Автономная некоммерческая образовательная организация высшего образования
«Научно-технологический университет «Сириус»

На правах рукописи



Дудко Наталия Александровна

РЕГУЛЯТОРНАЯ ГЕНОМИКА АГРЕССИВНОГО ПОВЕДЕНИЯ

1.5.7. Генетика

Диссертация
на соискание ученой степени
кандидата биологических наук

Научный руководитель:
доктор биологических наук, профессор,
академик РАН
Рогаев Евгений Иванович

Федеральная территория «Сириус» – 2026

ОГЛАВЛЕНИЕ

Введение ... 6
Актуальность темы исследования и степень ее разработанности ... 6
Цель исследования: ... 7
Задачи исследования: ... 7
Научная новизна ... 7
Теоретическая и практическая значимость ... 7
Методология и методы исследования ... 8
Положения, выносимые на защиту: ... 9
Глава 1. Обзор литературы ... 10
1. Генетическая архитектура агрессивного поведения ... 10
1.1. Кандидатные гены и полигенная архитектура ... 10
1.2. Транскриптомная архитектура агрессивного поведения ... 11
1.2.1. Транскриптомные изменения и регуляторные сети ... 11
1.3. Животные модели генетики поведения ... 12
1.3.1. Грызуны и мухи Drosophila melanogaster ... 13
1.3.2. Рыбы, птицы и домашние виды ... 14
1.3.3. Общие закономерности и ограничения моделей ... 14
1.4. Одомашнивание и поведенческая генетика пушных зверей: норка и серебристо-черная лисица. 14
1.4.1. История поведенческой селекции и эксперимент Д.К. Беляева ... 15
1.4.2. Американская норка: селекция темперамента и генетика агрессии ... 16
1.4.3. Серебристо-черная лисица: поведенческая и генетическая архитектура ... 17
1.4.4. Транскриптомные, нейрохимические и эндокринные корреляты ... 17
1.4.5. Нейроморфологические изменения мозга лисиц ... 18
1.5. Эпигенетическая регуляция агрессивного поведения у человека и животных ... 19
1.5.1. Метилирование ДНК и ранний опыт ... 19
1.5.2. Некодирующие РНК и природные модели ... 20
1.5.3. Гистоновые модификации и регуляторные элементы: H3K4me3, H3K27ac и ChIP-seq ... 21
Функциональное значение H3K4me3 и H3K27ac ... 21
Гистоновые модификации и агрессивное поведение ... 21
1.7. Интеграция геномных, транскриптомных и эпигеномных данных в исследовании сложных поведенческих фенотипов ... 22
1.7.1. От генетических вариантов к регуляции экспрессии ... 23
1.7.2. Совместный анализ экспрессии и регуляторных элементов ... 23
1.7.3. Анализ сетей коэкспресии и интеграция данных лисицы ... 24
Глава 2. Материалы и методы ... 25
2.1 Полногеномные данные ... 25
2.1.1. Коллекция биологического материала ... 25
2.1.1.1. Выделение ДНК, пробоподготовка геномных библиотек, секвенирование ... 25
2.1.2. Анализ данных полногеномного секвенирования ... 26
2.1.2.1. Выявление однонуклеотидных вариантов и генотипирование выборки ... 26
2.1.2.2. Фильтрация вариантов и фазирование ... 27
2.1.2.3. Анализ популяционной структуры ... 28
Импорт данных и прунинг (прореживание) по неравновесию по сцеплению ... 28
Анализ главных компонент ... 28

Анализ генетического происхождения с помощью ADMIXTURE....29
Визуализация....29
2.1.2.4. Выявление ОНВ, специфичных для агрессивной линии, окна ОНВ и гаплотипный анализ....29
Выявление ОНВ, специфичных для агрессивной линии (фильтр частот 80/10/10)....29
Визуализация плотности ОНВ по всему геному....30
Формирование и аннотирование ОНВ-окон....30
Гаплотипный тест в пределах окон....31
2.1.2.5. Анализ однонуклеотидных вариантов в геноме норки....31
2.2. Транскриптомы....32
2.2.1 Выделение тотальной РНК, приготовление и секвенирование транскриптомных библиотек....32
2.2.2 Анализ данных транскриптомного секвенирования....33
2.2.2.1. Первичная обработка транскриптомных данных....33
2.2.2.2. Анализ главных компонент и частичный дискриминантный анализ методом наименьших квадратов....33
2.2.2.3. Анализ дифференциальной экспрессии отдельных генов....34
2.2.2.4. Построение сетей коэкспрессии....35
2.2.2.3. Анализ сетей коэкспрессии генов и идентификация хаб-генов....36
2.3 Анализ регуляторных элементов генома....37
2.3.1. Исследуемая выборка....37
2.3.2. Получение иммунопреципитированного хроматина....37
2.3.3. Приготовление библиотек....38
2.3.4. Первичная биоинформатическая обработка данных секвенирования....38
2.3.5. Попарное сравнение локусов обогащения прочтениями между поведенческими группами....39
2.4. Мультиомиксный интегративный анализ геномных, транскриптомных и эпигеномных данных, полученных в рамках данной работы....40
Глава 3. Результаты и обсуждение....41
3.1. Полногеномные данные....41
3.1.1. Когорта и набор вариантов....41
3.1.2. Анализ главных компонент....41
3.1.3. Анализ предковых компонент (ADMIXTURE)....43
3.1.4. Родство внутри отдельных линий....44
3.1.5. Анализ FST по окнам между селектированными линиями....44
3.1.6. Выявление ОНВ, специфичных для животных агрессивной линии....45
3.1.7. Анализ распределения ОНВ в геноме....46
3.1.8. Окна, содержащие гены определенных функциональных групп: корреляция с ранее опубликованными данными....47
3.1.9. Анализ выявленных миссенс-вариантов, специфичных для агрессивных лисиц....48
3.1.11. Интеграция общей геномной популяционной структуры и окон ОНВ....55
3.2. Транскриптомный анализ фронтальной коры лисиц разных линий....58
3.2.1. Сходство транскриптомных профилей агрессивных и неселектированных лисиц....58
3.2.2. Дифференциально экспрессирующиеся гены у агрессивных лисиц....60
3.2.3. Анализ коэкспрессии и направленных регуляторных сетей....63
3.2.3.1. Модули коэкспрессии во фронтальной коре....65
3.2.3.2 Модули коэкспрессии в миндалевидном теле....68
3.3. Эпигеномный анализ....71
3.3.1. Первичная оценка данных секвенирования иммунопреципитированного хроматина....71
3.3.2. Попарное сравнение профилей иммунопреципитированного хроматина с синхронным

сигналом от обеих гистоновых меток.....73
3.3.3. Попарное сравнение профилей иммунопреципитированного хроматина с сигналом от гистоновой метки H3K4me3.....78
3.3.4. Попарное сравнение профилей иммунопреципитированного хроматина с сигналом от гистоновой метки H3K27ac.....80
3.3.5. Сопоставление с ранее опубликованными данными по экспрессии.....83
3.4. Интегративный анализ.....85
3.4.1.1. Специфичные для агрессивной линии OHB, совпадают с участками регуляторного хроматина с дифференциальным обогащением.....85
3.4.1.2. Дифференциально экспрессирующиеся гены внутри окон OHB.....87
3.4.1.3. Взаимно противоположные изменения экспрессии в локусе CTXN3–CCDC192.....88
3.4.1.4. Модули коэкспрессии объединяют гены-кандидаты в сети, коррелирующие с агрессивным поведением.....89
3.4.2. Регуляторная архитектура поведенческого отбора.....90
3.4.2.1. Протеолитическое ремоделирование нейронных цепей.....91
3.4.2.2. Гипоксическая и метаболическая сигнализация.....91
3.4.2.3. Ось «лизосомы – нейровоспаление» в миндалевидном теле: гены MFSD8, NLRP12, BST1, PPP1R17 и DRD4.....92
3.4.2.4. Внеклеточный матрикс и перинейрональные сети.....93
3.4.2.5. Нейрональная полярность и кортикальная идентичность.....93
3.4.2.6. Связь с геномикой одомашнивания лис.....94
ЗАКЛЮЧЕНИЕ.....95
ВЫВОДЫ.....97
Список литературы.....98
Приложение А
Список образцов с фенотипами, используемых в полногеномном анализе.....117
Приложение Б
Список образцов с фенотипами, используемых в транскриптомном анализе.....118
Приложение В
Статистика секвенирования геномных библиотек.....118
Приложение Г
Множественные выравнивания белковых последовательностей, содержащих миссенс-замены, специфичные для агрессивных линий, для 10 организмов.....121
Приложение Д
Дендрограммы построения генных сетей для (а) фронтальной коры и (б) миндалевидного тела. По оси ординат отмечена высота дендрограмм.....122
Приложение Е
Состав модулей коэкспрессии генов с наибольшим по модулю коэффициентом корреляции при сравнении фронтальной коры агрессивных с ручными и неселектированными животными.....123
Приложение Ж
Состав модулей коэкспрессии генов с наибольшим по модулю коэффициентом корреляции при сравнении миндалевидного тела агрессивных с ручными и неселектированными животными....124
Приложение З
Статистика секвенирования ChIP-seq библиотек, приготовленных из ДНК фронтальной коры головного мозга лис. Необогащенная ДНК - образец неиммунопреципитированной ДНК, использованный в качестве отрицательного контроля.....125
Приложение И
Аннотация локусов с одновременным дифференциальным уровнем метилирования и ацетилирования гистона H3, находящихся на расстоянии не более 3000 пн от TSS аннотированных генов лисы. Сравнение агрессивных особей (A) с ручными (T) и неселекционированными (N), logFCcp - усредненное значение logFC для 2 сравнений по 2 гистоновым меткам.....127

Приложение К
Пересечение локусов с дифференциальным уровнем метилирования или ацетилирования гистона H3, находящихся на расстоянии не более 3000 пн от TSS аннотированных генов лисы, с координатами ОНВ, специфичных для агрессивной линии. Сравнение агрессивных особей (A) с ручными (T) и неселектированными (N).....131

Введение

Актуальность темы исследования и степень ее разработанности

Поведение – это комплексный фенотип с множеством вариаций. Изучение механизмов возникновения и развития агрессивного поведения представляет собой крайне актуальную научную задачу. Многие нейродегенеративные и психиатрические заболевания ассоциированы с повышенной агрессивностью. Поведение является сложным признаком, и пока нет достаточного количества исследований, чтобы точно сказать, сколько генов вовлечено в процесс доместикации и как они взаимодействуют.

Традиционно агрессивное поведение представляет собой сложное социальное намеренное поведение, направленное на нанесение физического повреждения другому индивидууму [Smith, Leiras, 2018]. Оно характерно для большинства позвоночных и некоторых беспозвоночных, и обеспечивает выживание в условиях конкуренции за ограниченные ресурсы (пищу, территорию), соперничестве за партнеров при спаривании и защиту [Ferguson, 2010; Lischinsky, Lin, 2020; Traniello и др., 2023; Wei, Talwar, Lin, 2021]. Агрессия является эволюционно закрепленной формой поведения животных и человека, и сама по себе не является патологией [Haller, Kruk, 2006]. Однако такие психические расстройства, как биполярное аффективное расстройство (БАР), синдром дефицита внимания и гиперактивности (СДВГ), шизофрения, проявляющиеся во враждебности, угрозах и нападениях, ассоциированы с искаженными формами агрессивного поведения [Bhavana, Vydyanath, Lakshmi, 2023; Teksin и др., 2023]. Изучение агрессивного поведения, помимо вклада в фундаментальную нейрогенетику и реконструкцию генетических механизмов, контролирующих различные поведенческие паттерны, также позволит разработать подходы для прогнозирования предрасположенностей к разным формам социального поведения на основе характеристик генотипа, в перспективе – разработки новых подходов для коррекции патологических форм поведения, являющихся, в том числе частью проявления заболеваний.

Ранее на основе данных секвенирования транскриптомов нескольких тканей лис с разными типами поведения были получены результаты, свидетельствующие об изменении уровня экспрессии ряда генов у животных с ручным или агрессивным поведением [Kukekova и др., 2011]. Это говорит о том, что изменение поведения в линиях лис может быть связано не только с генетическими вариациями, но и, вероятно, с изменениями в механизмах регуляции работы генов, включая возможный вклад эпигенетических факторов.

На сегодняшний день нет достоверных данных об эпигенетической регуляции поведения на модели лисы *Vulpes vulpes*. Это делает актуальным применение эпигенетических методов, прежде всего анализа посттрансляционных гистоновых модификаций, и их интеграцию с

геномными и транскриптомными данными для выявления фундаментальных механизмов формирования агрессивного поведения.

Цель исследования:

Выявление генетических вариантов, хромосомных локусов и молекулярных механизмов, ассоциированных с агрессивным поведением у лисицы обыкновенной (*Vulpes vulpes*), путем интеграции данных полногеномного секвенирования селектированных и неселектированной линий с транскриптомными и эпигеномными профилями головного мозга.

Задачи исследования:

1. Получить полногеномные данные и провести на их основе поиск генетических вариантов в выборках лисиц, представляющих три линии (селектированные на агрессию/дружелюбие и неселектированную), и идентифицировать геномные регионы и генетические варианты, специфичные для агрессивной линии.
2. Получить транскриптомные данные для фронтальной коры и миндалины и провести поиск генов с дифференциальной экспрессией между агрессивной и ручной/неселектиуемой линиями.
3. Получить эпигенетические профили гистоновых меток H3K4me3 и H3K27ac фронтальной коры, сформировать карты регуляторных элементов (промоторы, энхансеры) и выявить локусы со статистически значимыми различиями эпигенетических профилей в разных линиях.
4. Провести интеграцию данных о генетических вариантах, экспрессии генов и состоянии хроматина для выявления потенциальных молекулярно-генетических механизмов регуляции поведения.

Научная новизна

Впервые на основе данных индивидуального полногеномного секвенирования 46 животных (со средним покрытием >4x) агрессивной, ручной и неселектированной линий лисицы с использованием актуальной хромосомной сборки генома VulVul3 выявлены генетические варианты, специфичные для животных агрессивной линии. Впервые для данной модели получены эпигенетические профили фронтальной коры и выявлены специфичные для агрессивной линии локусы, обогащенные гистоновыми метками H3K4me3 и H3K27ac. Впервые на основе интеграции геномных, транскриптомных и эпигеномных данных показана прямая связь между ОНВ, специфичными для агрессивной линии, и различиями в экспрессии генов.

Теоретическая и практическая значимость

В работе получены фундаментальные данные о генах и сигнальных путях, потенциально участвующих в регуляции агрессивного поведения на модели лис *Vulpes vulpes*. Набор ОНВ,

специфичных для агрессивной линии представляет собой потенциальные мишени для дальнейшего исследования в популяциях звероферм. Приоритетные кандидатные гены определяют конкретные направления дальнейших экспериментальных исследований, в частности, проверку ассоциации вариантов с поведенческими оценками и исследования белков.

Методология и методы исследования

Согласно поставленным задачам выбраны современные методологические подходы для их решения. В качестве объектов исследования использовались: фронтальная кора и миндалевидное тело лис *Vulpes vulpes*, принадлежащих к 3 разным поведенческим группам агрессивным, ручным и неселектированным по отношению к человеку, а также ранее полученные в лаборатории геномные данные норок *Neogale vison*, также принадлежащих к 3 аналогичным поведенческим группам.

Работа включала четыре основных этапа.

На первом этапе выделяли ДНК из фронтальной коры головного мозга лисиц, готовили геномные библиотеки, по результатам глубокого секвенирования которых получали полногеномные данные. Эти данные, дополненные полученными ранее, использовали для анализа популяционной структуры, поиска ОНВ и выявления геномных окон, специфичных для агрессивной линии. Также проводили аннотацию миссенс-вариантов, встречающихся с повышенной частотой в группе агрессивных животных.

На втором этапе выделяли тотальную РНК из фронтальной коры головного мозга и миндалевидного тела, готовили и секвенировали транскриптомные библиотеки, таким образом получали транскриптомные данные. Анализ дифференциальной экспрессии проводили отдельно для каждой ткани, определяли гены с дифференциальным уровнем экспрессии, отличающим агрессивных животных от ручных и неселектированных Определяли модули коэкспрессии генов и корреляции модулей с фенотипом.

