



На правах рукописи

САЛЬНИКОВ ПАВЕЛ АЛЕКСАНДРОВИЧ

**СВЯЗЬ ПРОСТРАНСТВЕННОЙ ОРГАНИЗАЦИИ ГЕНОМА И
РЕГУЛЯЦИИ ЭКСПРЕССИИ ГЕНОВ В ЛОКУСАХ
SLC29A3/UNC5B И *EPB41L4A* МЫШИ**

1.5.7. — Генетика (биологические науки)

АВТОРЕФЕРАТ

диссертации на соискание ученой степени
кандидата биологических наук

Новосибирск – 2026

Работа выполнена в лаборатории генетики развития ФГБНУ «Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук», г. Новосибирск.

Научный руководитель:

Фишман Вениамин Семенович

Доктор биологических наук, заведующий Сектором геномных механизмов онтогенеза ФГБНУ «Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук», г. Новосибирск

Официальные оппоненты:

Пчелина Софья Николаевна

доктор биологических наук, заведующая лабораторией молекулярной генетики человека Петербургского института ядерной физики им. Б.П. Константинова, НИЦ «Курчатовский институт», г. Санкт-Петербург

Куслий Мария Александровна

Кандидат биологических наук, старший научный сотрудник Лаборатории разнообразия и эволюции геномов ФГБНУ «Институт молекулярной и клеточной биологии Сибирского отделения Российской академии наук», г. Новосибирск

Ведущая организация:

ФГБУН Институт цитологии Российской академии наук (ИНЦ РАН), г. Санкт-Петербург

Защита диссертации состоится «__» _____ 2026 г. на утреннем заседании Диссертационного совета 24.1.239.01 на базе ФГБНУ «Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения РАН» в конференц-зале Института по адресу: 630090, г. Новосибирск, проспект Академика Лаврентьева, 10, т. (383)363-49-06, факс (383)333-12-78, e-mail: dissov@bionet.nsc.ru.

С диссертацией можно ознакомиться в библиотеке ИЦиГ СО РАН и на сайте Института <https://www.icgbio.ru/>

Автореферат разослан «__» _____ 2026 г.

Ученый секретарь
диссертационного совета,
доктор биологических наук

Т.М. Хлебодарова

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследования и степень ее разработанности. Развитие методов анализа архитектуры генома (включая технологии семейства 3C/Hi-C и картирование сайтов связывания белков хроматина) оказало значительное влияние на развитие современных представлений о роли механизмов компактизации хроматина в регуляции генной экспрессии (Lieberman-Aiden *и др.*, 2009; de Wit и de Laat, 2012; Rao *и др.*, 2014). Одним из ключевых уровней организации генома являются топологически ассоциированные домены (ТАДы), границы которых часто формируются кластерами сайтов связывания CTCF (CTCF binding sites, CTCF-bs) и контролируют частоты взаимодействий между цис-регуляторными генетическими элементами и промоторами генов (Rao *и др.*, 2014; Merckenschlager и Nora, 2016). Нарушение границ ТАДов мутациями сайтов CTCF или структурными перестройками может приводить к нарушению регуляции экспрессии генов и, как следствие, к патологиям (Ibn-Salem *и др.*, 2014; Lupiáñez *и др.*, 2015), однако выраженность и универсальность такого эффекта остаются дискуссионными: удаление отдельных CTCF-сайтов далеко не всегда вызывает значимые изменения экспрессии, а последствия зависят от контекста локуса и ткани, а деpleции основных структурных белков, участвующих в формировании ТАДов (CTCF или когезина), приводят к изменению экспрессии лишь небольшой доли генов (Nora *и др.*, 2017; Schwarzer *и др.*, 2017). Кроме того, подавляющая часть существующих исследований сфокусирована на локусах генов развития, имеющих узкий тканево-временной профиль экспрессии, и/или на моделях *in vitro* (Franke *и др.*, 2016; Kragestein *и др.*, 2018; Cova *и др.*, 2023), поэтому последствия локальных нарушений архитектуры хроматина на организменном уровне и в разных эпигенетических контекстах остаются недостаточно изученными.

В связи с этим в качестве модельных локусов для исследования были выбраны две ранее не исследованные границы ТАДов: в короткой межгенной области между *Slc29a3* и *Unc5b* (хромосома 10 мыши) и в 3'-терминальной области *Erb41l4a* (хромосома 18). Выбранные локусы содержат гены с широким, но вариативным профилем экспрессии среди тканей, имеющие медицинско-генетическую значимость и выраженные нокаутные фенотипы на модельных животных (Horii *и др.*, 1993; Bolz *и др.*, 2001; Hulková *и др.*, 2001; Taylor *и др.*, 2008; Kang *и др.*, 2010; Tadagavadi, Wang и Ramesh, 2010). Оба модельных локуса содержат границы ТАДов, имеющие высокую степень инсуляции, и консервативные среди позвоночных, что указывает на их потенциальную функциональную значимость. Кроме того, для границы локуса *EPB41L4A* человека описаны CNV, затрагивающие соответствующую границу и часто сопутствующие неврологическим нарушениям (Firth *и др.*, 2009). Исходя из этого было предположено, что мутации, ослабляющие архитектурную функцию CTCF-

сайтов в данной границе, могут приводить к измеримым эпигенетическим, транскрипционным и фенотипическим эффектам.

Таким образом, в рамках данной работы мы ставим перед собой **цель** оценить влияние нарушений пространственной архитектуры хроматина в локусах *Slc29a3/Unc5b* и *Epb41l4a* на регуляцию генной экспрессии и ее тканеспецифичность. Для достижения поставленной цели были сформулированы 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* мыши 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 мыши.
4. Получить генетически модифицированных животных с мутациями в регионе, содержащем границу ТАДов в локусе *Epb41l4a* и сравнить уровень экспрессии генов локуса *Epb41l4a* у мышей дикого типа и мышей, несущих мутации границы ТАДов в этом локусе.

