // *Nat Rev Mol Cell Biol*. Nature Publishing Group, 2023. T. 24, № 6. С. 383–395.
210. Pang B., Snyder M. P. Systematic identification of silencers in human cells: 3 // *Nat Genet*. Nature Publishing Group, 2020. T. 52, № 3. С. 254–263.
211. Pletenev I. A. и др. Extensive long-range polycomb interactions and weak compartmentalization are hallmarks of human neuronal 3D genome: 11 // *Nucleic Acids Res*. 2024. T. 52, № 11. С. 6234–6252.

## Приложения

Таблица 1. Последовательности олигонуклеотидов, использовавшихся в работе.

Название	Последовательность (от 5' к 3')	Описание
q-mUNC5B-F2	5'-CCACAGGCTTGCGGGT-3'	Праймеры для определения уровня экспрессии соответствующих генов методом количественной ПЦР
q-mUNC5B-R2	5'-CTCATGATCCAAGGGTACCTCC-3'	
Unc5b_mus_F	5'-CGAGTGGGTCAAGCCAGAATG-3'	
Unc5b_mus_R	5'-CCAGAGGCTCCTGGTCAAAGT-3'	
MusSlcF	5'-CCACCTTCACCAGCACAATC-3'	
MusSlcR	5'-GGACCCTGAGGACATGGAC-3'	
q-mEEF2-F	5'-CACTTACCATCCCCCGTCAC-3'	
q-mEEF2-R	5'-AAGGCATAGAAGCGGCCTTT-3'	
q-mGAPDG-UF	5'-GGGTTCTATAAAATACGGACTGC-3'	
q-mGapdhF	5'-AGGTCGGTGTGAACGGATTTG-3'	
q-mGapdhR	5'-TGTAGACCATGTAGTTGAGGTCA-3'	
q-mAKTB-F	5'-tattggcaacgagcgggtcc-3'	
q-mAKTB-R	5'-tgcatagaggtctttacgg-3'	
q-mGAPDH-PR	5'-FAM-AAAGTGGAGATTGTTGCCATCAACGACCC-BHQ-3'	Флуоресцентные зонды для определения экспрессии соответствующих генов методом количественной ПЦР
q-mSLC29a3-PR	5'-FAM-AACAGGGTCCAGGTGCACGTCCGT-BHQ-3'	
q-mUNC5b-PR	5'-FAM-AAGAGTCGCCGAGCCTACATCCGCATT-BHQ-3'	
q-mEEF2-PR	5'-FAM-ACTTATGATGTACATTTCCAAGATGGTGCC-BHQ-3'	
q-mAKTB-PR	5'-FAM-ccagccttcctcttgggtatggaat-BHQ-3'	
LLF	5'-AGAACTGGCAAGGACCTGTT-3'	Праймеры для генотипирования мышей по аллелю dM1R1
LLR	5'-GCTGGCCAGACATGTTCTCA-3'	
MLF	5'-GTTCTGAGGCACAACAGGGA-3'	
MLF2	5'-GGAATTCAGGTTGAGTGGGC-3'	
MLF2U	5'-GCAACACACGACACTATCCC-3'	
MLR	5'-CTTGTGGAGTTCATTTGCATTGTC-3'	
MLR2	5'-GATCCCAGGTATAATGAACGC-3'	
MRF	5'-ATGTGACTGTACCGCTGGTC-3'	
MRR	5'-GAATTGGAGCCTCAGCCGTT-3'	
RRF	5'-TCATCGCCGGAAGTGTTAC-3'	
RRL	5'-ATGGGCCAGGCTTGACTAAG-3'	
RRRD	5'-CATCTCTCCAGCCCGGTAAG-3'	
T7+ML-gRNA_F	5'-GTTAATACGACTCACTATAgtaactaagacaTCCA-3'	
	5'-GTAGTTTTAGAGCTAGAAATAGCAAGTTAA-3'	
T7+RR-gRNA_F	5'-GTTAATACGACTCACTATAGTAGAGAGAAC-3'	
	5'-AGGCGAACCTGTTTTAGAGCTAGAAATAGCAAGTTAA-3'	
T7+sgRNA_M2-Unc5b-CTCF	5'-GTTAATACGACTCACTATAgTACCGAGAGCTCCCACCAGAGTTTTAGAGCTAGAAATAGCAAGTTAA-3'	
T7+sgRNA_R2-Unc5b-CTCF	5'-GTTAATACGACTCACTATAgTGCCCCCTAGTGGCCACGGTTTTAGAGCTAGAAATAGCAAGTTAA-3'	