На третьем этапе выделяли клеточные ядра из тканей фронтальной коры головного мозга с дальнейшей иммунопреципитацией с антителами к гистоновым меткам H3K4me3 и H3K27ac, выделяли ДНК, готовили таргетные библиотеки с последующим секвенированием. Проводили поиск локусов, обогащенных прочтениями по гистоновым меткам, анализировали различия между различными линиями лисиц и определяли локусы, специфичные для агрессивной линии.

Четвертым этапом проводили интегративный анализ. Обогащение генами с дифференциальной экспрессией среди генов, в которых расположены ОНВ, оценивали с помощью критерия Фишера. Для мегарегионов ОНВ оценивали обогащение генами с дифференциальной экспрессией с помощью перестановочного анализа.

Положения, выносимые на защиту:

1. В результате анализа данных геномного секвенирования 46 лисиц трех линий (агрессивная, ручная, неселектированная), выявлены однонуклеотидные варианты, специфичные для животных агрессивной линии. Большинство вариантов локализовано в некодирующих областях, что указывает на преимущественно регуляторный характер изменений.
2. Во фронтальной коре и миндалевидном теле показаны различия в уровне экспрессии генов между агрессивной, ручной и неселектированной линиями; выявлены общие и тканеспецифичные гены, а также модули коэкспрессии, связанные с исследуемым фенотипом.
3. Впервые охарактеризованы маркеры гистоновых меток H3K4me3 и H3K27ac во фронтальной коре лисиц и выявлены межлинейные различия в распределении активного хроматина в генах, связанных с миелинизацией, нейрогенезом, синаптической и дофаминергической сигнализацией.
4. Совместный анализ геномных, транскриптомных и эпигенетических данных указывает на модель полигенной регуляции поведения с существенным вкладом некодирующих вариантов и изменений состояния хроматина.

Степень достоверности и апробация результатов

Степень достоверности полученных результатов является высокой и подтверждается достаточным количеством полученных данных, сопоставлением с опубликованными данными, использованием общепринятых современных экспериментальных методов, а также корректностью применяемых статистических способов анализа результатов.

Апробация результатов была осуществлена на научных мероприятия:

1) Конференции BGRS/SB-2024 с устным докладом по теме “Регуляторная геномика агрессивного поведения на неканонической модели лисы *Vulpes vulpes*” (Новосибирск, 2024)
2) Конференция “Генетика 2025” с устным докладом по теме “Регуляторная геномика агрессивного поведения на лисице *Vulpes vulpes*” (Москва, 2025)
3) Конференция “Генетические процессы в популяциях” с устным докладом по теме “Профиль активного хроматина клеток фронтальной коры мозга модели агрессивного поведения лисы *Vulpes vulpes*” (Москва, 2026)

Глава 1. Обзор литературы

1. Генетическая архитектура агрессивного поведения

Агрессивное поведение – сложный количественный фенотип, вариабельность которого определяется генетическими и средовыми факторами. Метаанализы близнецовых и семейных исследований оценивают его наследуемость примерно в 40–60%; величина зависит от возраста, формы поведения и способа фенотипирования [1–4]. Среди форм агрессии выделяют проактивную (спланированную) и реактивную (следствие импульса, состояния аффекта) [5]. Реактивная и проактивная агрессия имеют как общие, так и специфические генетические компоненты [6, 7], а генетическое влияние на агрессию лишь частично перекрывается с влиянием на другие проявления антисоциального поведения [8, 9]. Поэтому результаты необходимо интерпретировать с учетом операционального определения фенотипа; наследуемость не означает индивидуальной предопределенности поведения.

1.1. Кандидатные гены и полигенная архитектура

Ранние исследования были сосредоточены на генах серотонинергической, дофаминергической и нейроэндокринной систем. Среди наиболее изученных кандидатов можно выделить такие гены, как *MAOA*, кодирующий фермент катаболизма моноаминов; *SLC6A4*, кодирующий переносчик серотонина; *DRD4*, кодирующий дофаминовый рецептор D4; и *COMT*, участвующий в метаболизме катехоламинов. Однако результаты ассоциаций распространенных генетических вариантов в этих генах, включая 5-HTTLPR и MAOA-uVNTR, неоднородны и не образуют устойчивого набора воспроизводимых эффектов [10–13]. Систематический обзор исследований детской агрессии также отметил неоднородность выборок, фенотипов и результатов анализа генетип-средовых взаимодействий [13]. Данные о редких генетических нарушениях и экспериментальных моделях подтверждают, что отдельные гены могут влиять на проявление агрессии, но эти эффекты нельзя напрямую переносить на распространенную вариабельность поведения [14, 15].

Полногеномные ассоциативные исследования (GWAS) позволили отказаться от предварительного выбора генов-кандидатов, однако ранние исследования агрессии не выявили устойчивого набора общегеномно значимых ассоциаций [16–18]. В крупном метаанализе агрессии в детском возрасте, включавшем 87 485 детей и 328 935 наблюдений из 29 когорт, ни

один вариант не достиг общегеномного уровня значимости; наследуемость, оцениваемая по распространенным однонуклеотидным полиморфизмам, составила 3,31%, а полигенные показатели объясняли менее 0,5% вариации в независимых выборках [19]. В метаанализе широкого антисоциального поведения были выявлены ассоциации с интронными вариантами в гене *FOXP2*, но изучавшийся фенотип не был специфичен для агрессии [20].

Интегративный анализ генетических данных человека и грызунов выявил повторяющееся обогащение в сигнальном пути с участием G-белоксопряженных рецепторов; серотонинергической, дофаминергической, глутаматергической и ГАМК-ергической передачи, а также ERK/MAPK- и Rho-GTPase-каскадов, CREB-сигнализации, Reelin-пути и аксонального наведения [21]. Эти пути объединяют процессы нейромедиаторной передачи, внутриклеточного ответа и образования нейронных связей. Однако их обогащение в разных наборах генов не доказывает существования единого универсального механизма агрессии. В целом высокая наследуемость признака не сопровождается выявлением небольшого числа распространенных вариантов с крупными эффектами, что поддерживает полигенную модель и необходимость изучать функциональные последствия генетической изменчивости.

1.2. Транскриптомная архитектура агрессивного поведения

Работы с использованием РНК-секвенирования (RNA-seq) позволяют исследовать экспрессию генов и, при соответствующем дизайне, альтернативный сплайсинг. При агрессивном поведении выявляются изменения широкого круга функциональных категорий; они могут быть регионально-специфичными и формировать скоординированные транскрипционные сети [22–26].

1.2.1. Транскриптомные изменения и регуляторные сети

У человека прямые исследования транскриптома мозга немногочисленны, а многие работы посвящены смежным фенотипам. Так, в орбитофронтальной коре лиц с антисоциальным расстройством личности и/или расстройством поведения Piras и соавт. выявили 328 дифференциально экспрессируемых генов и изменения коэкспрессионных модулей, включая транскрипты возбуждающих нейронов и астроцитов и гены синаптической и

глутаматергической регуляции [22]. Исследование LINC01268 связало экспрессию этой некодирующей РНК в мозге с насильственным способом суицида, а генетический вариант, ассоциированный с ее уровнем, – с показателями агрессии у живых участников [23]. Изменения экспрессии генов иммунной регуляции обнаруживались и в периферических мононуклеарных клетках крови при перемежающемся взрывном расстройстве поведения, характеризующимся вспышками гнева и агрессии, но периферические профили нельзя считать прямым отражением транскриптома мозга [24].

В экспериментальных животных моделях можно лучше разделять устойчивые межлинейные различия и ответ на социальный опыт. При сравнении селектированных линий мышей (*Mus musculus*) и рыбок данио (*Danio rerio*) были выявлены изменения в генах *Fos*, *Dusp1*, *Hdac4*, *Bdnf* и *Nr4a1*, а также в генной регуляторной сети с участием белка Fos и в сигнальном пути MAPK [25]. *Fos* и *Nr4a1* относятся к генам раннего ответа, *Dusp1* регулирует MAPK-каскад, *Hdac4* связывает транскрипционную регуляцию с состоянием хроматина, а *Bdnf* участвует в нейропластичности. У рыбок данио, селектированных по уровню агрессии в зеркальном тесте, RNA-seq выявил многочисленные дифференциально экпрессируемые транскрипты и коэкспрессионные модули, связанные с группой и продолжительностью поведения [26]. Другие исследования показывают влияние на транскриптом таких факторов, как социальный опыт, стресс, пол и статус животных [27–29].

Молекулярные профили могут отражать и состояние нервных сетей после социального взаимодействия. У мышей с разным опытом социальных побед и поражений изменения экспрессии выявлялись в вентральной тегментальной области среднего мозга [27]. В модели перипубертатного стресса усиление агрессии во взрослом возрасте было связано с изменениями транскриптома гипоталамуса и префронтальной коры; функциональные эксперименты показали участие транстиретина [28]. В базолатеральной миндалине хомяков после социального конфликта профиль экспрессии зависел от пола и социального статуса [29]. Эти результаты показывают, что транскриптом агрессивного животного может отражать как устойчивые особенности, так и предшествующий опыт.

Результаты сетевого анализа дополняют данные по дифференциальной экспрессии: у селектированных рыбок данио метод WGCNA позволил выделить 12 модулей, связанных с экспериментальной группой и/или длительностью агрессивного взаимодействия [26]. В других моделях совместные изменения экспрессии выявляли сети, связанные с Fos и MAPK-сигнальными каскадами [25, 30].

1.3. Животные модели генетики поведения

Животные модели позволяют сочетать направленную селекцию, анализ локусов количественных признаков (QTL-картирование, в частности, экспрессии - eQTL-картирование), GWAS и функциональные генетические вмешательства. Эти подходы обеспечивают разные уровни доказательности: картирование выявляет связанные с фенотипом регионы, транскриптомика – молекулярные корреляты, а функциональные эксперименты позволяют проверять роль отдельных генов и путей [15, 21, 31–42].

1.3.1. Грызуны и мухи Drosophila melanogaster

У мышей QTL-картирование выявляло локусы агрессии на разных хромосомах; результаты зависели от генетической популяции и поведенческого показателя, а в F2-скрещивании не был обнаружен единичный локус крупного эффекта. Панель BXD выявила локус, влияющий на межсамцовую агрессию [31–33]. Было показано, что дефицит Maoa сопровождался изменениями агрессии и обмена серотонина и норадреналина, что подтверждает функциональное значение этого гена [15].

На крысах *Rattus norvegicus*, селектированных на ручное и агрессивное поведение, методы QTL-картирование и RNA-seq позволили выявить совпадающие области регуляции экспрессии и поведения; к кандидатам относили гены *Gltscr2*, *Lgi4*, *Zfp40* и *Slc17a7* [34, 43]. Это иллюстрирует переход от локализации поведенческого QTL к поиску генов, экспрессия которых зависит от генотипа.

У мушек Drosophila искусственная селекция на агрессивное поведение сопровождалась наследуемостью около 0,10, изменением 1539 транскрипционных зондов и выявлением 15 генетических кандидатов при исследовании мутантов [35]. GWAS естественной вариабельности с использованием DGRP и популяции расширенного скрещивания выявил множество кандидатов и эпистатические взаимодействия [36]. На функциональном уровне дефицит октопамина почти устранял характерные агрессивные выпады, а восстановление его синтеза частично возвращало поведение [37]. Манипуляции отдельными дофаминергическими нейронами кластеров T1 и PPM3 также изменяли частоту атак; их проекции связаны с рецепторами DopR и DD2R в центральном комплексе мозга [44]. Это показывает, что моноаминергические системы могут модулировать агрессию через отдельные нейронные контуры, причем необязательно линейно.

1.3.2. Рыбы, птицы и домашние виды

У колюшки Gasterosteus aculeatus сотни дифференциально экспрессируемых генов при сравнении агрессивных и неагрессивных рыб выявлялись в стволе мозга, тогда как в других отделах их было меньше [45]; у рыбок данио были определены коэкспрессионные модули генов, связанные с зеркальной агрессией [26]. У цихлид (*Mchenga conophoros* и *Petrotilapia chitimba*) различия в агрессивности сопровождались изменениями активности нейронов и экспрессии генов, относящихся к нейропептидной регуляции, синтезу стероидных гормонов и нейромедиаторным процессам [39]. У белогорлого воробья (*Zonotrichia albicollis*) был выявлен суперген, связанный с альтернативными поведенческими фенотипами; снижение экспрессии гена *ESR1* уменьшало различия территориальной агрессии между морфами, что поддерживает функциональное участие эстроген-рецепторного пути в этой конкретной модели [40]. На разных породах собак GWAS выявлял ассоциации страха и агрессии между породами [41].

1.3.3. Общие закономерности и ограничения моделей

Животные модели взаимодополняют друг друга, но их результаты нельзя объединять без учета фенотипа, пола, социального контекста и теста. Территориальная агрессия, оборонительная реакция на человека и атака отражения представляют разные операциональные признаки. Сходство молекулярных категорий у разных видов указывает на потенциально общие процессы, но не доказывает идентичность механизмов [21, 42]. В связи с этим, представляется важным использовать для изучения агрессивного поведения модели длительной селекции, в которых поведение можно сопоставлять с конкретными геномными и молекулярными признаками.

1.4. Одомашнивание и поведенческая генетика пушных зверей: норка и серебристо-черная лисица

Одомашнивание – длительный эволюционный процесс, который у разных видов проходил по различным сценариям и включал освоение антропогенной среды, неосознанный

отбор и последующее целенаправленное разведение. На ранних этапах преимущество могли получать животные, лучше переносящие постоянный контакт с человеком и условия поселений; позднее люди целенаправленно отбирали особей по продуктивным и поведенческим признакам [46–48]. Снижение страха и повышение стрессоустойчивости рассматриваются как важные компоненты поведенческих изменений при доместикации, однако последовательность и сопряженность этих изменений различаются между видами [49].

1.4.1. История поведенческой селекции и эксперимент Д.К. Беляева

Эксперимент Д.К. Беляева с серебристо-черной лисицей начался в 1959 г. в Институте цитологии и генетики СО РАН. В ручную линию отбирали животных с ослабленной боязнью человека и склонностью к контакту; позднее была создана линия с выраженной агрессивно-оборонительной реакцией, а фермерская популяция без поведенческой селекции сохранялась как контроль [50, 51]. Этот эксперимент стал моделью последствий направленного отбора по реакции на человека, а не воспроизведением всей истории одомашнивания собак или других видов.

Исходные животные происходили из фермерских популяций, а не из недавно выловленной дикой популяции. Исторические сведения и анализ митохондриальной ДНК указывают преимущественно на восточно-канадское происхождение основателей [52]. Следовательно, до начала строгой селекции уже существовал период разведения в неволе, в ходе которого могли действовать условия фермерского содержания и неосознанный отбор. Беляев описывал направленный отбор по поведению как «дестабилизирующий», предполагая, что он может затрагивать регуляцию развития и приводить к сопряженным морфологическим и физиологическим изменениям [53].

Гипотеза единого «синдрома одомашнивания» остается предметом дискуссии: Lord и соавт. показали, что история происхождения фермерских лисиц и данные по другим одомашненным видам не позволяют считать весь набор предполагаемых признаков универсальным следствием отбора на ручное поведение [54]. Поэтому в настоящем обзоре селектированные линии рассматриваются как модель наследуемых различий в реакции на человека, а не как универсальный эквивалент доместикации. Селекция по реакции на человека сопровождалась и некоторыми изменениями других форм социального поведения. Например, ручные лисята успешнее использовали отдельные человеческие коммуникативные жесты, хотя этот навык не был непосредственной мишенью отбора [55].

1.4.2. Американская норка: селекция темперамента и генетика агрессии

Американская норка (*Neogale vison*) – близкая по экспериментальной логике к лисице модель, в которой поведенческие признаки изучались в условиях звероводства. Hansen исследовал более 3000 животных в трех линиях, селектированных по исследовательскому, боязливому и агрессивному темпераменту. За шесть поколений различия между линиями усилились; около 90% животных линии, отбиравшейся на боязливость, устойчиво демонстрировали страх при контакте с человеком [56]. После десяти поколений отбора по уверенной или боязливой реакции потомство различалось не только в тесте контакта с человеком, но и при предъявлении нового объекта, незнакомой норки и новой пищи; уверенные животные быстрее вступали в неагрессивный контакт с незнакомыми особями [57]. Это указывает на генерализацию страха и исследовательского поведения.