Научная новизна. Впервые охарактеризованы эффекты мутаций, нарушающих структуру топологически ассоциированных доменов, для локусов *Slc29a3/Unc5b* и *Epb41l4a* мыши. Показано, что в локусе *Slc29a3/Unc5b* нарушения CTCF-опосредованной архитектуры приводят к тканеспецифичным изменениям регуляции экспрессии конститутивно активных генов, различающимся по величине и направлению эффекта. Установлено, что абберрации, затрагивающие границу ТАДов в 3'-терминальной области *Epb41l4a*, сопровождаются снижением экспрессии *Nrep* в некоторых частях мозга, а взаимодействие эпигенетических ландшафтов генов *Epb41l4a* и *Nrep* при слиянии ТАДов асимметрично.

Теоретическая и практическая значимость работы. Полученные результаты расширяют представления о роли пространственной организации хроматина в регуляции генной экспрессии *in vivo* и демонстрируют тканеспецифичный характер последствий локальных нарушений 3D-архитектуры генома. Практическая значимость работы заключается в разработке и апробации нового экспериментального подхода к исследованию эффектов мутаций цис-регуляторных элементов на основе использования гибридных животных и моноаллельного анализа экспрессии методами секвенирования нового поколения, что позволяет однозначно различать цис- и транс-эффекты и обеспечивает высокую чувствительность, сопоставимую с цифровой ПЦР. Предложенные решения, в первую очередь включая UMI-опосредованное таргетное секвенирование РНК для точной количественной оценки аллельного

дисбаланса, могут быть использованы более широко при изучении механизмов цис-регуляции в других геномных локусах и модельных системах. Дополнительно показана практическая применимость одновременного внесения коротких делеций и инверсий с помощью инструментов геномного редактирования на основе CRISPR/Cas9 в близко расположенные геномные сайты без обязательного возникновения протяженных делеций между ними, что расширяет инструментарий геномного редактирования для задач функциональной геномики.

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 *Epb41l4a*, разделяющую регуляторные ландшафты гена *Apc* 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.

Личный 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 (выполнена Торгунаковым Н.Ю. и Весной Э.).

Апробация результатов и публикации. Результаты диссертационной работы представлены в виде докладов на международных и всероссийских научных мероприятиях, включая Bioinformatics of Genome Regulation and Structure / Systems Biology (BGRS/SB-2024, 5–10.08.2024; BGRS/SB-2022, 4–8.07.2022), международные конференции «Хромосома 2023» (5–10.09.2023) и конгресс CRISPR-2023 (11–13.09.2023), а также на МНСК (2018–2020); по материалам работы опубликованы статьи в международных рецензируемых журналах (в т.ч. *Biochemistry (Mosc)*, 2018; *Int J Mol Sci*, 2023; *Scientific Reports*, 2024; *Epigenetics & Chromatin*, 2025).

Структура и объем работы. Работа включает в себя следующие разделы: введение, обзор литературы, материалы и методы, результаты, их обсуждение,

выводы. Общий объем диссертации составляет 148 страниц, 26 рисунков и 1 таблицу, библиография включает 211 научных источников.

СОДЕРЖАНИЕ РАБОТЫ

Глава 1. Обзор литературы

Современная генетика рассматривает регуляцию экспрессии генов эукариотических организмов как процесс, в высокой степени зависящий от цис-регуляторных механизмов, включая пространственные взаимодействия промоторов с удаленными регуляторными элементами (Furlong и Levine, 2018). Развитие методов анализа конформации хромосом показало, что частоты контактов этих элементов подчинены иерархии законов формирования архитектуры хроматина, и часто ограничены топологически ассоциированными доменами (ТАДами). Границы ТАДов у млекопитающих в общем случае формируются кластерами сайтов CTCF и ограничивают регуляторные ландшафты, а их нарушения могут приводить к сдвигам цис-регуляции и ассоциированы с патологиями и онкогенезом (Ibn-Salem *и др.*, 2014; Akdemir *и др.*, 2020). Механистически организация петель и доменов объясняется, в частности, моделью протягивания петли (loop extrusion) с участием когезина и CTCF, при этом первый связывается с ДНК в случайных местах и протягивает ее сквозь себя, а сайты CTCF стабилизируют границы петель (Sanborn *и др.*, 2015).

Экспериментальные данные показывают, что нарушения пространственной архитектуры способны приводить к патологическим фенотипам за счет перераспределения цис-регуляторных взаимодействий. Одним из классических примеров служат хромосомные перестройки в локусе *WNT6/INH/EPHA4/PAX3* у человека, при которых инверсии, делеции и дупликации, затрагивающие область границы ТАДов, приводят к врожденным нарушениям развития конечностей (Lupiáñez *и др.*, 2015). Аналогичные перестройки, смоделированные на синтетическом локусе мыши с помощью технологий редактирования генома, воспроизвели сходные фенотипы, при этом перестройки того же масштаба, но не затрагивающие граничные элементы, не имели фенотипического проявления. Дальнейшие исследования на других локусах показали, что нарушения функций границ ТАДов ассоциированы с наследственными заболеваниями и онкогенезом; предполагается, что такой механизм может объяснять заметную долю клинически значимых перестроек и некодирующих мутаций, поскольку изменение архитектуры доменов способно приводить к эктопической активации или репрессии генов за счет «перераспределения» энхансеров (Ibn-Salem *и др.*, 2014; Akdemir *и др.*, 2020; Pachano, Haro и Rada-Iglesias, 2022).

Вместе с тем к настоящему времени накоплены аргументы, ограничивающие универсальность гипотезы о важности пространственной архитектуры в регуляции генной экспрессии. Во многих случаях удаление CTCF-сайтов в границах ТАДов не приводит к существенным изменениям транскрипции, а глобальная деплеция CTCF или когезина вызывает значимые

сдвиги экспрессии лишь у крайне небольшой доли экспрессирующихся генов, что указывает на устойчивость регуляторных систем и наличие компенсаторных механизмов (Nora *и др.*, 2017; Rao *и др.*, 2017; Kabirova *и др.*, 2024). Показательно, что даже при последовательном удалении нескольких CTCF-сайтов и постепенном слиянии доменов (на примере локуса *Kcnj2/Sox9*), транскрипционный ответ остается умеренным, тогда как более радикальная перестройка топологии, такая как перенос границы ТАДов в новую позицию, способна приводить к резкой дисрегуляции транскрипции (Desrang *и др.*, 2019). В то же время значительная часть подобных работ выполнена на локусах генов развития и нередко в условиях *in vitro*, поэтому масштаб и значимость эффектов на организменном уровне могут быть недооценены.