Название	Последовательность (от 5' к 3')	Описание
T7+sgRNA_L2- <i>Unc5b</i> -CTCF	5'- gttaatacgaactcactatagTTTACAACCTGCAGCCCCTGCgttttagagctagaa atagaaggttaa-3'	
<i>Unc5b</i> -R2-R	5'-CTCCCACTGTATGCCAGACGA-3'	Праймеры для генотипирования мышей по мутациям в сайтах L2, M2 и R2
<i>Unc5b</i> -M2-F	5'-CTGGGTACCTGGGATTTGGC-3'	
<i>Unc5b</i> -M2-R	5'-TCCTGCTCACCATTGACCAC-3'	
<i>Unc5b</i> -L2-F	5'-ATTGGATGGGAGGTGCAGTG-3'	
<i>Unc5b</i> -L2-R	5'-CGGTGCTTCTTAGCACTCCT-3'	
q-m <i>UNC5B</i> -R2	5'-CTCATGATCCAAGGGTACCTCC-3'	
ssoR2_ <i>Unc5b</i>	5'- GCAAGTCTGAGAAGAGTTCCAAGAACATCAAGGGATTGA GCATGTGAGCATGAAATGaagcttCATGCCTATGGAACGTGTT ACAGGAGCCCTCACCTGCGGAAACAGGGTCCCCCTCC-3'	Одноцепочечный дезоксирiboолиго- нуклеотидный донор гомологии
ssoM2_ <i>Unc5b</i>	5'- TCATCGTCTATGGTGTGGATGGGTGTTTGTGGCATGGGCT GTGATTATCCCGGGTgaagcttCTCTCGGTACTGTAGGGGGGC CTGAAGCAATATATCCCTCCACTCTGCCCCACCCC-3'	
ssoL2_ <i>Unc5b</i>	5'- TGGAGAGGTGGTGTGAGAGGGATGGGAAAGAAAGATGGAAC CCAGGGAGCCCACAGCGAAagctTGCCAGGCAGGGgtccacaga tcattctagaaggacagaaatctctgtccctcat-3'	
MLRR_ssODN- template	5'-TATTTGGGTCCTCACTTTCCAAGACTG-3'	
	5'-TAGGTCTGCCTGCTCTCCCTTTTCTGCCCTTA-3'	
	5'-CTTCGCCTGTTCTCTCTACCCCACTCACC-3'	
	5'-CAACGATAACCTCTCTCTCAAGGCAAGGACCC-3'	
gRNA_REV	5'-AAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGAT-3'	Универсальные дезоксирiboолиго- нуклеотиды для синтеза хнРНК матрицы
	5'- AACGGACTAGCCTTATTTAACTTGCTATTTCTAGCTCTA- 3'	
REV	5'-AAAAGCACCGACTCGGTGCCACTTTTTCAAG-3'	
R2-ATGC-F	5'-ACATGCACCTCATGATCCAAGGGTACCTCC-3'	Праймеры, использованные для получения NGS-библиотек
R2-ATGC-R	5'-ACATGCAGCTGGAGGATTCCAGGCGA-3'	
M2-ATGC-F	5'-ACATGCAGAGTGTGGCTGGGTACCTG-3'	
M2-ATGC-R	5'-ACATGCAACTACACCTGTGTGGCCAAG-3'	
L2-ATGC-F	5'-ACATGCAGCACGTGGGGTGAGAG-3'	
L2-ATGC-R	5'-ACATGCCACATTAGGTCTACTTGAGCA>R2-CATG-F-3'	
R2-CATG-R	5'-ACCATGAGCTGGAGGATTCCAGGCGA-3'	
M2-CATG-F	5'-ACCATGAGAGTGTGGCTGGGTACCTG-3'	
M2-CATG-R	5'-ACCATGAACTACACCTGTGTGGCCAAG-3'	
L2-CATG-F	5'-ACCATGAGCACGTGGGGTGAGAG-3'	
L2-CATG-R	5'-ACCATGCCACATTAGGTCTACTTGAGCA-3'	
R2-GCAT-F	5'-ACGCATACCTCATGATCCAAGGGTACCTCC-3'	
R2-GCAT-R	5'-ACGCATAGCTGGAGGATTCCAGGCGA-3'	
M2-GCAT-F	5'-ACGCATAGAGTGTGGCTGGGTACCTG-3'	
M2-GCAT-R	5'-ACGCATAACTACACCTGTGTGGCCAAG-3'	
L2-GCAT-F	5'-ACGCATAGCACGTGGGGTGAGAG-3'	
L2-GCAT-R	5'-ACGCATCCACATTAGGTCTACTTGAGCA-3'	
R2-TGCA-F	5'-ACTGCAACCTCATGATCCAAGGGTACCTCC-3'	
R2-TGCA-R	5'-ACTGCAAGCTGGAGGATTCCAGGCGA-3'	