Количественные генетические исследования подтвердили наследуемость поведенческих признаков у норок. Для исследовательского/боязливого поведения в выборке из 26 371 животного оценка наследуемости составила 0,19 ± 0,03; генетическая корреляция между полами была высокой, а значимых генетических корреляций с качеством шкурки и плодовитостью выявлено не было [58]. Для стереотипного поведения оценка наследуемости составляла около 0,3, и его частоту удавалось изменять селекцией [59].

Для группового содержания особенно важны были непрямые генетические эффекты: количество следов укусов на шкурке зависит не только от генотипа пострадавшей особи, но и от поведения ее соседей. Модель, учитывающая прямые и непрямые эффекты, описывала данные лучше модели только с прямыми эффектами; совокупная наследуемая вариация для общего числа следов укусов достигала $T^2 = 0{,}61$ при классической $h^2 = 0{,}10$ [60]. Это указывает на взаимность агрессивных взаимодействий и необходимость учитывать социальное окружение при селекции. Другие исследования выявляли сопряженные изменения фолликулогенеза, сроков полового созревания и плодовитости у линий, селектированных на ручное или агрессивное поведение [61].

В фермерских популяциях описана связь некоторых мутаций окраски с частотой и выраженностью защитной реакции на человека: у норок окраса sapphire доля животных с признаками «доместикационного» поведения была выше [62]. Эти наблюдения позволяют допустить плейотропные эффекты отбора на поведение.

1.4.3. Серебристо-черная лисица: поведенческая и генетическая архитектура

Селектируемый фенотип лисиц отражает прежде всего реакцию на приближение и контакт с человеком, включая страх, избегание и оборонительную агрессию; он не эквивалентен внутривидовой агрессии и не охватывает все формы агрессивного поведения [51, 63–65]. Для количественного описания использовали систему из 50 бинарных наблюдений. Анализ главных компонент выявил градиент от ручного к агрессивно-оборонительному поведению и отдельную компоненту, связанную с активными и пассивными реакциями [63, 64]. QTL-картирование выявило поведенческий локус на хромосоме VVU12; эффекты зависели от генетического фона [64, 65].

В исследовании с помощью метода генотипирования секвенированием (GBS) были проанализированы 48 294 ОНВ; 68 вариантов различались между линиями, а десять из 11 дифференцированных ОНВ на VVU14 располагались в области около 1,3 Мб, перекрывающей ранее описанный локус QTL. Ген *NR2F2* был выделен как позиционный кандидат, но причинная роль конкретного варианта или гена не была установлена [66]. Полногеномное секвенирование ручной, агрессивной и неселектированной популяций выявило 103 региона с сигналами отбора; 30 перекрывались с ранее описанными поведенческими QTL. Среди позиционных кандидатов был также выделен ген *SORCS1*, кодирующий белок, участвующий в транспорте AMPA-рецепторов и нейрексинов, что указывает на возможную связь с синаптической пластичностью [67].

1.4.4. Транскриптомные, нейрохимические и эндокринные корреляты

Метод RNA-seq позволил выявить различия экспрессии в нескольких отделах мозга. В префронтальной коре и базальном переднем мозге ручных и агрессивных лисиц было обнаружено, соответственно, 146 и 33 дифференциально экспрессируемых гена; среди затронутых систем были серотонинергическая и глутаматергическая передача [68]. Наиболее значимые транскриптомные кандидаты включали гены *DKKL1* и *PCDHGA1*; изменения также затрагивали серотониновые рецепторы HTR5A-like, HTR3A и HTR7, компоненты cAMP/PKA и PI3K/AKT-сигнализации DUSP1 и AKT1, а также белки глутаматергического синапса GRIN2D,

ITPR3 и ADCY7. С помощью полногеномного исследования были выявлены варианты в генах *WNT3* и *WNT4* и других генах, связанных с развитием нервного гребня [67, 68].

В гипоталамусе было обнаружено 70 дифференциально экспрессируемых генов; семь из них – *DKKL1*, *FBLN7*, *NPL*, *PRIMPOL*, *PTGRN*, *SHCBP1L* и *SKIV2L* – демонстрировали однонаправленные изменения в трех отделах мозга. Модули коэкспрессии были обогащены путями GPCR, интерлейкиновой и цитокин-рецепторной сигнализации; у агрессивных лисиц также отмечалось обогащение mTOR-сигнализации, врожденного иммунного ответа и эндоцитоза [69]. Дополнительное исследование гиппокампа показало более низкий уровень стресс-индуцированного кортизола у ручных животных и более высокий – у агрессивных. У ручных лисиц повышалась экспрессия генов *CYP26B1* и *ADRA1A* в дорзальном гиппокампе и *NR3C2* в вентральном [70]. Эти результаты могут указывать на возможный вклад метаболизма ретиноевой кислоты, адренергической передачи и минералокортикоидной регуляции в различиях стресс-ответа.

Ранее нейрохимические исследования ручных лисиц описывали изменения серотонина и 5-HIAA, активности триптофангидроксилазы, МАО-А и плотности 5-HT1A-рецепторов в отдельных структурах мозга [71]. В исследовании гипоталамо-гипофизарно-надпочечниковой системы у ручных самцов после отлова и последующего контакта с людьми были снижены уровни кортизола, АСТН и экспрессии гена *POMC*; у агрессивных лисиц экспрессия гена *POMC* также была снижена, но уровни кортизола и АСТН были ближе к неселектированной линии [72]. Таким образом, изменения нейромедиаторной и эндокринной регуляции не формируют единого линейного градиента между группами.

1.4.5. Нейроморфологические изменения мозга лисиц

Дополнительно были изучены нейроморфологические особенности мозга исследуемых линий. Hecht и соавт. сравнили мозг 30 самцов – по 10 ручных, агрессивных и неселектированных лисиц – с использованием высокоразрешающей посмертной МРТ. Общий объем серого вещества был выше у обеих селектированных линий по сравнению с контрольной, тогда как различия общего объема мозга и белого вещества не достигали статистической значимости. Увеличенные области частично перекрывались и включали участки префронтальной, моторной и соматосенсорной коры, миндалину, гиппокамп и мозжечок. Анализ морфологической ковариации выявил сеть префронтальная кора–мозжечок, различавшую популяции вдоль поведенческой оси «ручная–агрессивная реакция», и сеть

префронтальная кора–гипоталамус, отличавшую обе селектированные линии от контрольной [73]. Результат показывает, что отбор по противоположным реакциям на человека может сопровождаться частично сходными изменениями морфологии мозга, а простая модель, согласно которой ручное поведение обязательно связано с уменьшением мозга, не подтверждается.

В отдельной работе Flattery и соавт. представили анатомический атлас мозга серебристой лисицы, объединив гистологические срезы и МРТ-изображения. Атлас был основан на животном из неселектированной фермерской линии и предоставляет анатомическую основу для маркировки структур и последующих сравнительных исследований [74].

1.5. Эпигенетическая регуляция агрессивного поведения у человека и животных

Эпигенетические механизмы – метилирование ДНК, модификации гистонов, перестройка хроматина и регуляция некодирующими РНК – изменяют активность генов без изменения последовательности ДНК. Они могут связывать генетическую предрасположенность с ранним стрессом и социальным опытом и различаться между тканями, отделами мозга и типами клеток [75–78]. Связь эпигенома с агрессией двунаправленна: эпигенетические изменения могут влиять на поведение, но также отражать его последствия.

1.5.1. Метилирование ДНК и ранний опыт

Метаанализ van Dongen и соавт., включавший 15 324 участника, выявил 13 сайтов метилирования в периферической крови 14 434 участников из 18 когорт, ассоциированных с агрессивным поведением. В пуповинной крови 2425 детей отдельные сайты метилирования не достигли общегеномной значимости, хотя для 83% сайтов направление ассоциации совпадало с результатами анализа крови взрослых. Часть сигналов была связана с курением, индексом массы тела, химическими воздействиями и генетическими вариантами, влияющими на метилирование [79]. Эти результаты не определяют единую эпигенетическую сигнатуру агрессии; анализ крови также не позволяет напрямую судить о состоянии хроматина в мозге. Соответственно, работы по взаимодействию генотипа и неблагоприятного детского опыта подчеркивают важность средового контекста [80].

Эксперименты на грызунах показывают, что ранний социальный опыт может длительно менять регуляцию генов стрессового ответа и нейропластичности. Материнское поведение у крыс связано с метилированием промотора гена *NR3C1* в гиппокампе потомства; ранние неблагоприятные воздействия также влияют на эпигенетическое состояние гена *BDNF* и чувствительность гиппокампа к среде [81–83]. Более прямые данные были получены на моделях стресса: перипубертатный стресс у крыс сопровождался усилением агрессии, повышением экспрессии гена *MAOA* и ацетилирования гистона H3 в промоторе гена *MAOA* в префронтальной коре, но аналогичные изменения не воспроизводились в миндалевидном теле [84]. У мышей были выявлены половые и региональные различия метилирования в гене *MAOA* и связывания с SIRT1 [85]; социальная изоляция сопровождалась гиперметилированием гена *PPARA*, снижением его экспрессии и изменениями нейровоспалительных процессов в модели агрессивного поведения [86].

1.5.2. Некодирующие РНК и природные модели

Длинная некодирующая РНК MAALIN связана с регуляцией экспрессии гена *MAOA*. В посмертных образцах мозга людей с импульсивно-агрессивными чертами повышение экспрессии MAALIN сопровождалось снижением экспрессии *MAOA* и изменениями метилирования и гистоновых модификаций. Эксперименты повышением экспрессии и выключением MAALIN в нейрональных клетках-предшественниках подтвердили обратную связь между уровнем этой РНК и экспрессией *MAOA*. Сверхэкспрессия MAALIN в гиппокампе мышей сокращала латентный период до первой атаки, но не увеличивала статистически значимо число атак [87]. Известно, что микроРНК также могут участвовать в регуляции агрессии: описаны взаимодействие miR-96 с регуляторной областью HTR1B и ассоциация варианта MIR124-1 с показателями агрессивности [88, 89].

В природных моделях показана связь метилирования регуляторной области гена VIP с альтернативными фенотипами у белогорлого воробья [90], а экспериментальная конкуренция у древесных ласточек (*Tachycineta bicolor*) сопровождалась изменениями в мозге птиц на транскриптомном и метиломном уровнях [91]. У мушек дрозофил ранее были исследованы белки MBD, участвующие в регуляции агрессивного поведения, а у медоносной пчелы (*Apis mellifera*) – родительские эффекты на транскрипцию и метилирование РНК [92, 93]. Таким образом, в целом, эпигенетические сигналы зависят от вида, фенотипа, пола, возраста, опыта и ткани.

1.5.3. Гистоновые модификации и регуляторные элементы: H3K4me3, H3K27ac и ChIP-seq

Посттрансляционные модификации гистонов влияют на организацию хроматина, связывание регуляторных белков и доступность ДНК. Их функция зависит от типа модификации, аминокислотного остатка и клеточного контекста [94, 95]. Для анализа активных регуляторных областей особенно важны гистоновые метки H3K4me3 и H3K27ac: первая преимущественно обогащена в промоторах около сайтов начала транскрипции, вторая – в активных промоторах и энхансерах [96–101].

Функциональное значение H3K4me3 и H3K27ac

Метка H3K4me3 характерна для промоторных областей активных генов, но не является прямой количественной мерой экспрессии: наличие метки не гарантирует высокий уровень транскрипции [102]. В нейрональном хроматине широкие домены H3K4me3 связаны с генами синаптической функции и ГАМК-ергической передачи [103]. Метка H3K27ac маркирует активные промоторы и энхансеры; ее уровень связан с активностью гистоновых ацетилтрансфераз, включая CBP/p300 [97, 101]. В сочетании с меткой H3K4me1 эта метка помогает отличать активные энхансеры от потенциально подготовленных к транскрипции участков [98–101]. При этом метка H3K27ac сама по себе не определяет тип регуляторного элемента: важны положение пика, дополнительные метки, доступность хроматина и функциональная проверка.

Гистоновые профили в мозге зависят от типа клеток и области. Girdhar и соавт. получили 157 карт меток H3K4me3 и H3K27ac для нейрональных, обедненных нейронами и цельных образцов двух областей коры; различия между нейрональными и ненейрональными фракциями были существенным источником различий [104]. Кроме того, нейрональная активность может сопровождаться изменениями транскрипции с энхансеров и уровня H3K27ac в регуляторных областях [105, 106]. Поэтому профили цельной ткани могут отражать как изменения хроматина внутри клеток, так и изменение доли клеточных типов.

Гистоновые модификации и агрессивное поведение

Прямые исследования гистоновых модификаций при агрессии немногочисленны. После перипубертатного стресса у крыс усиление агрессии и экспрессии гена *MAOA* сопровождалось повышенным ацетилированием гистона H3 в промоторе гена *MAOA*, однако авторы измеряли ацетилирование H3 в целом, а не специфически H3K27ac [84]. В ранее упомянутом исследовании некодирующей РНК MAALIN изменения хроматина в регуляторной области

включали снижение метки H3K27me3, тогда как различия H3K4me3 и энхансерных меток H3K4me1/H3K27ac не были значимыми [87]. В модели хронической социальной изоляции и острого стресса у мышей ингибиторы гистоновой деацетилазы HDAC I класса ослабляли агрессивные реакции и улучшали социальное взаимодействие [107]. У мышей с дефицитом белка Shank3 сочетанное ингибирование HDAC и LSD1 уменьшало повышенную агрессию и сопровождалось изменениями NMDA-рецепторной передачи [108]. Однако, стоит учитывать, что вышеописанные исследования затрагивают множество генов и не доказывают причинную роль отдельной гистоновой метки, а фенотип агрессивного поведения у мух относится к половому поведению и ухаживанию, а не к агрессии в строгом смысле [109].

Данные, полученные на модельных животных, свидетельствуют о сложной эпигенетической природе агрессивного поведения: сходные фенотипы могут обуславливаться изменениями разных генов и механизмов. Основные гены и/или белки, а также эпигенетические механизмы, участие которых в регуляции агрессии подтверждено на нескольких животных моделях, представлены на рисунке 1.1.

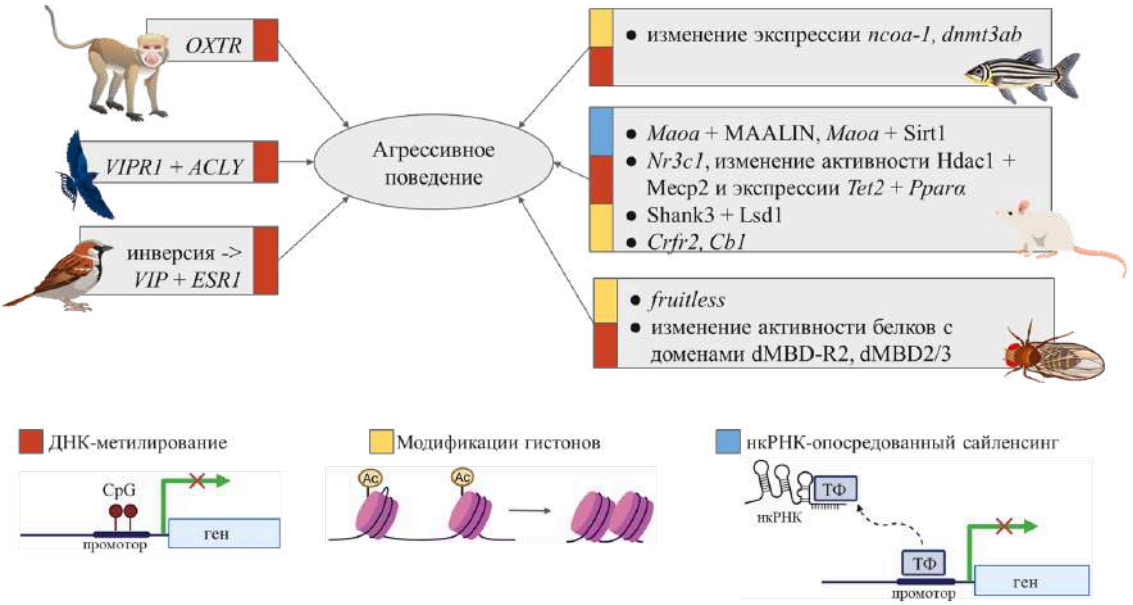


Рисунок 1.1. Животные модели для изучения эпигенетических механизмов агрессивного поведения и выявленные на них за последние несколько лет гены и/или белки, опосредованно или напрямую участвующие в эпигенетической регуляции. Цветом (красным, желтым и синим) обозначены ДНК-метилирование, модификации гистонов и нкРНК-опосредованный сайленсинг, соответственно.