Глава 2. Материалы и методы

Ключевыми методами, примененными нами в работе, стали метод геномного редактирования CRISPR/Cas9 для получения линий мышей с делециями в границах ТАДов (Mali *и др.*, 2013), а также метод анализа архитектуры хроматина capture Hi-C (cHi-C) (Downes *и др.*, 2022). Для количественной оценки транскрипционных эффектов мы разработали и применили подход, совмещающий таргетное UMI-опосредованное РНК-секвенирование, позволяющее с высокой чувствительностью проводить измерение аллельной экспрессии, с использованием групп гибридных животных, несущих экспериментальные и контрольные аллели на общем генетическом и эпигенетическом фоне. Это позволило нам проводить аллель-специфичное измерение экспрессии генов, нормируя экспрессию аллелей C57BL/6 с различающимся цис-окружением на экспрессию соответствующего аллеля CAST. Последний делит одинаковые транс-эффекты регуляции транскрипции с экспериментальным аллелем внутри одной особи, и делит одинаковое цис-окружение со всеми животными из разных групп. Такой дизайн позволил минимизировать влияние межиндивидуальных различий, условий содержания, возраста и клеточного состава тканей и преимущественно оценивать эффекты, обусловленные различиями в цис-окружении экспериментальных аллелей. Дополнительно автором использован новый подход для внесения коротких мутаций CRISPR/Cas9 в близко расположенные сайты, показавший отсутствие побочных мутаций в виде протяженных делеций. Кроме того, в работе применены современные методики секвенирования нового поколения, РНК-секвенирование и ChIP-Seq.

Глава 3. Результаты

3.1 Стратегия выбора модельных локусов

Для исследования эффектов мутаций границ ТАДов важно выбирать модельные локусы с выраженной пространственной архитектурой хроматина, а именно содержащие границу ТАДов, сформированную сайтами связывания CTCF, доступными для внесения генетических модификаций, то есть **не**

перекрывающиеся с критическими функциональными участками генов, чтобы минимизировать “побочные” эффекты, не связанные с изменениями в пространственной организации. Для анализа изменений экспрессии генов локуса в различных эпигенетических контекстах, гены этого локуса должны иметь функциональную значимость для организма, быть конститутивно активными, но в то же время подверженными тканеспецифичным различиям в регуляции в зависимости от ткани. Руководствуясь этими критериями и анализом публичных Hi-C карт, были выбраны две хорошо картированные и ранее не исследованные границы ТАДов: в короткой межгенной области между *Slc29a3* и *Unc5b* (хромосома 10 мыши) и в 3'-терминальной области *Epb41l4a* (хромосома 18).

Локус *Slc29a3/ Unc5b* у мыши и человека характеризуется наличием двух выраженных топологических доменов со сложной внутренней архитектурой, разделенных эволюционно консервативной сильной инсуляторной границей, что указывает на ее потенциальную функциональную значимость (Рис. 1А). Граница в этом локусе сформирована двумя дивергентными кластерами сайтов связывания CTCF, находящимся в межгенном районе и интронах *Unc5b* (Рис. 1Б). Кроме того, у человека гены *SLC29A3* и *UNC5B* имеют медицинскую значимость: нарушения их регуляции ассоциированы с онкологическими и наследственными заболеваниями, а также иммунными и метаболическими нарушениями. Так, мутации гена *CDH23* у человека являются частой причиной наследственной тугоухости и синдрома Ушера (Bork *и др.*, 2001); *SLC29A3* связан со спектром гистиоцитозов (Н-синдром, Faisalabad histiocytosis, семейная Rosai–Dorfman и другие) (Melki *и др.*, 2013, 2013); *PSAP* связан с лизосомными болезнями накопления с тяжелой нейродегенерацией (Hulková *и др.*, 2001); *UNC5B* является важным рецептором, регулирующим направление роста сосудов и аксонов периферических нейронов (Lu *и др.*, 2004; Navankasattusas *и др.*, 2008; Tadagavadi, Wang и Ramesh, 2010). Нокауты по большинству из этих генов летальны или вызывают серьезные фенотипические отклонения на моделях мышей (Akil *и др.*, 2006; Han *и др.*, 2012; Liu *и др.*, 2012; Nair *и др.*, 2019).

Локус *Epb41l4a*. Во второй модели мы сосредоточились на границе ТАДов, разделяющей регуляторные ландшафты клинически значимых генов *Nrep* (связь с неврологическими нарушениями (Li *и др.*, 2023)) и *Apc* (сильная ассоциация с колоректальным раком и другими опухолями (Rowan *и др.*, 2000)), расположенной в 3'-терминальной области *Epb41l4a* (Рис. 1В); эта граница консервативна между человеком и мышью и формируется массивным кластером сайтов CTCF. По данным полногеномного анализа инсуляции с использованием публичных данных, граница относится к числу наиболее сильных (в частности, в линии H1-hESC входит в 99,8-й перцентиль по индексу инсуляции), что может быть косвенным подтверждением ее функциональной значимости. Кроме того, в публичных базах данных описаны CNV человека с неизвестным статусом патогенности, затрагивающие данную границу и сопровождающиеся неврологическими проявлениями (Рис. 1Г).