Название	Последовательность (от 5' к 3')	Описание	
M2-TGCA-F	5'-ACTGCAAGAGTGTGGCTGGGTACCTG-3'		
M2-TGCA-R	5'-ACTGCAAACTACACCTGTGTGGCCAAG-3'		
L2-TGCA-F	5'-ACTGCAAGCACGTGGGGTGAGAG-3'		
L2-TGCA-R	5'-ACTGCACCACATTAGGTCTACTTGAGCA-3'		
R2-CGTA-F	5'-ACCGTAACCTCATGATCCAAGGGTACCTCC-3'		
M2-CGTA-F	5'-ACCGTAAGAGTGTGGCTGGGTACCTG-3'		
L2-CGTA-F	5'-ACCGTAAGCACGTGGGGTGAGAG-3'		
Unc_ad_F3	5'-GGCTTCCACCCTGTCAACTTC-3'		Олигонуклеотиды, использовавшиеся для цифровой ПЦР
Unc_ad_R3	5'-GGCCAGACCCAGAACCGAT-3'		
Unc_ad_P3snp	5'-HEX-CTGGAGTTATAGACCTTtATCTTGAGGCT-BHQ1-3'		
Unc_ad_P3wt	5'-FAM-CTGGAGTTATAGACCTTGATCTTGAGGCT-BHQ1-3'		
Slc_ad_F1	5'-AATAGCAGCGTTTACGGCCT-3'		
Slc_ad_R2	5'-GACTGCTGCCATGAGGAAGAA-3'		
Slc_ad_P2wt	5'-FAM-TGCTGTCCCGCACATCACTGGA-BHQ1-3'		
Slc_ad_F2snp	5'-HEX-TGCTGTCCCGCACGTCACTGGA-BHQ1-3'		
Psap_ad_F1	5'-GAAAGATAATGCTACGCAGGAGGA-3'		
Psap_ad_R1	5'-TTGAGCGCAGAGCACACTTC-3'		
Psap_ad_P1wt	5'-FAM-TCGTGCAAGGAGGTGGTTGACTCTT-BHQ1-3'		
Psap_ad_P1snp	5'-HEX-TCGTGTAAGGAGGCGGTTGACTCTT-BHQ1-3'		
Cdh23_ad_F1	5'-CCAAGGTGGACTCTACTGTGAAG-3'		
Cdh23_ad_R1	5'-TCTCAAAGGCCCCAGCAATG-3'		
Cdh23_ad_P1wt	5'-TAMRA-CGACTTCACCTCGGACTCAGCCATC-BHQ2-3'		
Cdh23_ad_P1snp	5'-ROX-CGACTTTACCTCGGACTCGGCCATC-BHQ2-3'		
Vsir_ad_F1	5'-CCCTGCTCCTTGCTATTTTCCT-3'		
Vsir_ad_R1	5'-ATCTGGACCTCGCCTCGT-3'		
Vsir_ad_P1wt	5'-TAMRA-TCTTGTAGATGGTCACATCGTGCCCTT-BHQ2-3'		
Vsir_ad_P1snp	5'-ROX-TCTTGTAGACGGTCACATCGTGCCCTT-BHQ2-3'		
Sgpl1_ad_F	5'-TGTGAAAACCTCGCTGCTCA-3'		
Sgpl1_ad_R	5'-ATATCTGGATGCAGTGGATTGCT-3'		
Sgpl1_ad_Pwt	5'-TAMRA-AGCTGTGTACAATGGGGAACCGAAGCTCA-BHQ2-3'		
Sgpl1_ad_Psnp	5'-ROX-AGCTGTGTACAACGGGGAACCGAAGCTCA-BHQ2-3'		
ActB_F	5'-TATTGGCAACGAGCGGTTCC-3'	Олигонуклеотиды, использовавшиеся для приготовления библиотек для UMI-опосредованного таргетного РНК-секвенирования	
ActB_R	5'-TGGCATAGAGGTCTTTACGG-3'		
ActB_P	5'-Cy5-CCAGCCTTCCTTCTTGGGTATGGAAT-BHQ2-3'		
microilluminaR	5'-CTGGAGTTCAGACGTGT-3'		
SQ-SLC-scstr1	5'-TCTTTCCCTACACGACGCTCTTCCGATCTTATCAGGAGGAGCCATGGGAG-3'		
SQ-SLC-rev1	5'-CTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNANNNNCNNNNCATAACAGAGCCCAAGGAAGACT-3'		
SQ-UNC-2-scstr1	5'-TCTTTCCCTACACGACGCTCTTCCGATCTGAACAAGCCAGTGGAACTGC-3'		
SQ-UNC-2-rev1	5'-CTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNANNNNCNNNNACACCTCGATCTGCACCTCT-3'		

Название	Последовательность (от 5' к 3')	Описание
SQ- <i>Vsir-2</i> -scstr1	5'- TCTTTCCCTACACGACGCTCTCCGATCTCTGGGGACGTCT TTTTCCA-3'	
SQ- <i>Vsir-2</i> -rev1	5'- CTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNANNNNC NNNNATAGGGGAGGGTCTGTGCAT-3'	
SQ- <i>Cdh23-1-dd</i> - scstr1	5'- TCTTTCCCTACACGACGCTCTTCCGATCTACATTGACAGTC GTGGCAGA-3'	
SQ- <i>Cdh23-1-dd</i> - rev1	5'- CTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNANNNNC NNNNGCCTTGACTGTGGCTACTCG-3'	
SQ- <i>Sgpl</i> -scstr1	5'- TCTTTCCCTACACGACGCTCTTCCGATCTACAACAGGTGA GCAAAGCCA-3'	
SQ- <i>Sgpl</i> -rev1	5'- CTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNANNNNC NNNNTTGCTCCACGTGAATTCTCCA-3'	
SQ- <i>Psap</i> -scstr1	5'- TCTTTCCCTACACGACGCTCTTCCGATCTTCACCGAAGCTG GGAAGTTG-3'	
SQ- <i>Psap</i> -rev1	5'- CTGGAGTTCAGACGTGTGCTCTTCCGATCTNNNNANNNNC NNNNTCCAGGATGACAGGCAGGTA-3'	
Epb_DEL_F	5'-CTCTTCTGCATGAAAGCATAGATGT-3'	
Epb_DEL_R	5'-CAAACAGACCAAAGCACCCTC	
Epb_INV_L_F	5'-GACTTACTGCACAAGGAAGGAC	
Epb_INV_L_R	5'-TGTTTGCTGTCTGACTTGGGA	
Epb_INV_R_F	5'-AGGACACATGGGCATGACTAAT	
Epb_DEL_R	5'-CAAACAGACCAAAGCACCCTC	Праймеры для генотипирования мышей по мутациям в локусе <i>Epb4114a</i>