1.6. Интеграция геномных, транскриптомных и эпигеномных данных в исследовании сложных поведенческих фенотипов

Агрессивное поведение имеет полигенную и многоуровневую природу: генетические варианты могут влиять на экспрессию и сплайсинг, а состояние регуляторных элементов – на активность генов и нейрональных сетей. Интеграция данных, полученных с помощью методов WGS, RNA-seq и эпигеномики позволяет связать генетические различия с молекулярными последствиями и приоритизировать гены и пути, потенциально участвующие в формировании фенотипа [110–114].

1.6.1. От генетических вариантов к регуляции экспрессии

Многие варианты сложных признаков расположены вне кодирующих областей и могут влиять на промоторы, энхансеры, связывание транскрипционных факторов, сплайсинг или стабильность РНК. eQTL-анализ позволяет выявить варианты, связанные с уровнем экспрессии, а sQTL и mQTL (локусы, генетические варианты в которых влияют на альтернативный сплайсинг РНК или ассоциированы с уровнем метилирования ДНК) – с альтернативным сплайсингом и метилированием. Атлас экспрессии GTEx, включающий 15 201 образцов RNA-seq из 49 тканей 838 доноров, показал выраженную тканевую специфичность генетических эффектов на экспрессию и сплайсинг [112]. Колокализация GWAS- и eQTL-сигналов помогает проверить, обусловлены ли оба сигнала одним и тем же вариантом [115]; полнотранскриптомный поиск ассоциаций (TWAS) позволяет выделить гены по предсказанному уровню экспрессии, но не доказывает, что именно их экспрессия является причиной признака [116].

1.6.2. Совместный анализ экспрессии и регуляторных элементов

Выполненное в рамках консорциума Roadmap Epigenomics исследование объединило данные 111 референсных эпигеномов, а база данных ENCODE расширила аннотацию функциональных элементов генома человека и мыши [110, 111]. Проекты PsychENCODE создали интегративные ресурсы человеческого мозга, объединяющие геномные, транскриптомные, эпигеномные и клеточные данные [113, 114]. Совместный анализ RNA-seq и ChIP-seq позволяет проверить, сопровождается ли изменение гистоновой метки в промоторной области изменением экспрессии соответствующего гена. Однако отсутствие полного совпадения

не обязательно противоречиво: экспрессия зависит от нескольких регуляторных элементов и может не меняться синхронно с гистоновой меткой.

Варианты, расположенные в областях с признаками активного хроматина, могут влиять на регуляторную функцию. Для дистальных энхансеров необходимо отдельно устанавливать ген-мишень: ближайший ген не всегда является правильным. Вычислительная модель Activity-by-Contact, предсказывающая, какие энхансеры регулируют конкретные гены в определенном типе клеток, объединяет активность элемента и его физический контакт с промотором; ее прогнозы проверялись функциональными CRISPRi-экспериментами [117]. На ее основе построены карты взаимодействий энхансеров и генов для 131 типа клеток и тканей человека [118].

1.6.3. Анализ сетей коэкспресии и интеграция данных лисицы

Интеграцию можно проводить на уровне генов, биологических процессов и коэкспрессионных сетей. Метод WGCNA позволяет выделить модули совместно экспрессируемых генов, связанные с фенотипом, выделяя центральный “хаб”-ген в модуле [119]. Мультиомиксный факторный анализ (MOFA) позволяет выделять общие и специфичные источники изменчивости в нескольких омиксных наборах [120]. Если профили разных типов получены не от одних и тех же животных, их сопоставляют на уровне групп, генов, локусов или функциональных категорий. Так, в исследовании на крысах объединение поведенческих QTL и eQTL помогло приоритизировать гены-кандидаты [43].

В настоящем исследовании было решено использовать метод WGS для характеристики генетических различий между ручной, агрессивной и неселектированной популяциями; метод RNA-seq для выявления изменений экспрессии во фронтальной коре и миндалевидном теле и метод ChIP-seq для оценки состояния активности хроматина во фронтальной коре. Фронтальная кора и миндалевидное тело были выбраны как ключевые узлы нейронной сети регуляции эмоций и агрессивного поведения, дисфункция связей между которыми лежит в основе реактивной агрессии [121, 122]. Соответственно, далее было решено провести прямое сопоставление трех уровней для фронтальной коры, а данные миндалевидного тела интерпретировать в контексте WGS и транскриптома, без выводов о гистоновых метках в этой ткани. Сравнение трех популяций помогает отделять различия между крайними фенотипами от особенностей каждой селекционной линии, но генетическую структуру и демографическую историю необходимо учитывать [67, 123].

Глава 2. Материалы и методы

Методологическая база настоящего исследования включает в себя комплексный анализ геномных, транскриптомных и ChIP-seq данных. Описание используемых подходов в данном разделе структурировано по принципу раздельного рассмотрения каждого молекулярно-генетического уровня. Такой порядок изложения обусловлен существенными различиями в протоколах пробоподготовки образцов для секвенирования и алгоритмах биоинформатического анализа. Индивидуальное описание каждого метода необходимо для детального изложения специфических этапов контроля качества и первичной обработки данных.

2.1 Полногеномные данные

2.1.1. Коллекция биологического материала

Лисы

В работе были использованы полногеномные данные от 37 самцов и 9 самок рыжей лисицы *Vulpes vulpes*, которые содержались на экспериментальной звероферме ИЦиГ СО РАН и подверглись отбору на агрессивное и ручное поведение, все исследуемые особи были охарактеризованы по оборонительной реакции на человека. Было использовано 33 образца (10 агрессивных, 12 дружелюбных и 11 неселектированных особей), полученные ранее в лаборатории, а также 13 образцов фронтальной коры, из которых была выделена ДНК, приготовлены полногеномные библиотеки и получены полногеномные данные (5 агрессивных, 3 дружелюбных и 5 неселектированных особей). Фенотипы всех 46 образцов представлены в Приложении А.

2.1.1.1. Выделение ДНК, пробоподготовка геномных библиотек, секвенирование

Выделение ДНК проводили с использованием стандартного протокола со спин-колонками производителя QIAGEN (QIAamp DNA Mini). Для приготовления библиотек из образцов ДНК использовали набор реагентов NEBNext Ultra II DNA Library Prep Kit for Illumina

(New England BioLabs) и протокол, рекомендованный фирмой-производителем набора. Для этого по 100 нг ДНК каждого образца фрагментировали с помощью ультразвукового соникатора и далее использовали в реакциях достраивания концов, аденилирования 3'-концов и лигирования NEBNext адаптеров. После необходимых промежуточных этапов, а также готовые фрагменты библиотек очищали с помощью магнитных частиц NEBNext Sample Purification Beads (New England BioLabs), после чего библиотеку амплифицировали с использованием двойных индексов NEBNext (New England BioLabs). Оценку качества приготовленных библиотек осуществляли с помощью наборов DNA High Sensitivity Chip (Agilent) и High Sensitivity D5000 (Agilent) на приборах Agilent 2100 Bioanalyzer (Agilent) или TapeStation 4150 (Agilent), соответственно. Молярность библиотек оценивали методом qPCR с использованием набора KAPA Library Quantification Kit (KAPA Biosystems). Секвенирование полученных библиотек проводили на приборе Illumina NovaSeq в режиме парно-концевых прочтений (151+151 п.н.).

Норки

В работе были использованы биологические образцы, полученные от самцов американской норки *Neogale vison* стандартной темно-коричневой окраски, которые содержались на экспериментальной звероферме ИЦиГ СО РАН и подверглись отбору на агрессивное поведение, все исследуемые особи были охарактеризованы по оборонительной реакции на человека.

2.1.2. Анализ данных полногеномного секвенирования

С помощью команды bwa mem из программного пакета BWA полученные после глубокого секвенирования прочтения были картированы на референсную сборку лисы VulVul3 [124]. Маркирование ПЦР-дупликатов в картированных данных секвенирования проводили с использованием команды MarkDuplicates из пакета программ Picard [125].

2.1.2.1. Выявление однонуклеотидных вариантов и генотипирование выборки

Однонуклеотидные варианты (ОНВ) выявляли по выровненным BAM-файлам с помощью GATK HaplotypeCaller v4.5.0 в режиме оценки достоверности референсного генотипа

(*--emit-ref-confidence GVCF*) [126, 127]. В этом режиме для каждого образца сохраняются блоки gVCF, содержащие информацию о достоверности соответствия генотипа референсному геному, что позволяет на этапе определения генотипов всех образцов из выборки отличать истинные гомозиготные референсные генотипы от отсутствия прочтений. Выявление вариантов выполняли отдельно для каждой комбинации «образец – хромосома». Минимальные значения качества оснований и качества картирования составляли 20. Данный порог качества прочтений соответствует значению по умолчанию в GATK, а порог качества картирования позволяет исключить прочтения, для которых положение в геноме определено неоднозначно. Расчет правдоподобия с помощью PairHMM, являющийся основным ресурсоемким этапом HaplotypeCaller, выполняли в нативном многопоточном режиме. Затем gVCF-файлы, полученные для отдельных хромосом, объединяли в один gVCF-файл для каждого образца и индексировали.

Генотипирование всей когорты проводили отдельно для каждой хромосомы в три этапа. Сначала gVCF-файлы всех образцов импортировали в рабочую область GenomicsDB с помощью GenomicsDBImport, обрабатывая по 25 образцов за один запуск. Такой подход позволял ограничить пиковое потребление памяти. Затем совместные генотипы определяли с помощью GenotypeGVCFs [128] с параметрами по умолчанию; на этом этапе фильтрацию по качеству на уровне сайтов и генотипов не проводили, перенося все процедуры фильтрации на следующий этап. Наконец, VCF-файлы с результатами совместного генотипирования для отдельных хромосом объединяли в единый VCF-файл для всей выборки.

2.1.2.2. Фильтрация вариантов и фазирование

В отсутствие валидированного набора истинных вариантов для генома лисицы качество выявленных вариантов оценивали с помощью жесткой фильтрации, процедуру перекалибровки качества вариантов не использовали. ОНВ извлекали с помощью функции *SelectVariants* и фильтровали с помощью функции *VariantFiltration*, используя рекомендуемые GATK пороговые значения для немодельных организмов: QD < 2.0, FS > 60.0, MQ < 40.0, MQRankSum < −12.5 и ReadPosRankSum < −8.0. Сохраняли только сайты, прошедшие все фильтры; после фильтрации для оставшегося набора повторно рассчитывали количество и частоты аллелей в выборке .

Затем отфильтрованный набор ОНВ ограничивали строго биаллельными ОНВ с помощью bcftools v1.20 [129, 130] (*view -m2 -M2 -v SNP*). Фазирование аллелей образцов выполняли отдельно для каждой хромосомы с помощью SHAPEIT2 r904, используя окно

фазирования размером 1 Mb, установленное по умолчанию [131]. Поскольку валидированная генетическая карта лисицы отсутствовала, задавали равномерную частоту рекомбинации 1 cM/Mb в требуемом SHAPEIT2 трехколоночном формате (позиция в п.н., частота рекомбинации в cM/Мб, кумулятивная генетическая позиция в cM). Фазированные гаплотипы преобразовывали обратно в формат VCF (*shapeit -convert*), сжимали, индексировали и объединяли по хромосомам с помощью bcftools concat.

Затем к фазированному VCF применяли два фильтра на уровне всей выборки: удаляли сайты, в которых доля отсутствующих генотипов превышала 20%, а также сайты с частотой минорного аллеля ниже 0.05 (bcftools filter). В результате получали готовый к анализу набор вариантов, включавший фазированные биаллельные ОНВ со статусом PASS, MAF ≥ 0.05 и долей отсутствующих генотипов на уровне сайта ≤ 0.2.

2.1.2.3. Анализ популяционной структуры

Импорт данных и прунинг (прореживание) по неравновесию по сцеплению

Фазированный VCF использовали в качестве входного файла в PLINK2 [132]. Поскольку SHAPEIT2 оставляет столбец идентификаторов вариантов пустым, при импорте каждому варианту присваивали уникальный идентификатор в формате «хромосома:позиция» (*--set-all-var-ids '@:#'*). Допускались названия скаффолдов, не входящие в стандартный набор хромосом PLINK (*--allow-extra-chr*), а неполные генотипы (half-calls) преобразовывали в отсутствующие (*--vcf-half-call missing*).

Для популяционного анализа с использованием методов PCA и ADMIXTURE рекомендуют использовать несцепленные хромосомные локусы, так как участки, находящиеся в неравновесии по сцеплению (Linkage Desquilibrium, LD) влиять на результат анализа, проводили удаление (прунинг) сцепленных вариантов с помощью PLINK2 *--indep-pairwise [133, 134]*. Использовали окно размером 50 Кб, шаг в один вариант (при задании окна в килобазах шаг должен быть равен единице) и порог $r^2 = 0.1$. Поскольку PLINK2 не обеспечивает надежную оценку LD для выборок, содержащих менее 50 основателей, для использованной в настоящем исследовании полногеномной выборки из 46 образцов требовалось принудительное переопределение *--bad-ld*. Оставшиеся варианты (prune.in) извлекали в набор данных для последующего анализа.

Анализ главных компонент

PCA выполняли на наборе данных после удаления сцепленных вариантов с помощью PLINK2 --*pca 10*, реализующего точное сингулярное разложение. PLINK2 не оценивает частоты аллелей для выборок менее чем из 50 образцов; поэтому частоты аллелей рассчитывали отдельно (--*freq*) и передавали в процедуру PCA (--*read-freq*). Сохраняли десять ведущих собственных векторов (коэффициентов проекций образцов) и соответствующие собственные значения. Долю объясненной дисперсии для каждой компоненты рассчитывали как отношение ее собственного значения к сумме всех собственных значений.

Анализ генетического происхождения с помощью ADMIXTURE

Доли предковых популяций в геномах оценивали с помощью ADMIXTURE v1.3.0 на наборе данных после прореживания по неравновесию по сцеплению [135]. Для работы ADMIXTURE требуется бинарный формат PLINK 1 (.bed) с целочисленными кодами хромосом. Поэтому идентификаторы скаффолдов NC_* перекодировали в последовательные целые числа в порядке следования в файле, а варианты сортировали по хромосомным позициям перед преобразованием в формат .bed. Параметр переопределения количества хромосом --chr-set задавали равным 16, чтобы обеспечить распознавание всех перекодированных хромосом.

Для каждого значения K от 2 до 5 ADMIXTURE запускали десять раз с фиксированными, различными случайными начальными значениями генератора (seed = 100·K + replicate) и пятикратной перекрестной проверкой (--*cv*). Повторные запуски позволяли снизить риск сходимости к локальным оптимумам. Ошибку перекрестной проверки для каждого запуска извлекали из журналов выполнения; оптимальное значение K выбирали по минимальной средней ошибке перекрестной проверки, а для визуализации сохраняли повторный запуск с наименьшей ошибкой CV при выбранном K.

Визуализация

Диаграммы рассеяния PCA (PC1 в зависимости от PC2, с цветовым обозначением по линии селекции и указанием доли объясненной дисперсии для каждой оси), а также столбчатые диаграммы долей предковых компонент ADMIXTURE (матрицы Q, упорядоченные по линиям селекции, с отдельной панелью для каждого K) строили в R 4.0.0 с использованием пакетов ggplot2 и ggrepel [136].