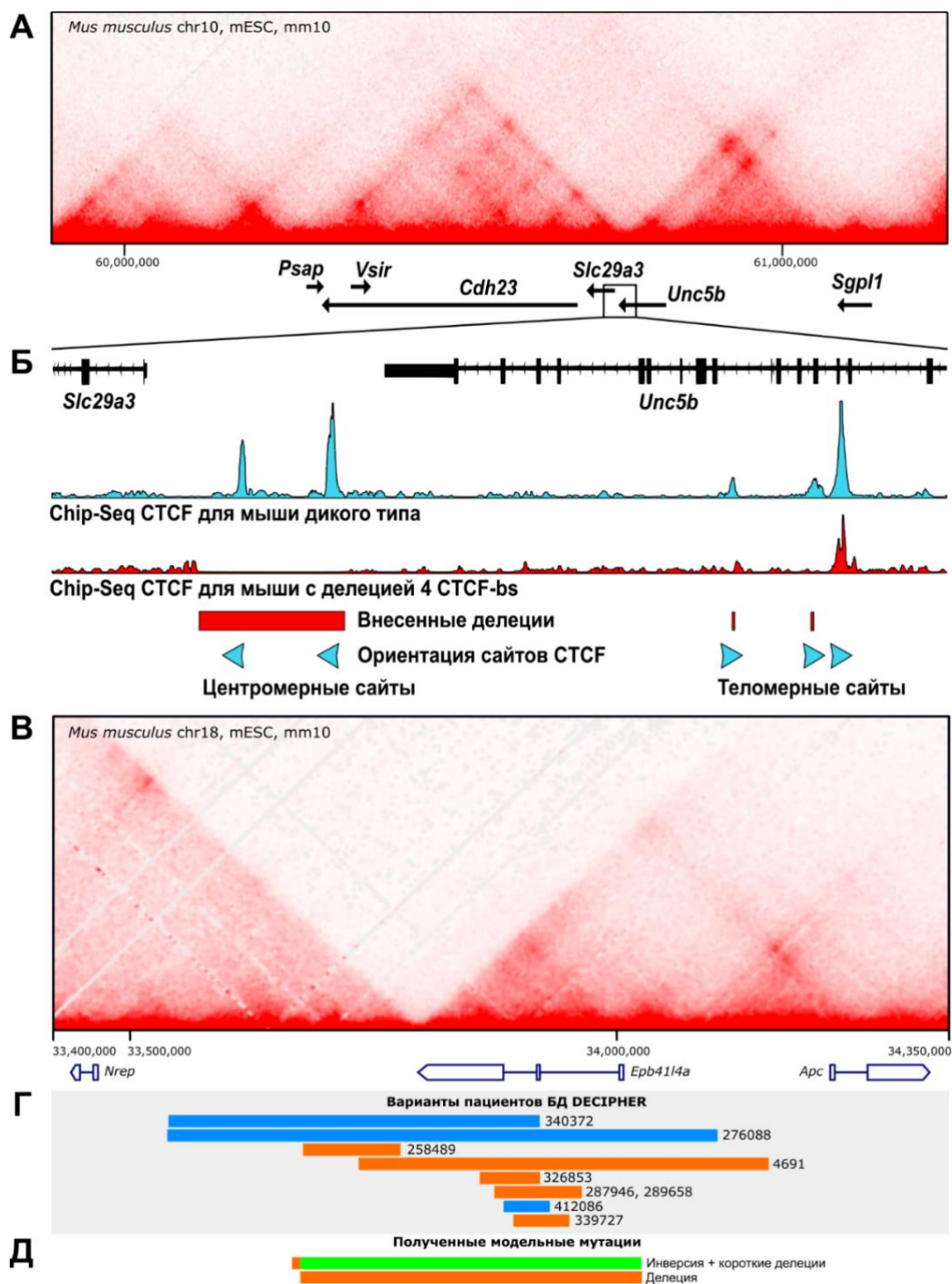


Рисунок 1. Пространственная организация выбранных локусов. А – карта Hi-C локуса *Slc29a3/Unc5b* мыши и местоположения кодирующих генов. Б – положение и ориентация сайтов связывания CTCF в локусе *Slc29a3/Unc5b*, треки CTCF ChIP-Seq из ткани печени мышей дикого типа и гомозиготных по dM1R1L2M2 и координаты внесенных мутаций аллеля dM1R1L2M2. В – карта Hi-C локуса мышинного гена *Erp41l4a* и местоположения кодирующих генов. Г – координаты мутаций у пациентов из базы данных DECIPHER. Д – координаты полученных мутаций у мышинных линий. Синий – дупликации, оранжевый – делеции, зеленый – инверсии.

Таким образом, выбранные модельные локусы не относятся к классическим для подобных исследований моделям с гомеозисными генами, при этом включают гены с широким и вариативным спектром экспрессии и биомедицинской значимостью, что делает их удобными для поиска тканеспецифичных транскрипционных последствий нарушения пространственной архитектуры хроматина.

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

Мы получили мышей с нарушением CTCF-опосредованной границы ТАДов между *Slc29a3* и *Unc5b* с использованием CRISPR/Cas9 в результате двух раундов редактирования генома. В первом мы спроектировали направляющие РНК, фланкирующие два межгенных сайта CTCF, выполнили микроинъекцию компонентов системы в зиготы C57BL/6J и получили мутантных животных. В результате генотипирования отобрали животное, несущее аллель с целевой протяженной делецией ~5,5 т.п.н. (dM1R1), и после возвратного скрещивания и скрещивания потомства получили стабильную гомозиготную линию. Затем на полученной линии выполнили мультиплексную модификацию близко расположенных CTCF-bs, расположенных в интронах гена *Unc5b* с использованием трех наборов направляющих РНК и длинных одноцепочечных олигонуклеотидов, используемых в качестве доноров гомологии для репарации. В результате мы выяснили, что в таком экспериментальном дизайне среди мутаций преобладали исходы репарации по микрогомологичному соединению концов (ММЕJ), при отсутствии побочных мутаций в виде делеций между разными сайтами внесения двуцепочечных разрывов. Несмотря на это нами была получена гомозиготная линия мышей, несущих компаунд-аллель (dM1R1L2M2) с удаленными четырьмя из пяти образующих целевую границу ТАДов CTCF-bs. Для данного аллеля мы верифицировали утрату связывания CTCF в целевом регионе методом ChIP-seq. В последующих экспериментах использовали линии dM1R1 и dM1R1L2M2 (Рис. 1Б).