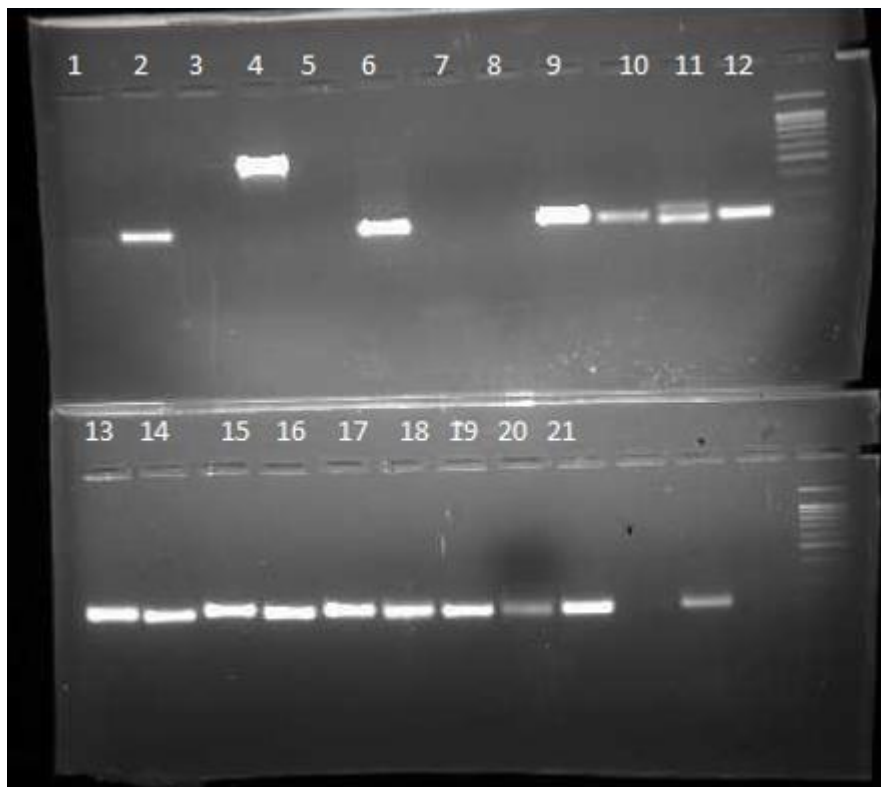


Рисунок 1. ПЦР-генотипирование мышей F0 по аллелю dM1R1 в эксперименте по удалению центромерных сайтов связывания CTCF в локусе *Slc29a3/Unc5b* и получению аллеля dM1R1 (праймеры MLF2 и RRL).



Рисунок 2. ПЦР-генотипирование мышей F0 по аллелю инверсии (верхний ряд, праймеры MLF2 и RRF) и дупликации (нижний ряд, праймеры MLR2 и RRF) в эксперименте по удалению центромерных сайтов связывания CTCF в локусе *Slc29a3/Unc5b* и получению аллеля dM1R1.