2.1.2.4. Выявление ОНВ, специфичных для агрессивной линии, окна ОНВ и гаплотипный анализ

Выявление ОНВ, специфичных для агрессивной линии (фильтр частот 80/10/10)

Варианты, специфичные для агрессивной линии, выявляли с использованием фильтра частоты наличия альтернативного аллеля на основе набора биаллельных ОНВ, прошедших жесткую фильтрацию (результат до фазирования). Для каждого ОНВ в каждой линии рассчитывали частоту присутствия альтернативного аллеля как долю генотипированных животных, несущих хотя бы одну его копию. ОНВ относили к специфичным для агрессивной линии, если альтернативный аллель присутствовал не менее чем у 80% животных агрессивной линии и не более чем у 10% животных ручной неселектированной контрольной линий (пороговые значения 0.8/0.1/0.1). Данный фильтр намеренно ориентирован на направление фенотипического различия и не опирается на предположения о распределении частот аллелей.

Аннотацию функциональных последствий вариантов в кодирующих областях проводили с помощью утилиты bcftools csq относительно геномной референсной последовательности GCF_048418805.1. Этот инструмент был выбран, поскольку он учитывает гаплотипный контекст при определении последствий вариантов для отдельных транскриптов. Каждый ОНВ дополнительно характеризовали по его положению относительно геномных функциональных областей на основании геномной аннотации. Выделяли следующие категории: CDS, экзон (не кодирующая последовательность), интрон, промотор (±3 кб от сайта начала транскрипции, TSS) и межгенная область. Для ОНВ, расположенных вне генов, указывали ближайший ген и расстояние до него, а также описание гена.

Визуализация плотности ОНВ по всему геному

Геномное распределение и кластеризацию ОНВ, специфичных для агрессивной линии, визуализировали с помощью программного пакета CMplot в среде R [137] в режиме отображения плотности ОНВ (plot.type = "d"). Использовали интервалы размером 1 Мб и цветовой градиент. Данная визуализация позволяла выявить области с высокой плотностью ОНВ, на основании которых определяли аналитические окна.

Формирование и аннотирование ОНВ-окон

Аналитические окна формировали путем разделения генома на неперекрывающиеся интервалы размером 1 Мб, соответствующие разбиению, использованному при визуализации плотности. Каждый интервал, содержащий хотя бы один ОНВ, специфичный для агрессивной линии, сохраняли в качестве отдельного окна; его границы уточняли по крайним позициям ОНВ внутри соответствующего интервала.

Каждое окно аннотировали следующими характеристиками: номер и название хромосомы согласно сборке VulVul3, длина хромосомы, количество и символы перекрывающихся генов, ближайший ген и расстояние до него для окон, не содержащих генов, а также плотность генов на уровне хромосомы и отдельного окна. Полученные окна далее

использовались в качестве входного файла для гаплотипного теста и функциональной аннотации.

Гаплотипный тест в пределах окон

Для каждого окна фазированные с помощью SHAPEIT2 генотипы в позициях ОНВ, входящих в это окно, объединяли в одну аллельную последовательность для каждой копии хромосомы каждого образца. Например, последовательность 01101 обозначала сочетание референсных и альтернативных аллелей в позициях ОНВ, расположенных в пределах окна. Цифры 0 и 1 соответствовали референсным и альтернативным аллелям VCF соответственно; при этом такое кодирование не отражало направленность аллеля относительно агрессивного фенотипа. Нефазированные гетерозиготные генотипы и отсутствующие генотипы трактовали как пропущенные данные для соответствующей копии хромосомы и исключали из расчета частот и статистических тестов. Количество копий хромосом, включенных в анализ, регистрировали отдельно для каждого окна.

Доминирующий гаплотип агрессивной линии для каждого окна определяли как наиболее часто встречающуюся аллельную последовательность среди копий хромосом животных агрессивной линии. В случае равенства частот выбирали последовательность, первую в лексикографическом порядке. Для каждого окна в каждой линии селекции рассчитывали долю валидных копий хромосом, несущих данный доминирующий гаплотип. Различия между агрессивной и ручной линиями проверяли с помощью точного теста Фишера, применяемого к копиям хромосом и основанного на сравнении носителей доминирующего гаплотипа с неносителями. Полученные p-значения корректировали по методу Бенджамини – Хохберга для контроля доли ложноположительных результатов (FDR) с учетом множественного тестирования по всем окнам.

В анализ включали также окна, содержащие только один ОНВ: в этом случае гаплотипный тест сводился к тестированию отдельного аллеля. При интерпретации результатов такие окна рассматривали отдельно от окон, содержащих несколько сайтов.

2.1.2.5. Анализ однонуклеотидных вариантов в геноме норки

Для 11 агрессивных норок были получены данные полногеномного секвенирования на платформе Illumina HiSeq 2000/2500, в качестве группы сравнения использовались геномы 8 норок, не подвергавшихся отбору по поведению, полученные нами ранее [138–140]. Полученные данные геномного секвенирования с помощью программы BWA mem [124, 141]

были картированы относительно референсного генома самки американской норки (GCF_020171115.1). Генетические варианты в области гена *MAOA* были выявлены с помощью функции HaplotypeCaller (*--sample-ploidy 1*) из пакета программ GATK v4 [142] и анализировались в программе PLINK [143] .

2.2. Транскриптомы

Для исследования были использованы образцы фронтальной коры головного мозга и миндалевидного тела самцов лисиц из разных поведенческих групп (дружелюбные, агрессивные, неселектированные), предоставленные зверофермой Института цитологии и генетики СО РАН в Новосибирске. Итого в анализе было использовано 45 образцов от 30 животных: 30 образцов фронтальной коры (10 образцов на поведенческую группу), 15 образцов миндалевидного тела (5 образцов на поведенческую группу). При этом, для 3-4 животных из каждой поведенческой группы в анализе использовались образцы из обоих отделов мозга.

2.2.1 Выделение тотальной РНК, приготовление и секвенирование транскриптомных библиотек

Методом экстракции РНК с помощью Тризола с последующей обработкой от ДНКаз и очисткой на колонках RNeasy были получены препараты РНК из двух отделов мозга. Для каждого препарата были получены профили РНК с каноническими субъединицами (18S и 28S) и оценен показатель RIN с помощью капиллярного фореза на приборе TapeStation 4150 (Agilent) с использованием набора RNA ScreenTape.

Для образцов с показателем RIN, большим или равным 7,5 было проведено обогащение мРНК и приготовление транскриптомных библиотек с помощью наборов NEBNext Poly(A) mRNA Magnetic Isolation Module (NEB #E7490) и NEBNext Ultra II Directional RNA Library Prep Kit for Illumina (NEB #E7760, #E7765).

Оценку качества приготовленных библиотек осуществляли с помощью набора D1000 (Agilent) на приборе TapeStation 4150 (Agilent). Молярность библиотек оценивали методом qPCR с использованием набора KAPA Library Quantification Kit (KAPA Biosystems).

Тестовое секвенирование полученных библиотек проводили на приборе MiSeq с последующей первичной биоинформатической обработкой для оценки качества библиотек (в

частности, доли прочтений, картируемых в пределах аннотированных генов). Глубокое секвенирование транскриптомов проводилось на приборе NovaSeq.

2.2.2 Анализ данных транскриптомного секвенирования

2.2.2.1. Первичная обработка транскриптомных данных

Прочтения для 30 библиотек из фронтального кортекса и 15 библиотек из миндалевидного тела картировали на референсную сборку генома лисы VulVul3 с помощью программы HISAT2 [144]. Провели подсчет прочтений в рамках координат экзонов аннотированных генов (с использованием базовой опции *-t exon*) и получили матрицы экспрессии с помощью команды *featurecounts* из пакета программ *Subread* [145]. Для получения данных о доле прочтений, картированных на кодирующие, интронные, межгенные и нетранслируемые (UTR) области, а также о покрытии генов, использовалась программа CollectRnaSeqMetrics из пакета Picard [125].

Анализ выполняли отдельно для каждого отдела мозга, на матрицах чисел прочтений по экзонам с помощью пакета программ Deseq2 на языке R [146]. Гены фильтровали по правилу наличия не менее 10 прочтений как минимум в 3 образца, далее отфильтрованные матрицы использовали для анализа дифференциальной экспрессии.

2.2.2.2. Анализ главных компонент и частичный дискриминантный анализ методом наименьших квадратов

Для оценки различий в профилях экспрессии генов между линиями лис проводили анализ главных компонент (PCA) с помощью пакета Deseq2 и частичный дискриминантный анализ методом наименьших квадратов (PLS-DA) с использованием пакета mixOmics [147] отдельно для фронтальной коры и миндалевидного тела. Анализ выполняли на основе VST-нормализованных значений экспрессии генов.

PCA использовали для оценки общей структуры вариабельности данных и распределения образцов без учета принадлежности к линиям. PLS-DA применяли для оценки разделения образцов по линиям с максимизацией различий. Для обоих методов использовали

матрицы экспрессии, включающие образцы в качестве наблюдений и гены в качестве переменных. Для PLS-DA оценивали разделение групп по латентным компонентам.

2.2.2.3. Анализ дифференциальной экспрессии отдельных генов

Для каждой ткани фактор *strain* (поведенческая группа) задавали так, чтобы референсным уровнем была неселектированння линия, а модель строили с дизайном по фактору *strain*. Подгонку модели выполняли стандартным алгоритмом DESeq2 с оценкой нормировочных коэффициентов по медианному отношению, оценкой дисперсий и Wald-тестами. Из одной подогнанной модели для каждой ткани извлекали два попарных сравнения: агрессивная группа с ручной и агрессивная группа с неселектированной. При извлечении результатов использовали Wald-тест, поправку Бенджамини–Хохберга. Дифференциально экспрессированными считали гены с padj < 0.05. VST также вычисляли отдельно для каждой ткани на тех же отфильтрованных матрицах чисел прочтений и с тем же дизайном по фактору *strain*, с учетом дизайна (*blind = FALSE)*, после подгонки модели, чтобы переиспользовать уже оцененные нормировочные коэффициенты и дисперсии. Попарные сравнения поведенческих групп извлекались в виде статистики Вальда с p-value, скорректированными по методу Бенджамина - Хохберга (padj).

Для обеих тканей каждый ген был классифицирован на основе двух попарных сравнений. Гены со специфичным для агрессивной линии паттерном экспрессии были определены на двух уровнях:

1) Набор с высоким уровнем достоверности (строгий фильтр). Для включения гена в этот набор требовалось, чтобы padj < 0,05 одновременно в сравнении агрессивных особей как с ручными, так и с неселектированными, а направления изменения экспрессии в этих сравнениях совпадали.
2) Расширенный набор (мягкий фильтр). Для этого набора критерий значимости во втором сравнении с участием агрессивной линии был ослаблен: требовалось padj < 0,05 только в одном из сравнений и номинальное p-value < 0,05 в другом, при совпадающем направлении изменения экспрессии. Поскольку такая схема отбора допускает более высокую долю ложноположительных результатов, расширенный набор рассматривали только в качестве потенциально значимого результата.

2.2.2.4. Построение сетей коэкспрессии

Построение сетей коэкспрессии генов проводили отдельно для каждого отдела мозга. Для построения сетей коэкспрессии использовали биоинформатический пакет Bioconductor BioNERO, реализующий соответствующие алгоритмы WGCNA для модулей коэкспресии [148]. Для стабилизации дисперсии данных экспрессии использовали параметр vstransform=TRUE. Проводили фильтрацию пропущенных данных с помощью функции *removena* пакета BioNERO. Удаляли гены, не демонстрировавшие экспрессии, а также гены с медианным значением количества считываний менее пяти. К преобразованным данным применяли фильтрацию по среднему значению, после чего для каждого условия отбирали 8000 генов с наибольшей вариабельностью экспрессии. Далее выбросы среди образцов удаляли с помощью метода ZKfiltering [149]. После успешного преобразования матрицы и стабилизации дисперсии полученные данные передавали в конвейер анализа WGCNA.

Для начала задавали диапазон значений степени (power) от 3 до 20 с шагом 1. С помощью функции pickSoftThreshold пакета WGCNA для каждого набора данных рассчитывали индекс соответствия топологии безмасштабной сети (scale-free topology fit index) и среднюю связность (mean connectivity). Рассматривали значения степени, при которых индекс соответствия топологии безмасштабной сети превышал 0.80. Кроме того, контролировали, чтобы средняя связность для каждого набора данных оставалась ниже 100. С учетом используемых фильтров были подобраны оптимальные величины power, равные 9 и 11 для фронтальной коры и миндалевидного тела, соответственно (Рисунок 2.1). В качестве входного параметра использовали тип сети signed, поскольку при этом типе сети отрицательные межгенные корреляции смещаются к нулю, благодаря чему модули содержат только те гены, для которых наблюдается изменение в одном направлении. Гены с отрицательной корреляцией, как правило, относятся к функционально различным программам (одна из которых подавляется, а другая – активируется), поэтому их объединение в один модуль затрудняет интерпретацию результатов).

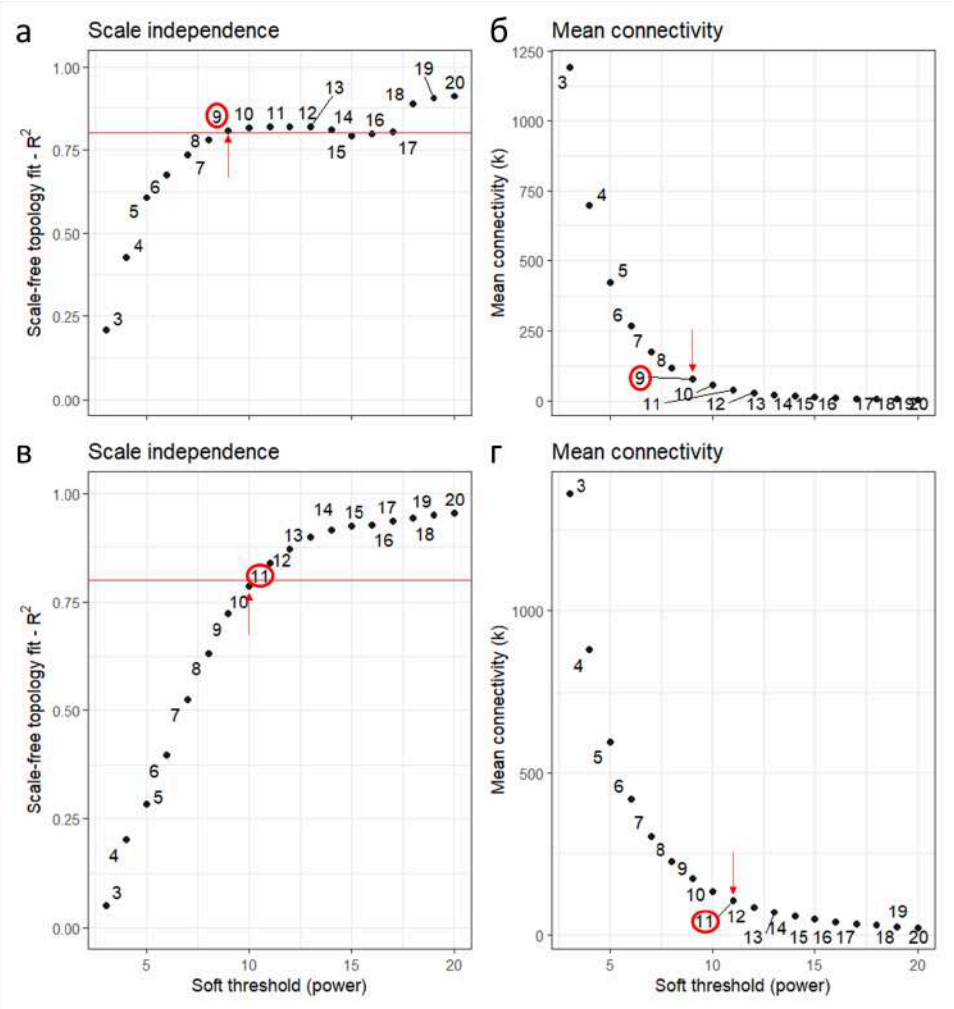


Рисунок 2.1 – Подбор оптимального значения ß для (1) фронтальной коры: а) по коэффициенту детерминации R^2. б) по средней связности; (2) миндалевидного тела: а) по коэффициенту детерминации R^2. б) по средней связности. Красной линией обозначен порог коэффициента детерминации, равный 0.8. Красными стрелками обозначены оптимальные значения ß и R^2.