3.3 Изменения пространственной архитектуры локуса *Slc29a3/Unc5b* у мышей, гомозиготных по аллелю dM1R1 и dM1R1L2M2

Мы проанализировали пространственную организацию локуса *Slc29a3/Unc5b* у полученных линий мышей, методом cHi-C в печени, почке и мозжечке и показали, что делеции CTCF-сайтов приводят к ослаблению инсультаторной функции границы и перестройке архитектуры петель хроматина (Рис. 2А-Д). В частности, во всех тканях исчезают петли, которые в норме формируются удаленным кластером CTCF и направлены в сторону ТАДа *Slc29a3* (контакты промоторной области *Slc29a3* с участками в теле *Cdh23* и областью *Psap*), тогда как петли, направленные в противоположную сторону, сохраняются.

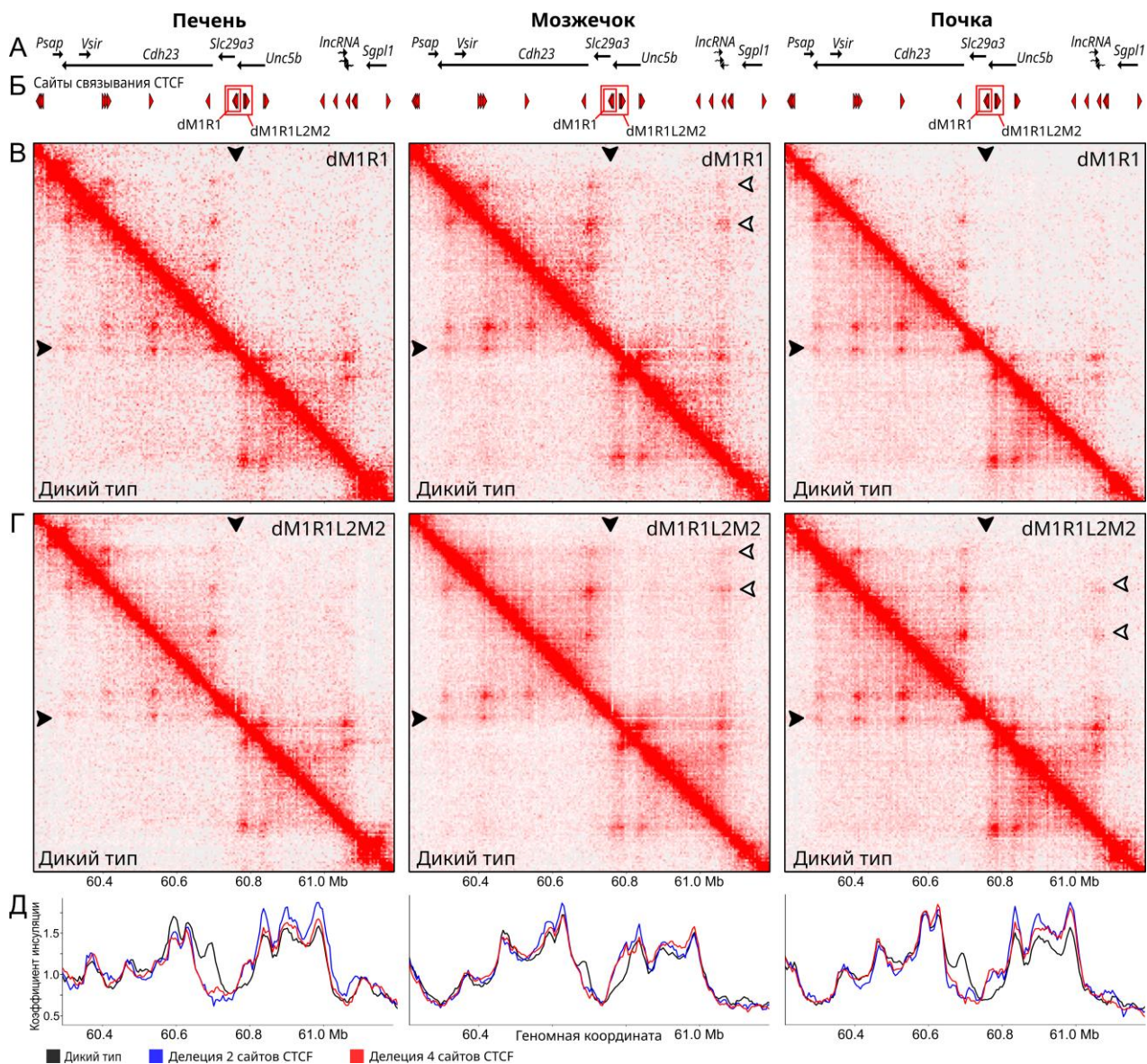


Рисунок 2. А — Положение и направление транскрипции генов локуса. Б — Положение и ориентация сайтов связывания CTCF, показаны удаленные сайты для аллелей dM1R1 и dM1R1L2M2. В, Г — Карты пространственных контактов локуса *Slc29a3/Unc5b* для печени, мозжечка и почки в координатах chr10:60103000-61356000, mm10. Над главной диагональю представлены карты локуса для мышей, гомозиготных по аллелю dM1R1 (В) или dM1R1L2M2 (Г), под диагональю — для мышей дикого типа. Черные стрелки отмечают исчезнувшие петли модифицированной границы ТАДов, белые стрелки — новые петли, преодолевающие границу ТАДов.

Одновременно участок от исходной границы до промотора *Cdh23*, включающий *Slc29a3*, инсулируется от своего исходного цис-окружения, и приобретает выраженные взаимодействия со всем ТАДом *Unc5b*. Таким образом, функциональная граница доменов смещается к промотору *Cdh23*, переводя *Slc29a3* в новое пространственное окружение ТАДа *Unc5b*. Кроме того, мы наблюдаем появление новых эктопических петель между внутренними CTCF-

сайтами в гене *Cdh23* и дистальной границей ТАДа *Unc5b*; эти взаимодействия пересекают границу ТАДов и демонстрируют тканеспецифичный характер, усиливаясь от печени к мозжечку, что согласуется с более высокой общей активностью локуса в ткани, и сонаправлены с ослаблением границы ТАДов на фоне делеции двух и четырех CTCF-bs в аллелях dM1R1 и dM1R1L2M2.