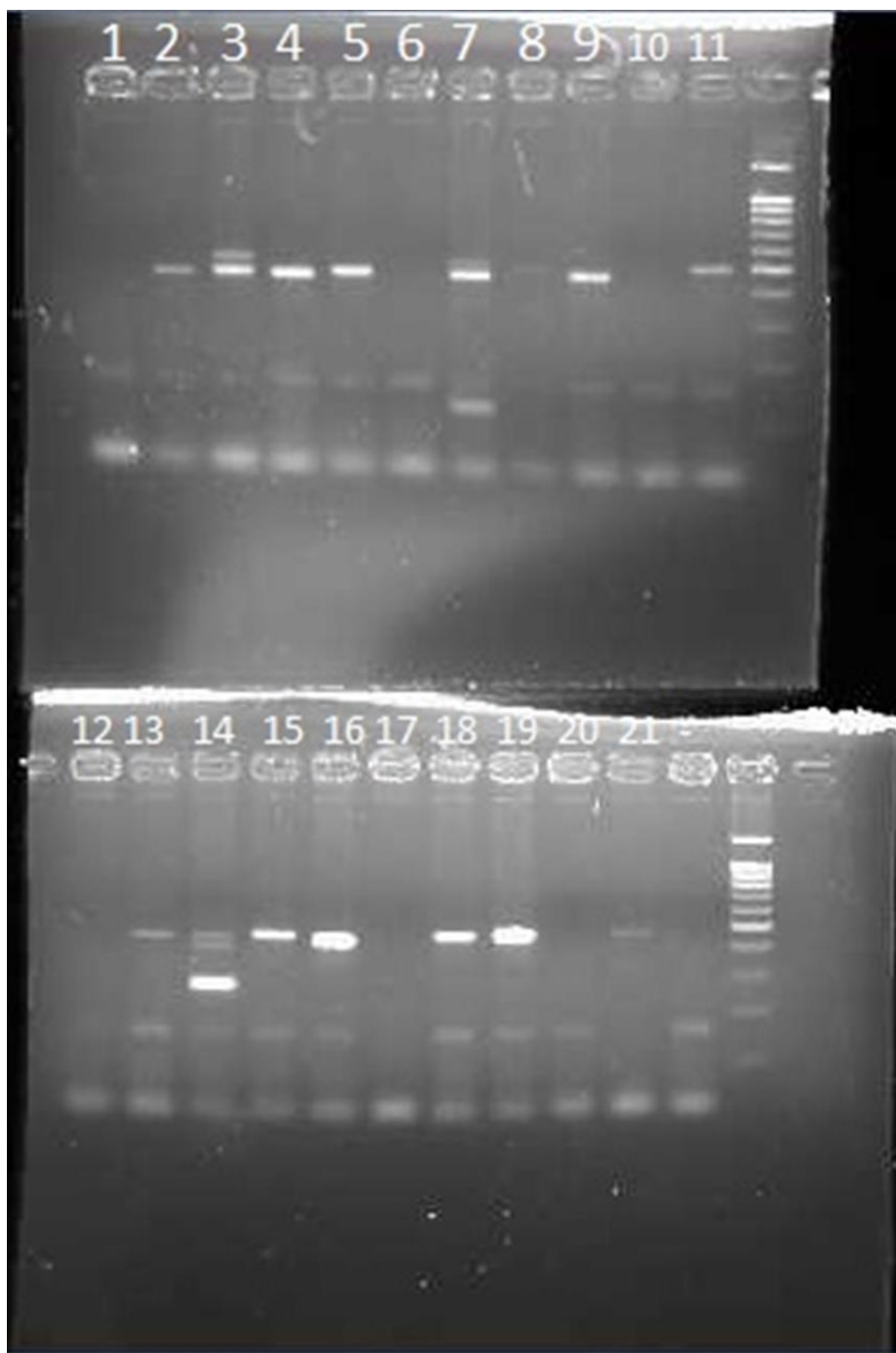


Рисунок 3. ПЦР-генотипирование мышей F0 по левому сайту внесения DSB (chr10:60755585) в эксперименте по удалению центромерных сайтов связывания CTCF в локусе *Slc29a3/Unc5b* и получению аллеля dM1R1 (праймеры MLF2 и MLR2).

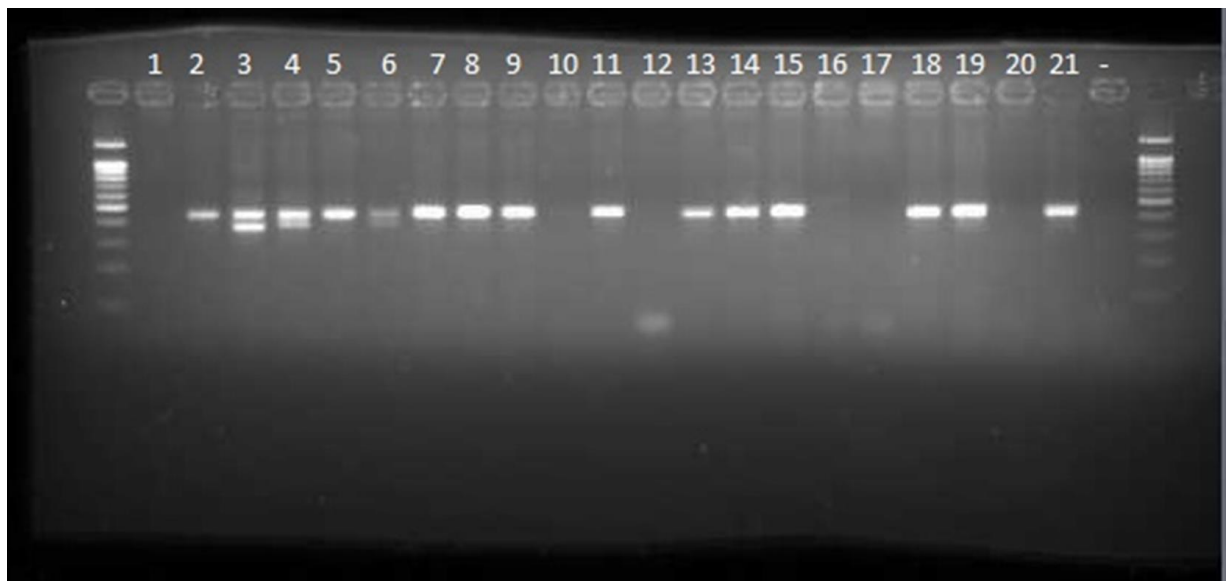


Рисунок 4. ПЦР-генотипирование мышей F0 по правому сайту внесения DSB (chr10:60761088) в эксперименте по удалению центромерных сайтов связывания CTCF в локусе *Slc29a3/Unc5b* и получению аллеля dM1R1 (праймеры RRF и RRRD).