Построение взвешенных сетей коэкспрессии генов для каждого объекта выполняли по стандартной процедуре с использованием функции exp2gcn пакета BioNERO. Силу связей между генами рассчитывали на основе матрицы смежности (adjacency matrix), задавая тип сети signed. Порог объединения модулей (module merging threshold) устанавливали на уровне 0.8.

2.2.2.3. Анализ сетей коэкспрессии генов и идентификация хаб-генов

После построения сетей для всех групп визуализировали дендрограммы и цветовое обозначение модулей. На графиках отображали как исходные, еще не объединенные модули, так и модули после объединения. Для более наглядного представления размеров модулей определяли количество генов в каждом из них. Строили тепловую карту корреляций между модулями и признаками. Для каждого условия определяли модули, демонстрирующие наиболее сильную корреляцию с фенотипом; серый модуль исключали из рассмотрения.

Центральные гены (“хаб”-гены) в соответствующих модулях определяли с помощью функции get_hubs_gcn пакета BioNERO. К “хаб”-генам относили гены с наибольшей степенью корреляции с остальными членами модуля. Для визуализации наиболее интересующих модулей с “хаб”-генами получали подграфы с помощью функции get_edge_list с фильтром слабых корреляций (rcutoff = 0.7) с последующей отрисовкой модулей в виде подграфов с помощью функции plot_gcn.

2.3 Анализ регуляторных элементов генома

2.3.1. Исследуемая выборка

Для исследования были использованы образцы фронтальной коры головного мозга лисиц из групп, селектированных на разное поведение по отношению к человеку (ручные, агрессивные, неселектированные), предоставленные зверофермой Института цитологии и генетики СО РАН в Новосибирске (Приложение Б). Для анализа был использован биологический материал 18 животных (самцы в возрасте 8-9 месяцев), по 6 образцов из каждой поведенческой группы.

2.3.2. Получение иммунопреципитированного хроматина

Для иммунопреципитации хроматина использовали протокол для нативного хроматина [150]. Ядра выделяли из 250 мг ткани префронтальной коры головного мозга лисиц. На фрагментацию хроматина с помощью микрококковой нуклеазы (Mnase, 9013-53-0, Sigma, США) брали по 1 млн. ядер от каждого образца. Половину использовали на иммунопреципитацию с

антителами к гистоновым меткам триметилирования четвертого лизина гистона H3 (H3K4me3 (Tri-Methyl-Histone H3 (Lys4) (C42D8) Rabbit mAb, “Cell Signaling”, США) и половину на иммунопрецицпитацию с антителами к гистоновым меткам ацетилирования двадцать седьмого лизина H3K27ac (Anti-Histone H3 (acetyl K27) antibody, “Abcam”, США). Для образца FBP85 дополнительно отобрали аликвоту ДНК до иммунопрецципитации для дальнейшего использования в качестве контроля. Полученные продукты очищали от неспецифических фрагментов, РНК и белков и выделяли ДНК фенол-хлороформ-изоамиловым методом.

2.3.3. Приготовление библиотек

Для приготовления библиотек из образцов ДНК после иммунопреципитации и из неиммунопреципитированной ДНК использовали набор реагентов NEBNext Ultra II DNA Library Prep Kit for Illumina (New England BioLabs) и протокол, рекомендованный фирмой-производителем набора. Для этого по 30 нг ДНК каждого образца использовали в реакциях достраивания концов, аденилирования 3'-концов и лигирования NEBNext адаптеров. Фрагменты библиотек очищали с помощью магнитных частиц NEBNext Sample Purification Beads (New England BioLabs), после чего библиотеку амплифицировали с использованием двойных индексов NEBNext (New England BioLabs). Оценку качества приготовленных библиотек осуществляли с помощью наборов DNA High Sensitivity Chip (Agilent) и High Sensitivity D5000 (Agilent) на приборах Agilent 2100 Bioanalyzer (Agilent) или TapeStation 4150 (Agilent), соответственно. Молярность библиотек оценивали методом qPCR с использованием набора KAPA Library Quantification Kit (KAPA Biosystems). Секвенирование полученных библиотек проводили на приборе Illumina HiSeq 2500 в режиме парно-концевых прочтений (101+101 п.н.).

2.3.4. Первичная биоинформатическая обработка данных секвенирования

Полученные в результате секвенирования прочтения были картированы на референсный геном лисицы (GCF_048418805.1) с помощью программы BWA mem [124, 141] (Приложение B). ПЦР-дупликаты маркировали с использованием команды MarkDuplicates из пакета программ Picard [125]. Для каждого образца был сформирован список локусов обогащения

картированными прочтениями (пиков, участков с повышенной плотностью картированных прочтений) с помощью программы MACS2 [151] с использованием параметров *-narrow, -q 0.001, --broad-cutoff 0.01,* потенциально соответствующих промоторам и энхансерам, маркированным эпигенетическими метками H3K4me3 и H3K27ac. В качестве контрольного образца при запуске MACS2 использовали библиотеку из неиммунопреципитированной ДНК для фильтрации ложноположительных пиков и учета уровня шума.

С помощью пакетов ensembldb и txdb для версии генома VulVul3 (GCF_048418805.1) были сформированы аннотации структуры генов лисицы в виде SQLite-баз данных [152]. С использованием полученных аннотаций и утилиты ChIPseeker [153] оценивали локализацию пиков прочтений относительно генов для каждого образца по координатам структуры генов для оценки соответствия полученных локусов промоторным регионам. Для получения координат точек начала транскрипции (TSS) аннотированных генов и определения координат промоторной области (на расстоянии ± 3000 п.н. от начала транскрипции (в направлении 5’ и 3’)[154]) также использовали аннотацию генома VulVul3. Координаты TSS были использованы для сопоставления выявленных локусов аннотированным генам и их типам.

2.3.5. Попарное сравнение локусов обогащения прочтениями между поведенческими группами

Полученные для всех образцов пики (локусы обогащения), соответствующие каждой из двух исследованных гистоновых меток (H3K4me3 и H3K27ac) объединяли с помощью пакета программ rtracklayer [155]. Затем, с помощью функции *regionCounts* из пакета csaw [156] для каждого образца производили пересчет числа прочтений, картированных в пределах координат объединенных локусов. Затем, с помощью функций *normFactors* и *normOffsets* из пакета программ csaw, проводили нормализацию полученных данных. Для выявления дифференциально обогащенных локусов (локусов со статистически значимым различием в уровне связывания (интенсивности сигнала) гистоновой модификации между группами животных) проводили межгрупповое попарное сравнение (агрессивные особи против ручных и агрессивные особи против неселектированных) и статистический анализ, в том числе расчет p-value с поправкой на множественное тестирование, с помощью пакетов программ *IHW* и edgeR [157, 158]. Для выявления наиболее значимых регионов, связанных с агрессивным поведением, проводился поиск локусов, для которых при попарном сравнении агрессивных особей с ручными и неселектированными наблюдается статистически значимая разница (p-value<0.05) для обоих сравнений с одинаковым направлением изменения.

2.4. Мультиомиксный интегративный анализ геномных, транскриптомных и эпигеномных данных, полученных в рамках данной работы

В предыдущих разделах настоящей диссертации приводилось описаний методов, применяемых для выявления агрессиво-специфичных отличий от двух других исследуемых линий на разных генетических уровнях (популяционная структура, однонуклеотидные варианты (2.1), экспрессия мРНК (2.2), активность регуляторных элементов (2.3)). Далее было решено сосредоточиться на выявлении корреляции между этими уровнями и ее интерпретацией.

Глава 3. Результаты и обсуждение

3.1. Полногеномные данные

3.1.1. Когорта и набор вариантов

Полногеномные данные были проанализированы для 37 самцов и 9 самок серебристой лисицы, принадлежащих к трем линиям: селектированной на агрессивное поведение (15 особей), ручное поведение (15 особей) и неселектированной контрольной (16 особей). После совместного генотипирования, жесткой фильтрации, ограничения набора бивалентными ОНВ, фазирования с помощью SHAPEIT2 и применения фильтров к когорте (MAF $\geq$ 0,05; доля пропущенных генотипов в локусе $\leq$ 0,2) варианты были прорежены по неравновесию по сцеплению (окно 50 kb, $r^2 = 0,1$) для анализа популяционной структуры. Ограничение когорты самцами исключает пол как дополнительный фактор при анализе структуры.

3.1.2. Анализ главных компонент

Анализ главных компонент (PCA), проведенный на наборе ОНВ после прореживания по неравновесному сцеплению, четко разделяет три исследуемые линии (Рисунок 3.1). На PC1 приходится 23,0%, а на PC2 – 14,9% общегеномной дисперсии (в сумме около 38%). Линия ручных лисиц образует компактный, хорошо обособленный кластер на одном из краев оси PC1, тогда как агрессивная и неселектированная линии занимают соседние положения и разделяются между собой вдоль PC2. При построении PCA по компонентам PC1 и PC3 наблюдается четкая дифференциация по PC1 между кластером ручной линии и смешанным кластером агрессивной и неселектированной линий.

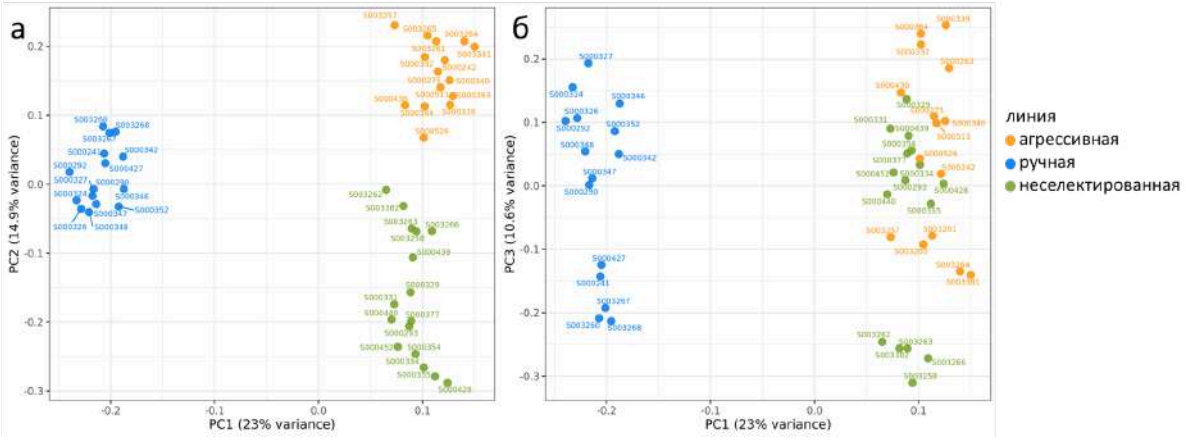


Рисунок 3.1 – Популяционная структура геномов исследуемых линий в виде главных компонент для (а) первой и второй компонент; (б) первой и третьей компонент. Цветами обозначены исследуемые линии: оранжевый – агрессивная, зеленый – неселектированная, синий – ручная.

Такая структура согласуется с разделением линий, описанным для данного эксперимента по селекции в исследовании генома лисицы [67]. Поэтому мы рассматривали ее как уже установленную характеристику, а не как новый результат. В данном случае анализ популяционной структуры служит для подтверждения того, что линии достаточно дифференцированы для последующего анализа селекции по геномным окнам.

Две особенности геометрии PCA имеют существенное значение для интерпретации. Во-первых, наиболее выраженная общегеномная ось разделяет ручную линию и две остальные: геномы ручных лисиц наиболее обособлены среди трех линий. Во-вторых, агрессивная линия располагается близко к неселектированной линии в пространстве общегеномных осей: общегеномная дифференциация не определяется преимущественно агрессивным фенотипом. Таким образом, можно предположить, что генетические различия, связанные с агрессивным поведением, затрагивают отдельные участки генома, а не приводят к масштабным изменениям генетического профиля агрессивной линии в целом. Следовательно, ожидается, что сигнал, связанный с агрессивностью, будет локализован в отдельных островах селекции, а не проявится в виде однородного сдвига по всему геному.

3.1.3. Анализ предковых компонент (ADMIXTURE)

Выбор модели на основе пятикратной перекрестной проверки для K = 2–5 (десять повторов для каждого K) показал, что средняя ошибка перекрестной проверки была минимальной при K = 2 (0,748; лучший повтор – 0,747) и монотонно возрастала с увеличением K (0,767 при K = 3; 0,836 при K = 4; 0,943 при K = 5). Формально предпочтительный выбор K = 2 отделяет ручную линию от кластера, включающего агрессивную и неселектированную линии, что соответствует результатам по PC1. Поскольку схема разведения предусматривает три линии, мы также приводим разложение на компоненты при K = 3: каждая линия преимущественно представлена собственной компонентой происхождения; компонента ручной линии является наиболее однородной, а неселектированная линия содержит небольшую долю компоненты, характерного для агрессивной линии.

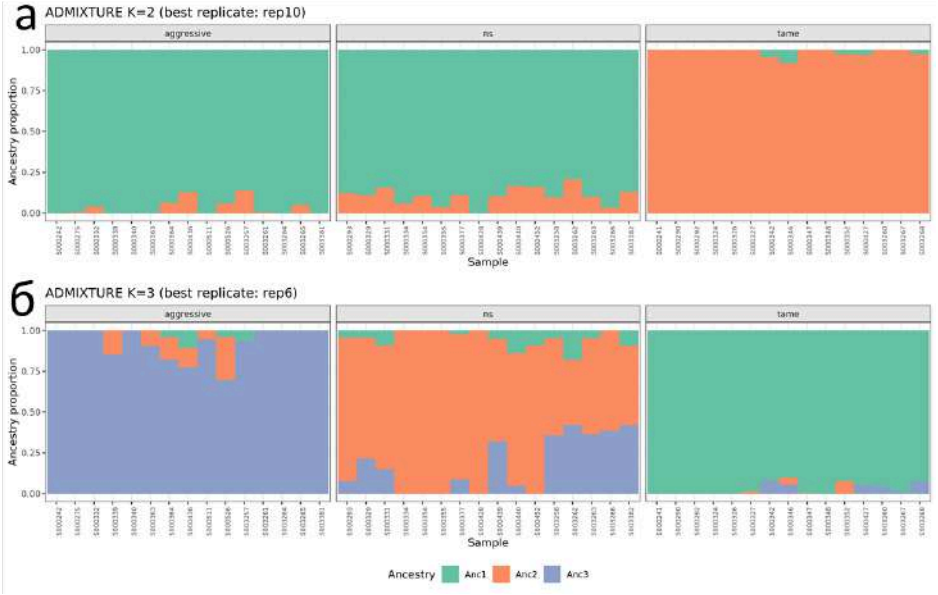


Рисунок 3.2 – Популяционная структура исследуемых линий в виде предковых компонент для (а) K=2; (б) K=3. Цветами обозначены линии: оранжевый – агрессивная, зеленый – неселектированная, синий – ручная.

Данная картина может объясняться сохранением генетической изменчивости, присутствующей в исходном фермерском поголовье, из которого были отобраны основатели неселектированной линии, вследствие чего она не является генетически обособленной популяцией. Монотонное увеличение ошибки перекрестной проверки дополнительно подтверждает интерпретацию PCA: структура происхождения определяется прежде всего отделением ручной линии, тогда как различия между агрессивной и неселектированной линиями, хотя и отчетливо проявляются при K = 3 и по PC2, слабее выражены на уровне всего генома.

3.1.4. Родство внутри отдельных линий

Анализ родства методом KING-robust выявил только одну пару родственных особей среди 46 животных: двух полных сибсов из агрессивной линии (коэффициент родства 0,185, превышающий порог для родства третьей степени – 0,0625). Родственных особей в линиях неселектированных и ручных лисиц выявлено не было. Было решено оставить всю когорту из 46 животных в основном анализе, поскольку одна пара родственных особей не может обусловить кластеризацию на уровне линий.