3.4 Изменения экспрессии генов *Psap*, *Vsir*, *Cdh23*, *Slc29a3*, *Unc5b*, и *Sgpl1* на фоне делеционных аллелей dM1R1 и dM1R1L2M2

Поскольку ожидаемые транскрипционные последствия нарушения границы ТАДов в негемеозисном локусе *Slc29a3/Unc5b* предполагались умеренными (Despang *и др.*, 2019), для их выявления потребовался подход, обеспечивающий точную оценку небольших изменений экспрессии генов. Экспериментальный дизайн сочетал использование гибридных животных и аллель-специфическое измерение экспрессии. Были получены три группы гибридов F1 C57BL/6 × CAST, у которых один аллель локуса происходил от линии CAST, а второй был представлен аллелем дикого типа C57BL/6, dM1R1 или dM1R1L2M2 (WT/CAST, dM1R1/CAST и dM1R1L2M2/CAST соответственно). Это позволило нам проводить аллель-специфичное измерение экспрессии генов, нормируя экспрессию аллеля C57BL/6, находящегося в цис-положении с мутантной или дикого типа границей ТАДов, на экспрессию аллеля CAST того же гена. В пределах одной особи оба аллеля находились в общем транс-регуляторном окружении, тогда как локальное цис-окружение референсного аллеля CAST оставалось неизменным во всех экспериментальных группах. Такой дизайн позволил минимизировать влияние средовых различий и преимущественно оценивать эффекты, обусловленные различиями в цис-окружении экспериментальных аллелей. Для точной оценки аллельной экспрессии мы разработали и применили UMI-опосредованное таргетное РНК-секвенирование. Межлинейные SNP позволяют отнести прочтения к аллелям C57BL/6 или CAST, а UMI – провести точную оценку уровня экспрессии гена.

Используя данный подход, мы количественно измерили соотношение аллельных транскриптов шести генов локуса (*Psap*, *Vsir*, *Cdh23*, *Slc29a3*, *Unc5b*, *Sgpl1*) в пяти органах (почка, печень, мозжечок, обонятельная луковица, мочевой пузырь) и выявили 20 статистически значимых случаев изменений из 60 проанализированных комбинаций (ген/орган/генотип). Наблюдаемые сдвиги экспрессии, как правило, не превышали ~50%, а их паттерн оставался тканеспецифичным и нередко разнонаправленным; при этом часть эффектов воспроизводилась в обоих генотипах, а часть проявлялась только на фоне аллеля dM1R1, то есть при более слабом нарушении инсуляции в границе ТАДов, что может указывать на неизвестные механизмы, опосредующие связь пространственной архитектуры хроматина и регуляции генной экспрессии (Рис. 3А,Б).

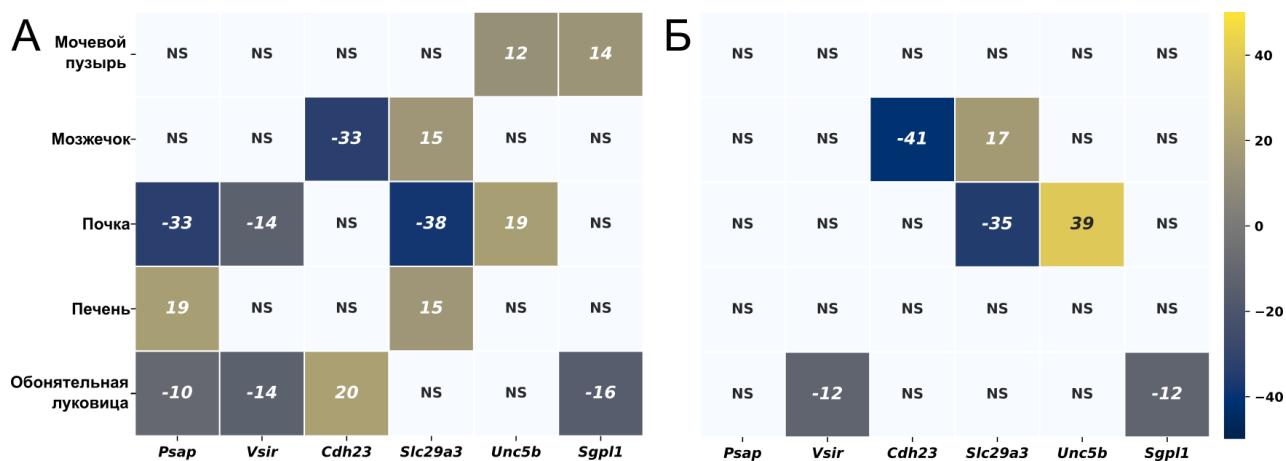


Рисунок 3. Результаты измерения экспрессии генов локуса *Slc29a3/Unc5b* у мышей, несущих аллели dM1R1 (А) или dM1R1L2M2 (Б), полученные с помощью UMI-опосредованного таргетного РНК-секвенирования. Представлены значения изменений экспрессии аллелей генов на фоне делеции сайтов связывания CTCF относительно аллелей с цис-окружением дикого типа в процентах. Отмечены статистически значимые результаты (FDR <0,1).

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

В рамках экспериментов с вторым модельным локусом, мы получили две линии мышей с хромосомными перестройками, нарушающими границу ТАДов в локусе *Erb41l4a*, используя систему генетического редактирования CRISPR/Cas9, нацеленную на геномные сайты, фланкирующие протяженный фрагмент, включающий массивный границеобразующий кластер CTCF-bs (чтобы исключить компенсаторную функцию соседних сайтов при делеции (Barutcu *и др.*, 2018; Despang *и др.*, 2019)). В результате эксперимента по редактированию генома было получено 4 животных F0, и в последующих скрещиваниях получены стабильные гомозиготные линии с делецией ~350 тыс. п.н. (chr18:33673967–34025702) и с инверсией того же региона, однако совмещенные с побочными незначительными для эксперимента делециями вблизи сайтов внесения двуцепочечных разрывов (~7,8 тыс. п.н. и 27 п.н.), обнаруженные после оптимизации ПЦР-систем генотипирования (Рис. 1Д). Несмотря на внесенные мутации, животные обеих линий были жизнеспособны и фертильны, не имели очевидных фенотипических отклонений, а контроль потенциальных off-target мутаций в сцепленных кандидатных сайтах не выявил нежелательных побочных мутаций.