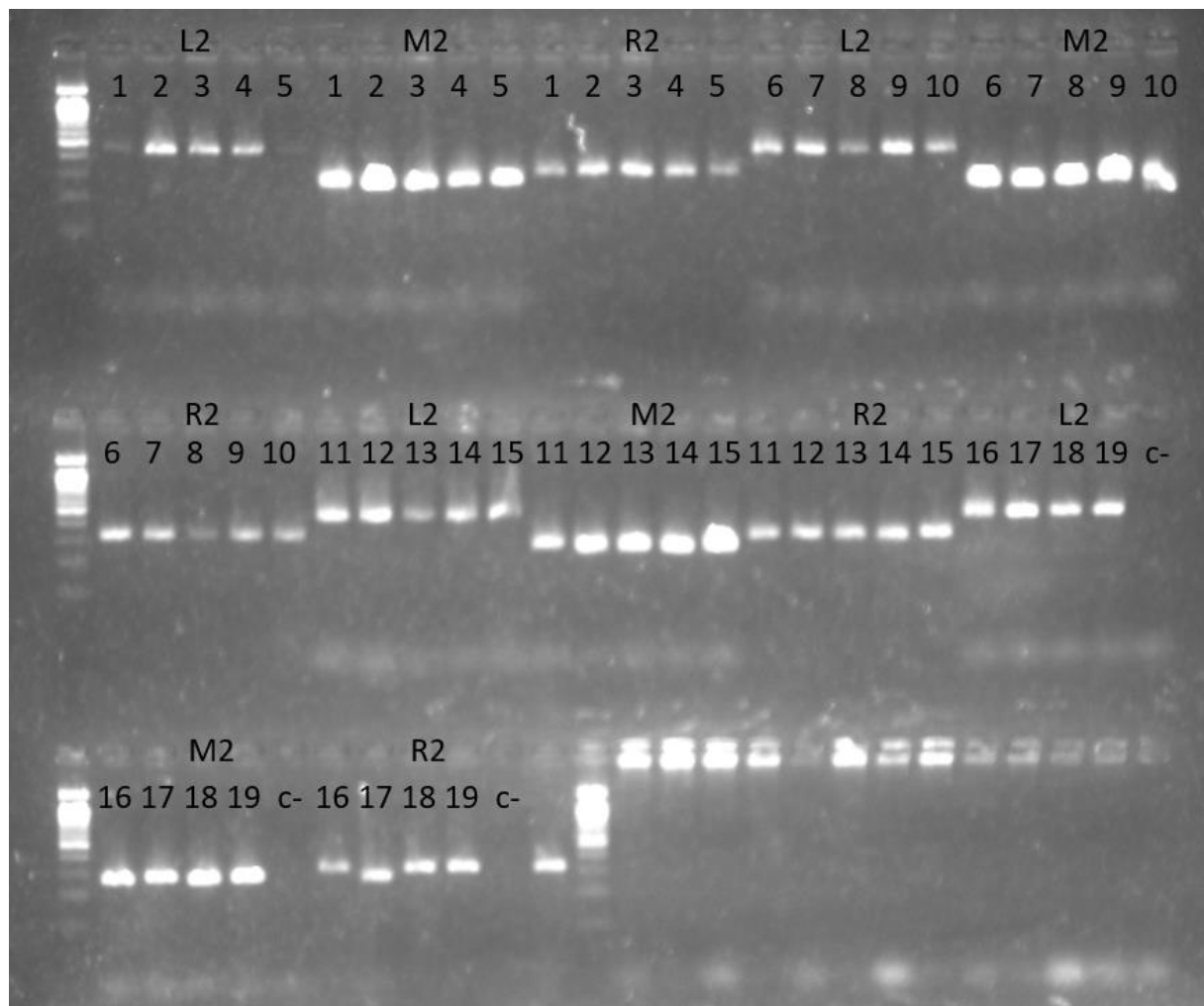


Рисунок 5. ПЦР-ПДРФ генотипирование мышей F0 по сайтам внесения DSB (L2 — chr10:60775689, M2 — chr10:60778732, R2 — chr10:60779829) в эксперименте по удалению теломерных сайтов связывания CTCF в локусе *Slc29a3/Unc5b* и получению мутаций в сайтах L2, M2 и R2 (праймеры Unc5b-R2-F, Unc5b-R2-R, Unc5b-M2-F, Unc5b-M2-R, Unc5b-L2-F, Unc5b-L2-R). ПЦР продукты перед обработкой ферментами рестрикции.