3.1.5. Анализ FST по окнам между селектированными линиями

Для выявления участков генома, подвергшихся действию отбора в исследуемых линиях был проведен скользящий анализ окон FST по методу Weir–Cockerham (окна размером 50 Кб, шаг 10 kb; ~231,400 перекрывающихся окон для каждого попарного сравнения: FST_{AN} (агрессивная линия против неселектированной), FST_{AT} (ручная против агрессивной), FST_{TN} (ручная против неселектированной) показавший следующие средние значения по геному: FST_{AN}=0,075 , FST_{TN}=0,134 и FST_{AT}=0,153. Полученный результат соответствует геометрии PCA: животные агрессивной линии наиболее близки к неселектированному генетическому фону по аутосомным маркерам, тогда как ручная линия демонстрирует наибольшую дифференциацию. FST_{AT} превышает значения для обоих сравнений с неселектированной линией. Это указывает на то, что две селектированные линии во многих локусах отклонялись от неселектированного фона в противоположных направлениях. Все три полученных значения FST высоки для сравнений

внутри одного вида, что отражает десятилетия разведения замкнутых линий с повторяющимися эффектами бутылочного горлышка на уровне основателей [67].

Чтобы определить, какие участке генома обогащены различиями, 50 окон с наибольшими значениями FST для каждого попарного сравнения были объединены в регионы (FST_{AN}: 12 регионов общей протяженностью 1,05 Мб; FST_{AT}: 20 регионов – 1,34 Мб; FST_{TN}: 15 регионов – 1,11 Мб). Большинство таких регионов специфично для отдельных сравнений. При этом, только один регион является строго специфичным для агрессивной линии – он входит в число окон с наибольшими значениями FST в обоих сравнениях с участием агрессивной линии, но отсутствует в сравнении ручной и неселектированной линий: NC_132787.1 (chr8), 31,54–31,69 Мб. Этот участок содержит гены *ATF7IP* и *PLBD1*; взвешенные значения FST: FST_{AN}=0,70; FST_{AT}=0,88.

3.1.6. Выявление ОНВ, специфичных для животных агрессивной линии

Данные полногеномного секвенирования 46 лисиц из двух селектированных линий (агрессивной и ручной) и неселектированной были использованы для выявления однонуклеотидных вариантов. Отличающиеся от референсного варианты, специфичные для животных агрессивной линии, определяли на основе частот встречаемости альтернативного аллеля, учитывающего направление фенотипического различия: альтернативный аллель должен был присутствовать не менее чем у 80% животных агрессивной линии и не более чем у 10% животных ручной/неселектированной линий (фильтр 80/10/10). В результате такого отбора вариантов было выявлено 3 582 ОНВ, специфичных для животных агрессивной линии. Этот фильтр намеренно не основывали на параметрических предположениях: он не предполагает определенного распределения частот аллелей, а непосредственно отражает историю селекции. Аллель, достигший фиксации или близкий к фиксации в агрессивной линии, но редко встречающийся в обеих других линиях, с меньшей вероятностью возник вследствие генетического дрейфа в условиях общей фермерской среды. При этом контрольная неселектированная линия позволяет учитывать дрейф и эффекты основателей, общие для селектированных линий.

3.1.7. Анализ распределения ОНВ в геноме

Выявленные 3 582 ОНВ распределены по геному неравномерно. Кластеризация плотности в окнах размером 1 Мб показывает, что основная часть сигнала сосредоточена в небольшом числе хромосомных регионов (Рисунок 3.3).

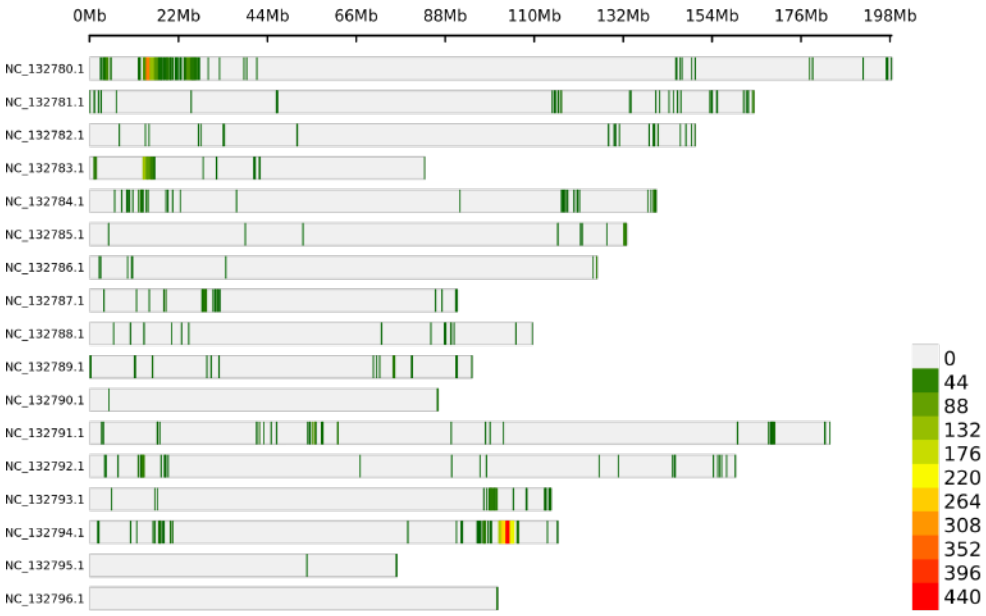


Рисунок 3.3 – Распределение специфичных для агрессивной линии ОНВ по хромосомам в выявленных окнах. Цветовая шкала отражает количество ОНВ в каждом окне: от 0 (серый) до 440 (красный).

Геном разделили на окна размером 1 Мб и сохранили все окна, содержащие хотя бы один ОНВ; границы каждого окна затем скорректировали по крайним положениям ОНВ. В результате получили 242 окна общей протяженностью 57,5 Мб. Основная часть сигнала приходится на три мегарегиона (Таблица 2.1).

Таблица 2.1 – Распределение специфичных для агрессивной линии ОНВ по мегарегионам.

Мегарегион	Протяженность	Число ОНВ	Доля
NC_132794.1 (chr15)	101–106 Мб	около 1 100	31%
NC_132780.1 (chr1)	13–18 Мб	около 630	18%
NC_132783.1 (chr4)	1–17 Мб	около 310	9%

3.1.8. Окна, содержащие гены определенных функциональных групп: корреляция с ранее опубликованными данными

В настоящем исследовании были выявлены окна ОНВ, специфичных для агрессивной линии, содержащие следующие гены:

- Мегарегион chr15 (51 ген): *SHTN1/shootin1* (поляризация нейронов и синаптогенез), *SLC18A2/VMAT2* (везикулярный транспорт моноаминов – потенциально значимый компонент моноаминергических механизмов агрессии), *ADRB1*, *GFRA1*, *KCNK18*, *TCF7L2*, *EMX2* (формирование корковых областей), *VAX1*, *ABLIM1*, *CASP7*, *HABP2*, *AFAP1L2*; регион chr15 (1 ген): *SORCS1*.
- Мегарегион chr1, 13–18 Мб (189 генов): *SHANK1* (структурный белок постсинаптического уплотнения), *SLC17A7/VGLUT1* (транспорт глутамата), *GRIN2D* (субъединица NMDA-рецептора), *KCNC3*, *SYT3* (синаптотагмин), *SLC8A2*, *NPAS1*, кластер генов калликреинов *KLK1–KLK13*, *PNMA8A/B/C*, *CALM3*, *CPT1C*, *TRPM4*, *BAX/BBC3/CARD8*.
- Мегарегион chr4 (13 генов): *MFSD8/CLN7* (нейрональный цероидный липофусциноз, или болезнь Баттена), *FAT4*, *PLK4*.
- Регион chr1, 23–27 Мб (72 гена): *NOL8*, *WNK2*, *SHC3*, *SEMA4D* (направление роста аксонов), *S1PR3*, *SYK*, *TRIM28*, *BICD2*, *CHMP2A*.
- Регион chr14, 99–100,8 Мб (32 гена): *ACAN*, *NTRK3* (рецептор нейротрофинов), *KIF7*, *POLG*, *RLBP1*.

Таким образом, выявленные мегарегионы и регионы, обогащенные ОНВ, специфичными для агрессивных лисиц, содержат гены, связанные с функционированием нервной системы, формированием и пластичностью нейронных связей, синаптической передачей, нейромедиаторной регуляцией и развитием мозга.

Ранее в работе Kukekova et al. (2018) были выявлены геномные регионы, ассоциированные с ручным и агрессивным поведением у лисиц той же экспериментальной модели. Было установлено, что в кандидатных регионах обогащены гены, связанные с неврологическими расстройствами у человека, поведением у мышей и процессами, ассоциированными с одомашниванием у собак [67]. В качестве сильного позиционного кандидата на роль в формировании ручного поведения был выделен ген *SorCS1*, кодирующий важный белок транспорта AMPA-рецепторов и нейрексинов. При сопоставлении выявленных нами специфичных для агрессивной линии регионов FST и окон OHB с опубликованными данными был выявлен 21 ген, присутствующий в обоих наборах. При этом, 15 из 21 гена (~71%) расположены на 8 хромосоме (NC_132787.1). Так, гены *ATF7IP* и *PLBD1* находятся в единственном выявленном нами регионе FST, специфичном для агрессивной линии (31,54–31,69 Мб), а ген *GRIN2B* содержит локус со специфичным для агрессивной линии гаплотипом «11» (32,24–32,33 Мб). Эти три гена также входят в окна FST, выявленные в опубликованном исследовании при сравнении ручной и агрессивной линий.

Также пересечение было обнаружено для шести генов: гена *SORCS1*, связанный с ручным поведением, — находится внутри нашего мегарегиона OHB на chr15 (101–106 Мб), а также для генов *GMNC* (chr3), *NDUFA3* (chr1), *TACR3* (chr4), *CRADD* и *SOCS2* (chr10). В рамках полученных окон OHB нами также было выявлено несколько других «нейрональных» генов, не вошедших в опубликованные регионы FST (в частности, *SHANK1*, *SLC17A7*, *SLC18A2*, *MFSD8*, *NTRK3*, *GRIN2D*, кластер генов семейства KLK). В совокупности эти результаты указывают на синаптическую пластичность как на возможный общий механизм, связанный с различиями в поведении.

Таким образом, в специфичных для агрессивной линии окнах впервые были обнаружены гены, ассоциированные с неврологическими заболеваниями, включая *MFSD8/CLN7*, кластер *KLK*, *POLG* и *TCF7L2*. Также, были выявлены гены, связанные с синаптической пластичностью, – *SHANK1*, *SLC17A7*, *GRIN2D*, *SYT3*, *SLC18A2*, *NTRK3* и *SEMA4D*, функционально сопоставимые с *SorCS1*, а также небольшое окно из 5 OHB, 2 из которых находятся в интронных областях гена *SORCS1*.

3.1.9. Анализ выявленных миссенс-вариантов, специфичных для агрессивных лисиц

Среди 3 582 OHB преобладали варианты, расположенные в некодирующих участках генома: интронные – 1 756 (49,0%), межгенные – 1 407 (39,3%), расположенные в генах

некодирующих РНК – 382 (10,7%), синонимичные – 22 (0,6%), затрагивающие сайты сплайсинга – 5 (0,1%) и миссенс-варианты – 10 (0,3%).

Десять миссенс-ОНВ были локализованы в семи генах. Для каждого варианта проводили аннотацию, включающую реконструкцию белковой последовательности на основе геномной аннотации лисицы. Для анализа консервативности позиций, в которых были выявлены миссенс-варианты, положения аминокислотных замен у лисицы проецировали на соответствующие позиции человеческих белков с использованием множественного выравнивания полноразмерных белков RefSeq десяти видов (рыжая лисица, песец, собака, серый волк, золотистый шакал, енотовидная собака, кошка, большая панда, мышь; канонический белок человека использовали в качестве системы координат; выравнивание MAFFT L-INS-i). Затем объединяли предсказания VEP SIFT/PolyPhen для канонических транскриптов человека, оценки различий по шкале Grantham, гидропатические и зарядовые различия, сведения о доменах UniProt и информацию о консервативности позиций по выравниванию 10 видов. Модели, основанные на характеристиках аминокислотных замен, позволяют ранжировать десять ОНВ по степени химических различий между исходной и измененной аминокислотами (Таблица 3.1).

Стоит, однако, учесть, что «радикальная» замена в гене *ACAN* представляет собой восстановление предкового, референсного состояния: аллель агрессивной линии восстанавливает аминокислотный остаток, соответствующий референсной последовательности человека. При этом выравнивание 10 видов показывает, что именно аминокислотные остатки, присутствующие у лисиц, являются производными состояниями (Приложение Г). Следовательно, большие расстояния по шкале Grantham могут отражать дивергенцию последовательности у лисиц и не демонстрируют возникновение производного функционального изменения в агрессивной линии.

Данные о консервативности позволяют разделить десять ОНВ на три группы:

(а) Замены в полностью консервативных позициях — *ACAN* P772S (пролин у всех 10 видов), *SHTN1* V579I (валин у всех 10 видов), *NOL8* S359C (серин у всех 9 видов, вошедших в выравнивание; у мыши эта область делетирована) и *HABP2* F24Y (фенилаланин присутствует у всех девяти видов, кроме мыши).

(б) Замены, нарушающие консенсусное состояние у представителей отряда хищных — *AFAP1L2* S673L (серин у всех 8 представителей отряда хищных; человек, у которого в этой позиции находится пролин, и мышь, у которой находится аланин, являются исключениями) и *PNMA8A* P132A (пролин у 9 из 10 видов; серин у человека представляет собой производное состояние).

(в) Замены, соответствующие производному состоянию, обнаруженному у другого вида — *AFAP1L2* A679T (треонин соответствует аминокислотному состоянию, характерному для кошки, собаки, волка и шакала), *ACAN* L793S (лейцин встречается только у двух видов рода Vulpes; серин в аллеле агрессивной линии восстанавливает консенсусное состояние у псовых и соответствует состоянию у человека) и *ACAN* V1812M (валин у всех 8 представителей отряда хищных; метионин в аллеле агрессивной линии соответствует состоянию у человека).

Таблица 3.1 – Ранжированный список специфичных для агрессивной линии миссенс-мутаций с оценками от разных предсказателей.

ген	замена а/к у лисы	замена а/к у человека	расстояние по Грэнтему	класс категории по Грэнтему	показатель SIFT	показатель PolyPhen
AFAP1L2	S673L	P626	145	radical	tolerated (0.13)	benign (0.01)
AFAP1L2	A679T	V632	58	moderate	-	-
NOL8	S359C	S361	112	radical	deleterious (0.01)	benign (0.191)
ACAN	P772S	P763	74	moderate	deleterious (0.02)	unknown
SHTN1	V579I	V579	29	conservative	tolerated (0.26)	possibly damaging (0.631)
LOC112929753	R189W	нет ортолога	101	radical	-	-
PNMA8A	P132A	S136	27	conservative	tolerated (0.17)	benign (0.006)
HABP2	F24Y	F24	22	conservative	tolerated (0.43)	benign (0.072)
ACAN	L793S	S796	145	radical	-	-
ACAN	V1812M	M1974	21	conservative	-	-

Предсказанные структуры белков дикого типа (WT) и мутантных вариантов (MUT) – модели AlphaFold v2 для мономеров шести ортологичных белков и двух участков домена ACAN, а также модель ESMFold для белка лисицы LOC112929753 – сравнивали после суперпозиции по алгоритму Кабша. Анализировали значения RMSD (мера средней разницы между координатами соответствующих атомов в двух или более совмещенных трехмерных молекулярных структурах) структурного ядра (для остатков с высокой уверенностью

предсказания), RMSD окрестности сайта замены, изменение значения pLDDT (показатель локальной уверенности модели AlphaFold в правильности предсказания структуры белка для каждой отдельной аминокислоты) в этой области и изменения контактов.

Поскольку структуры WT и MUT предсказываются независимо друг от друга, у неупорядоченных белков глобальная конфигурация петель может различаться между двумя расчетами даже при отсутствии существенных структурных изменений. Для компактной структуры LOC112929753 глобальный RMSD составил 0,69 Å; это значение использовали для калибровки нижнего предела структурного шума.

Результаты для отдельных OHB представлены ниже.