3.6 Транскрипционный эффект хромосомных перестроек, затрагивающих границу ТАДов в локусе *Erb41l4a*

Мы оценили транскрипционные последствия делеции и инверсии региона, содержащего границу ТАДов в локусе *Erb41l4a*, с помощью RNA-seq в мозжечке и обонятельной луковице у самцов, гомозиготных по внесенным мутациям, используя дикий тип как контроль. При делеции экспрессия *Erb41l4a* не

детектировалась в обоих органах, тогда как при инверсии экспрессия *Epb41l4a* в мозжечке возрастала более чем в четыре раза, что согласуется с механизмом «кражи энхансера» после перемещения промотора в новое цис-окружение (Рис. 4А,Б) (Northcott и др., 2014; Pachano, Haro и Rada-Iglesias, 2022). Одновременно экспрессия *Nrep* снижалась более чем на 70% в мозжечке как при делеции, так и при инверсии, и снижалась в обонятельной луковице при инверсии (Рис. 4В), при этом экспрессия *Arc* значимо не изменялась ни при одном из условий (Рис. 4Г).

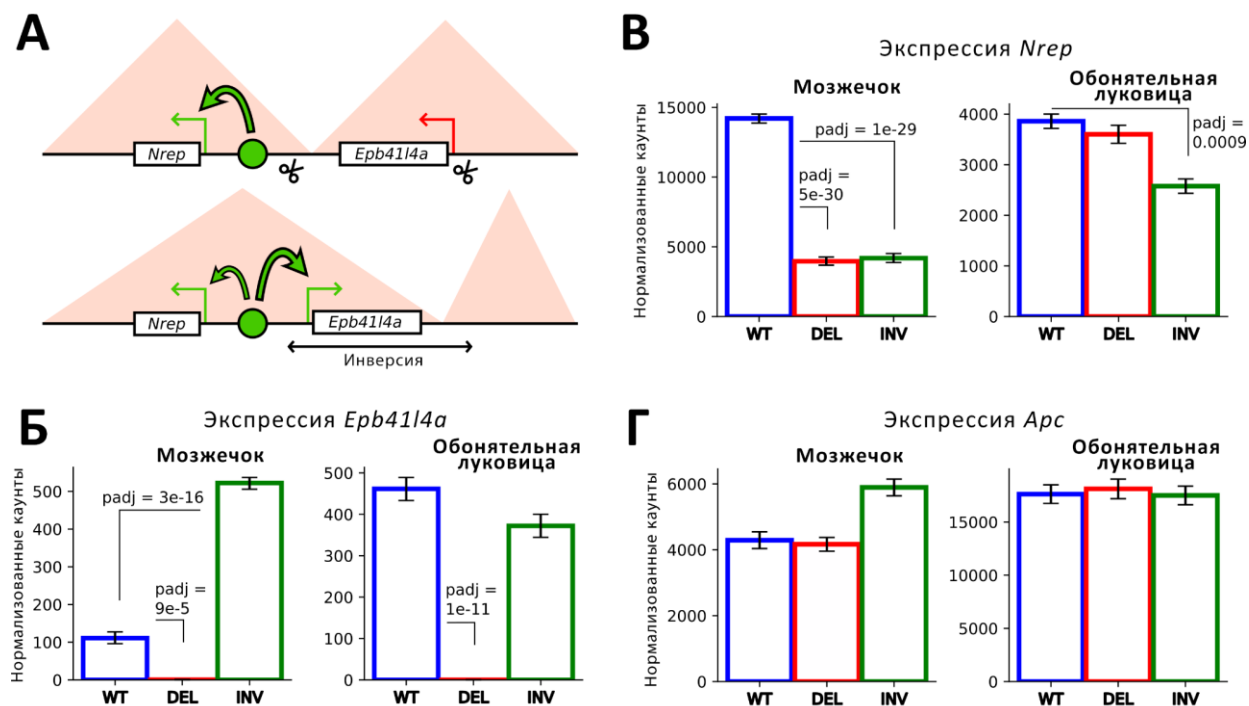


Рисунок 4. А — Схема механизма кражи энхансера на модели инверсии в локусе *Epb41l4a*. Зеленый круг обозначает энхансер *Nrep*, символами ножниц отмечены мишени CRISPR/Cas9. Зеленый и красный цвета промоторов обозначают высокую и низкую транскрипционную активность соответственно. Б-В — уровни экспрессии *Epb41l4a* (Б), *Nrep* (В) и *Arc* (Г) в мозжечке и обонятельной луковице для всех экспериментальных групп. WT — уровень экспрессии у мышей дикого типа, INV — у мышей с инверсией, DEL — с делецией целевого участка. По оси Y приведены нормализованные значения, рассчитанные с помощью DESeq2.

Глава 4. Обсуждение

Полученные результаты выступают подтверждением гипотезы о том, что функциональная роль архитектурных цис-элементов генома, таких как сайты связывания CTCF и границы ТАДов, и их роль в регуляции генной экспрессии, определяются эпигенетическим контекстом и исходной активностью генов. Таким образом, мутации, нарушающие эту функцию, могут приводить к различным, в том числе разнонаправленным, транскрипционным ответам в разных органах.

В ходе получения генетически модифицированных животных, мы применили мультиплексное редактирование соседних сайтов связывания CTCF, планируя ввести короткие мутации с использованием длинных олигонуклеотидов, что обеспечивает внесение предсказуемых мутаций. Однако в зиготах мыши данный дизайн оказался малоэффективным: большинство событий репарации происходило по пути ММЕJ. Тем не менее, нам удалось получить аллель с короткими мутациями, одновременно нарушающими два сайта связывания CTCF. Таким образом, для подобных задач более перспективным является дизайн, изначально ориентированный на ММЕJ (Shen *и др.*, 2018) и, как показал данный эксперимент, применимый для редактирования нескольких близкорасположенных геномных сайтов.