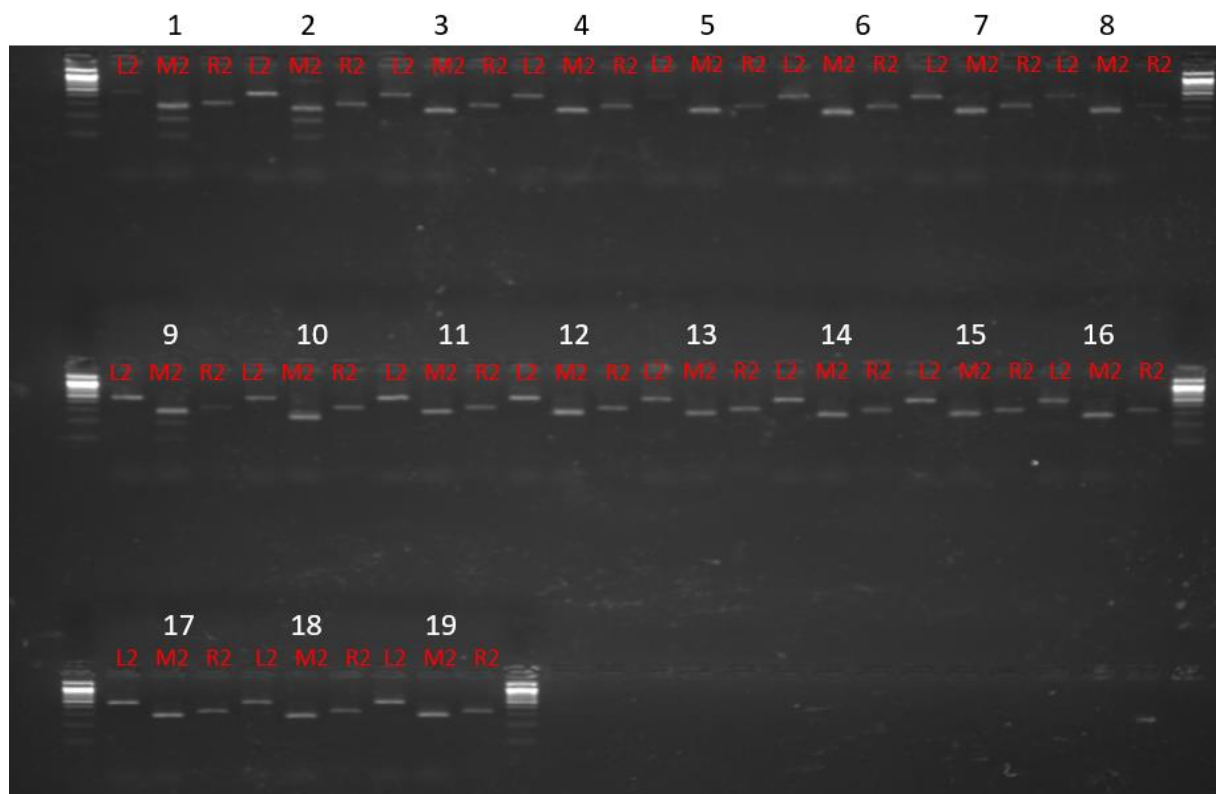


Рисунок 6. ПЦР-ПДРФ генотипирование мышей F0 по сайтам внесения DSB (L2 — chr10:60775689, M2 — chr10:60778732, R2 — chr10:60779829) в эксперименте по удалению теломерных сайтов связывания CTCF в локусе *Slc29a3/Unc5b* и получению мутаций в сайтах L2, M2 и R2 (праймеры Unc5b-R2-F, Unc5b-R2-R, Unc5b-M2-F, Unc5b-M2-R, Unc5b-L2-F, Unc5b-L2-R). ПЦР продукты после обработки ферментом рестрикции *HindIII*.

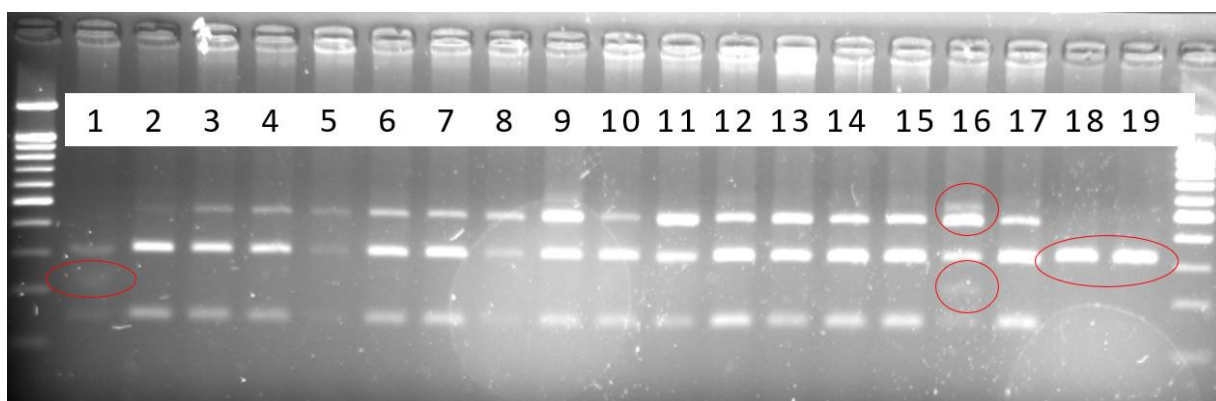


Рисунок 7. ПЦР-ПДРФ генотипирование мышей F0 по сайту внесения DSB L2 в эксперименте по удалению теломерных сайтов связывания CTCF в локусе *Slc29a3/Unc5b* и получению мутаций в сайтах L2, M2 и R2 (праймеры *Unc5b-L2-F* и *Unc5b-L2-R*). ПЦР продукты для сайта L2 после обработки ферментом рестрикции *PstI*.