В локусе *Slc29a3/Unc5b* внесенные нами делеции CTCF-сайтов приводят к ослаблению функции границы и реорганизации петлевой архитектуры. Разделенные этой границей ТАДы не сливаются, однако граница смещается к сайту связывания CTCF внутри домена, переводя область *Slc29a3* в новое цис-окружение ТАДа *Unc5b*. Параллельно возникают новые эктопические контакты, пересекающие область границы (Рис. 2). Их происхождение плохо укладывается в простые модели пространственной архитектуры и указывает на вовлечение иных механизмов, таких как фазовая сепарация типов хроматина, различающихся по белковому составу (Nuebler *и др.*, 2018).

Изменения транскрипции продемонстрировали несколько сценариев. Во-первых, «усреднение» экспрессии генов, разделенных границей ТАДа: после ослабления инсуляции регуляторные ландшафты начинают конкурировать и взаимно влиять на экспрессию генов. При этом репрессированные гены попадают под активирующее воздействие цис-регуляторных факторов активных генов и *vice versa* (например, разнонаправленные изменения *Slc29a3* и *Unc5b* в почке, Рис. 3А,Б). Во-вторых, сценарий «кражи энхансера», при котором более активная регуляторная среда одного домена активирует соседний промотор без снижения активности своих исходных мишеней (например, увеличение экспрессии *Slc29a3* в мозжечке на фоне высокой активности окружения *Unc5b* и наличия кандидатного суперэнхансера). Часть наблюдений все еще остается труднообъяснимой, что подчеркивает сложность связи регуляторных механизмов и пространственной организации хроматина.

В нашем эксперименте большинство выявленных транскрипционных эффектов умеренны и обычно не превышают 50%, а различия порядка 10–20% близки к нижнему порогу детекции (Рис. 3А,Б). Тем не менее, такая амплитуда молекулярного ответа согласуется с данными об устойчивости регуляторных ландшафтов к пертурбациям, вызванным ослаблением границ ТАДов (Despang *и др.*, 2019), но противоречит высокой эволюционной консервативности расположения этих границ (Dixon *и др.*, 2012; Kentepozidou *и др.*, 2020). Это может объясняться тем, что компенсаторные механизмы способны нивелировать эффект мутаций до тех пор, пока они не будут устранены рекомбинацией или

отбором. Кроме того, даже такая дисрегуляция транскрипции может быть эволюционно значимой и находиться под давлением отбора. Наконец, функция границ может быть критична лишь в узкий период развития организма или для небольшой группы клеток, транскрипционный ответ которых мы не смогли оценить. В таком случае тканеспецифичность функции границ разрешает это противоречие.

Отдельно мы подчеркиваем ценность разработанного подхода к высокочувствительной количественной оценке малых цис-эффектов с помощью аллель-специфичного УМП-опосредованного таргетного секвенирования РНК у гибридных животных. Этот подход расширяет возможности анализа эффектов мутаций архитектурных и цис-регуляторных элементов.

Более крупные перестройки, удаляющие либо инвертирующие область сильной консервативной границы локуса *Epb41l4a*, дают более отчетливые эффекты. Как делеция, так и инверсия приводят к значительной дисрегуляции *Nrep* (Рис. 4В). Хотя в обоих случаях *Nrep* потенциально получает доступ к регуляторным элементам соседнего ТАДа, происходит слияние домена *Nrep* с разными участками ТАДа *Apc*. Поэтому маловероятно, что эктопические взаимодействия с различными наборами цис-элементов могли бы приводить к практически одинаковому снижению экспрессии *Nrep* в мозжечке. Гипотеза о действии репрессивных элементов (Kraft *и др.*, 2022; Pang *и др.*, 2023) соседнего ТАДа на расстоянии более 300 тыс. п. н. также выглядит маловероятной с учетом отсутствия признаков гетерохроматина в домене *Apc*. Кроме того, снижение транскрипции *Nrep* в обонятельной луковице при инверсии противоречит ожиданию, поскольку в этом случае *Nrep* попадает в активное эпигенетическое окружение *Epb41l4a*.

С учетом этого более правдоподобно, что снижение экспрессии *Nrep* связано не столько с приобретением новых контактов, сколько с утратой взаимодействия с областью границы. Мы предполагаем, что в регионе границы могут находиться цис-регуляторные элементы, контролирующие экспрессию *Nrep* или необходимые на ранних этапах формирования регуляторной программы нервной ткани. Альтернативно, граница может выполнять не только инсульторную функцию, но и быть критически необходимой для пространственной координации цис-регуляторных элементов локуса (Symmons *и др.*, 2014). В целом результаты подчеркивают, что нарушение пространственной архитектуры хроматина само по себе может снижать экспрессию генов со сложной тканеспецифической регуляцией. В нашем исследовании это наиболее наглядно проявляется для *Nrep* и *Cdh23* в тканях с их повышенной экспрессией. Эти данные позволяют предположить молекулярный механизм действия части CNV пациентов, перекрывающих данную границу и сопровождающихся неврологическими симптомами (Firth *и др.*, 2009).

Модель инверсии позволила протестировать гипотезу о симметричности взаимодействия цис-регуляторных ландшафтов разных генов. Предполагается, что если цис-регуляторные элементы одного гена могут активировать промотор второго, то и элементы второго смогут активировать промотор первого, поскольку оба промотора относятся к одному классу в терминах промотор-энхансерной специфичности. Инверсия переносит промотор *Epb41l4a* в домен *Nrep*, что приводит к активации экспрессии *Epb41l4a* в мозжечке, где активен *Nrep*, и хорошо объясняется существующими моделями (Рис. 4А) (Pachano, Nago и Rada-Iglesias, 2022). Одновременно экспрессия *Nrep* снижается как при инверсии, так и при делеции, что можно интерпретировать как проявление взаимного влияния регуляторных ландшафтов. При этом в обонятельной луковице, где, напротив, активен *Epb41l4a*, активации *Nrep* не происходит, что указывает на асимметричный активаторный потенциал регуляторных ландшафтов этих генов.

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 основных публикаций по теме диссертационной работы:

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. IF= 2.2
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. IF= 4.9
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. IF= 3.9
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. IF= 3.5