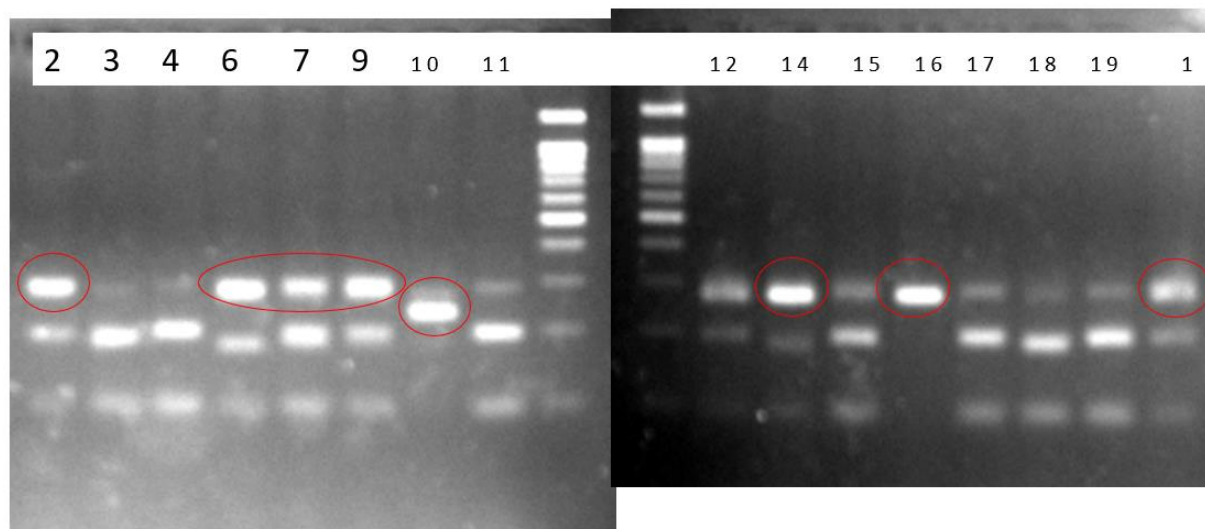


Рисунок 8. ПЦР-ПДРФ генотипирование мышей F0 по сайту внесения DSB M2 в эксперименте по удалению теломерных сайтов связывания CTCF в локусе *Slc29a3/Unc5b* и получению мутаций в сайтах L2, M2 и R2 (праймеры *Unc5b-M2-F* и *Unc5b-M2-R*). ПЦР продукты для сайта M2 после обработки ферментом рестрикции *SacI*.

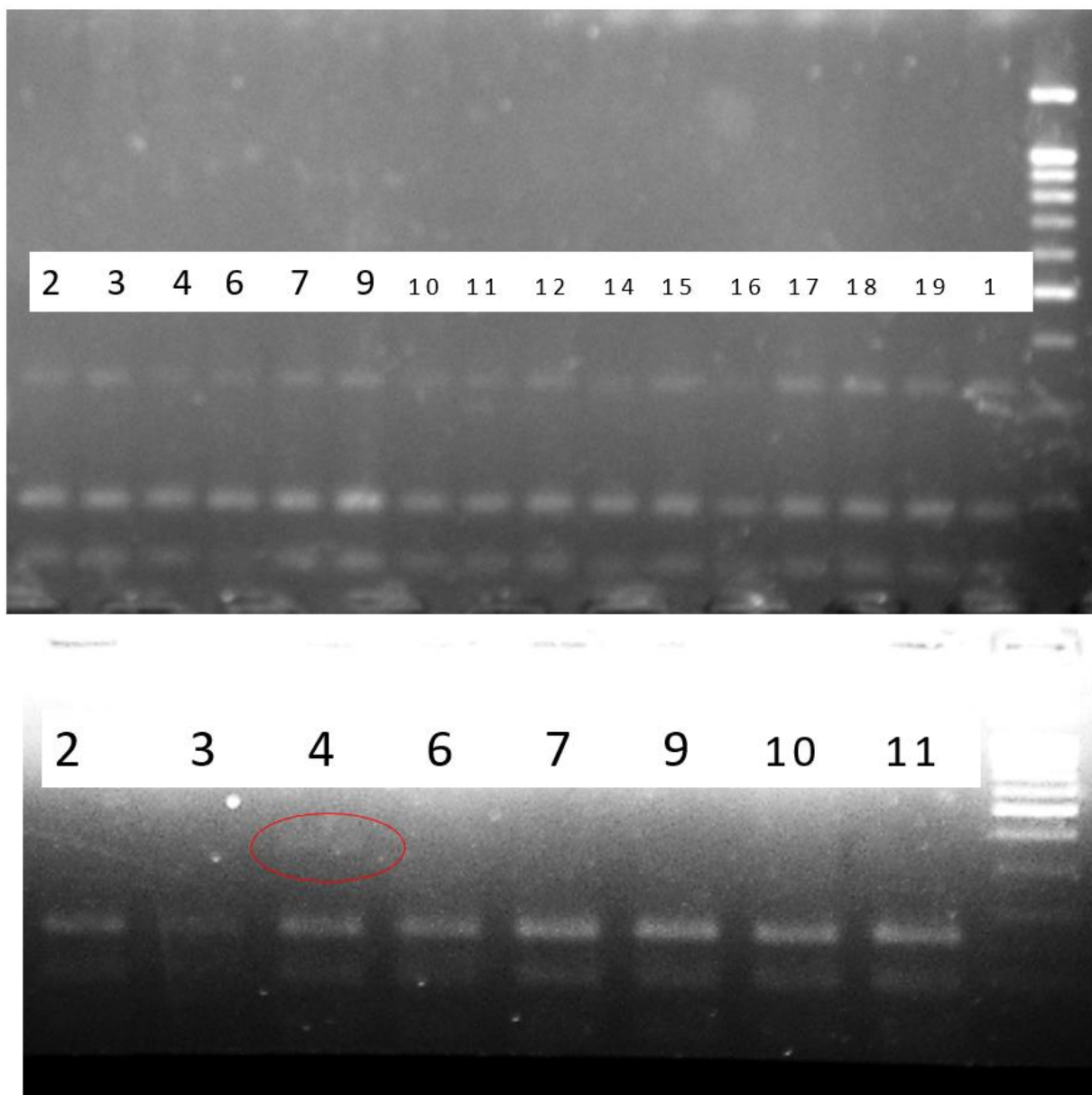


Рисунок 9. ПЦР-ПДРФ генотипирование мышей F0 по сайту внесения DSB R2 в эксперименте по удалению теломерных сайтов связывания CTCF в локусе *Slc29a3/Unc5b* и получению мутаций в сайтах L2, M2 и R2 (праймеры *Unc5b-R2-F* и *Unc5b-R2-R*). ПЦР продукты для сайта R2 после обработки ферментом рестрикции *SphI* (верхний ряд) и *HaeIII* (нижний ряд).