1. Аминокислотная замена P772S в белке ACAN (Рисунок 3.4) была определена как наиболее интересная. Пролин консервативен у всех 10 исследованных таксонов. Замена устраняет изгиб полипептидной цепи, обусловленный пролином, на границе домена KS. Значение pLDDT в области замены снижается на 17 пунктов – это наибольшее изменение уверенности предсказания среди всех исследованных вариантов. При этом структурное ядро домена KS воспроизводится с RMSD 0,80 Å. Предсказание SIFT указывает на вероятно неблагоприятный эффект (0,02).

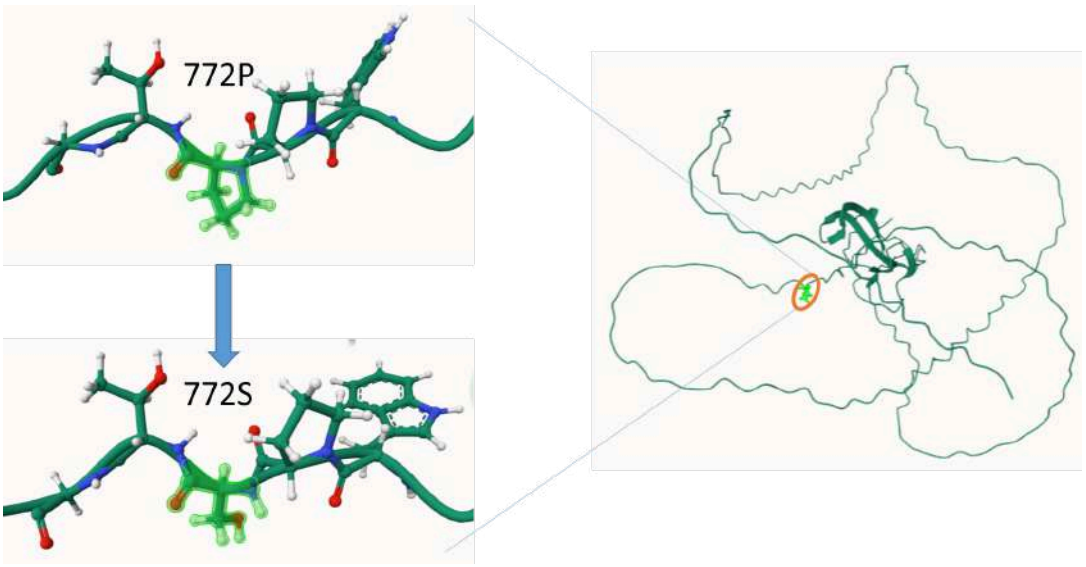


Рисунок 3.4 – Предположительная 3D-структура домена KS белка ACAN (справа), аминокислотная замена P772S (слева): референсная аминокислота и специфичная для агрессивной линии.

2. Консервативная аминокислотная замена V579I в в неупорядоченном линкере белка SHTN1 (Рисунок 3.5) одной гидрофобной аминокислоты на другую (Grantham = 29) в позиции, где валин консервативен у всех 10 видов. Согласно PolyPhen, замена, возможно, оказывает повреждающее действие (0,631), тогда как SIFT классифицирует ее как переносимую (0,26). Поскольку этот участок не образует устойчивой пространственной структуры, которую можно было бы нарушить, аргументация основана на эволюционной консервативности остатка и предсказании PolyPhen.

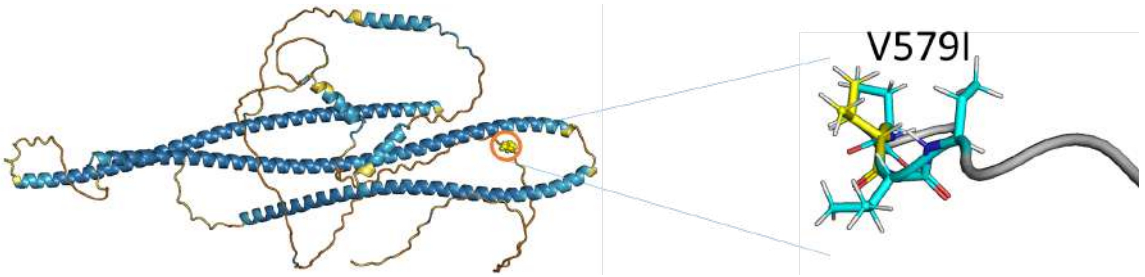


Рисунок 3.5 – Предположительная 3D-структура белка SHTN1 (слева), аминокислотная замена V579I (справа): аминокислота, специфичная для агрессивной линии.

3. Аминокислотная замена серина на цистеин S359C в белке NOL8 (Рисунок 3.6) находится в консервативном участке, обогащенном серином и треонином (серин у всех 9 видов, вошедших в выравнивание; у мыши соответствующая область не выровнена из-за делеции). Гидроксильная группа, потенциально участвующая в фосфорилировании, заменяется тиольной группой. Значение pLDDT в области замены снижается на 5,9 пункта; также утрачивается один контакт.

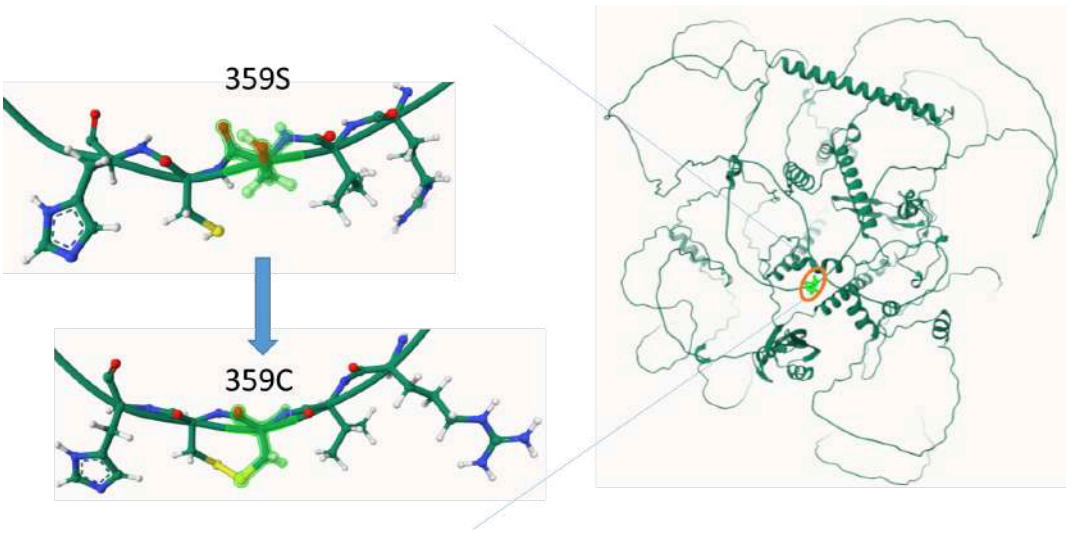


Рисунок 3.6 – Предположительная 3D-структура белка NOL8 (справа), аминокислотная замена S359C (слева): референсная аминокислота и специфичная для агрессивной линии.

4. Ген, кодирующий белок AFAP1L2, – единственный ген, в котором изменения кодирующей последовательности S673L и A679T совпадают с различиями в экспрессии: уровень экспрессии этого гена повышен во фронтальной коре агрессивной линии. Вариант S673L заменяет серин, консервативный у всех 8 представителей отряда хищных, на аминокислоту, не встречающуюся в соответствующей позиции выравнивания. Тем не менее оба предсказательных алгоритма, примененные к последовательности в системе координат человека, классифицируют замену как переносимую (P626L: PolyPhen — “доброкачественный” вариант, 0,01; SIFT — “переносимая замена”, 0,13). Вариант A679T соответствует аминокислотному состоянию, обнаруженному у кошки, собаки, волка и шакала (для соответствующей замены V632T в белковой аннотации человека отсутствует путь, обусловленный заменой одного нуклеотида, поэтому предсказание не выполнялось).

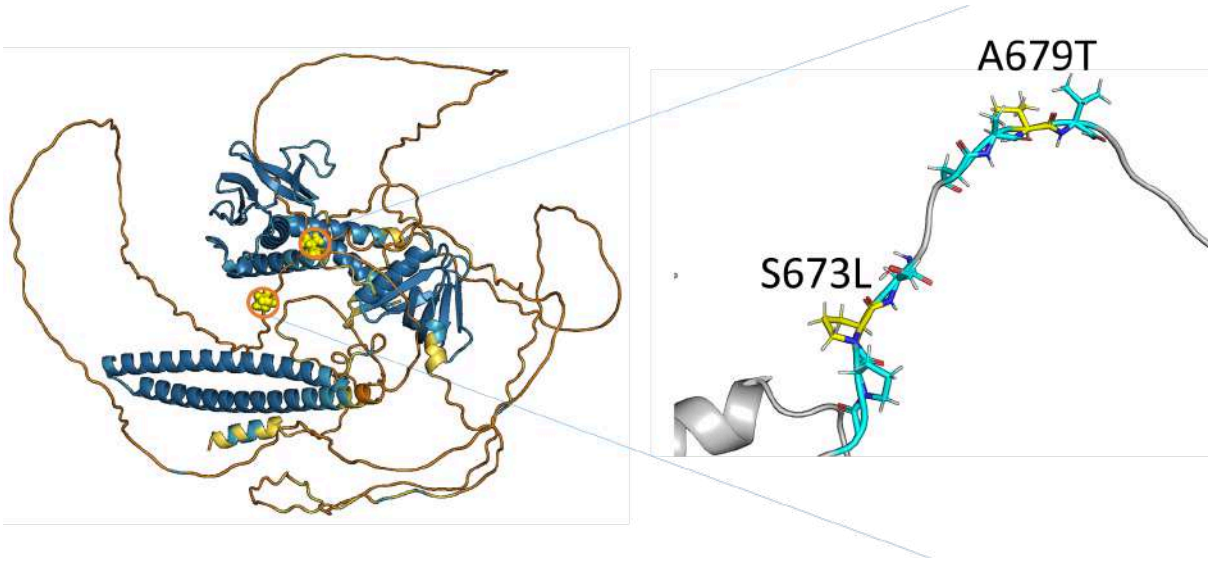


Рисунок 3.7 – Предположительная 3D-структура белка AFAP1L2 (слева), аминокислотные замены S673L и A679T (справа): аминокислоты, специфичные для агрессивной линии.

5. Вариант R189W в белке LOC112929753 является единственным, затрагивающим участок со свернутой структурой белка. Внутри кармана положительный заряд (+1) заменяется нейтральным. Общая структура при этом сохраняется: RMSD структурного ядра составляет 0,50 Å, глобальный RMSD – 0,69 Å. Формируется один дополнительный контакт, а pLDDT в области замены повышается на 4 пункта. Однако, учитывая недоступность сопоставления с ортологами, степень функциональной значимости этой замены остается неопределенной.

Мутации в генах, кодирующих белки PNMA8A (P132A), HABP2 (F24Y), ACAN (L793S, V1812M), являются консервативными либо восстанавливающими предковое состояние заменами, в неупорядоченных участках. Их структурный эффект невозможно надежно отличить от фоновых расхождений между предсказанными моделями.

Таким образом, набор ОНВ, специфичных для агрессивной линии, воспроизводит основной вывод исследования генома лисицы, полученный в независимом анализе с более строгими критериями соответствия направлению фенотипических различий: в геномных регионах, ассоциированных с селекционными линиями, представлены гены, связанные с синаптической пластичностью, неврологическими расстройствами и поведением (Kukekova et al., 2018). Частотный фильтр 80/10/10 не требует статистических предположений и непосредственно отражает историю селекции, а контрольная неселектированная линия позволяет учитывать альтернативное объяснение, связанное с генетическим дрейфом. Однако такой подход выявляет

сравнительно мало изменений в кодирующих последовательностях: среди 3 582 ОНВ обнаружено лишь десять миссенс-ОНВ (0,3%). Это согласуется с представлением о том, что отбор по поведенческим признакам может действовать преимущественно через регуляторные генетические варианты. В частности, 49% ОНВ расположены в интронах, 39% – в межгенных областях и 11% – в генах некодирующих РНК. Среди миссенс-вариантов только *ACAN* P772S объединяет все линии доказательств высокой значимости: консервативность пролина у всех десяти видов, механистически обоснованную структурную гипотезу (устранение изгиба полипептидной цепи на границе KS-домена), значительное снижение уверенности предсказания для данного участка при сохранении воспроизводимости структурного ядра и предсказание SIFT о повреждающем эффекте замены. Варианты *SHTN1* V579I и *NOL8* S359C также заслуживают внимания, однако их обоснование опирается на консервативность позиции и локальные химические свойства, а не на структурные данные: для *SHTN1* это предсказание PolyPhen о возможном повреждающем эффекте замены валина, консервативного у всех десяти видов; для *NOL8* — предсказание SIFT о повреждающем эффекте замены гидроксильной группы, способной подвергаться фосфорилированию. Два варианта *ACAN*, восстанавливающие состояние референсной последовательности, и радикальная, но классифицированная как переносимая замена *AFAP1L2* S673L наглядно показывают, почему одной лишь величины расстояния Грантема недостаточно для вывода о функциональном эффекте: важны направление замены относительно консенсусного состояния, определяемого выравниванием, и согласованность предсказаний.

3.1.11. Интеграция общей геномной популяционной структуры и окон ОНВ

Исследуемые линии лисиц демонстрируют выраженную дифференциацию на общегеномном уровне: первые две главные компоненты полностью разделяют их, первая и третья компонента разделяют кластер ручных с кластером агрессивных и неселектированных. Средние значения окон FST между линиями составляют 0,075–0,153, что является достаточно высоким показателем для сравнений внутри одного вида. Это ожидаемое следствие десятилетий разведения замкнутых линий, основанных на небольших группах основателей, и воспроизводит структуру популяции, описанную в исследовании сборки генома лисицы для данного эксперимента [67].

И геометрия PCA, и кривая перекрестной проверки ADMIXTURE показывают, что геномная дифференциация определяется преимущественно ручной линией: PC1 и

предпочтительное решение K = 2 выделяют ручную линию, тогда как агрессивная линия остается близкой к неселектированному фону (среднее FST = 0,075 против 0,134–0,153 для сравнений с участием ручной линии). Таким образом, поведенческий фенотип, поддерживаемый интенсивным направленным отбором, не сопровождается общегеномной дивергенцией несущей его линии. Это соответствует архитектуре, выявленной в экспериментальной модели ручного поведения у крыс, где ручное поведение ассоциировано с ограниченным числом геномных областей, а не с однородным сдвигом по всему геному [34]. Результат также согласуется с геномными исследованиями самого эксперимента по селекции лисиц [67]. Следовательно, ожидается, что сигнал, связанный с агрессией, будет концентрироваться в локальных регионах.

Примечательно, что единственный регион с окнами с максимальным значением FST, строго специфичном для агрессивной линии, не содержит ни одного из 3 582 OHB, специфичных для агрессивной линии. Это означает, что участок характеризуется выраженной дифференциацией, но не содержит OHB, достигающих порогов частотного фильтра 80/10/10. Следовательно, дифференциация здесь, по-видимому, связана с гаплотипным фоном, а не с фиксацией отдельных вариантов в агрессивной линии.

Перекрытие островов FST с набором OHB, специфичных для агрессивной линии, носит избирательный, а не общегеномный характер. В четырех из 47 регионов с наиболее высокими значениями FST присутствует как минимум один OHB; три из них представляют особый интерес:

1. NC_132794.1 (chr15), 103,50–103,59 Mb — участок внутри крупнейшего мегарегиона OHB (101–106 Mb). В одном окне с высоким FST, расположенном в области гена *ATRNL1*, обнаружено 134 агрессиво-специфичных OHB.
2. NC_132787.1 (chr8), 32,29–32,35 Mb — участок в области гена *GRIN2B*, несущего гаплотип, специфичный для агрессивной линии и выявленный при фазировании данных (80% в агрессивной линии, 0% в ручной и неселектированной линиях; FDR = $4{,}0 \times 10^{-9}$).
3. NC_132796.1 (chr17), 100,82–100,94 Mb — участок в области гена *GPC4*, содержащий 9 агрессиво-специфичных OHB.

Остальные регионы с наиболее высокими значениями FST, включая наиболее выраженные регионы для сравнения ручной и агрессивной линий на NC_132791.1 (chr12), 62–90 Mb, и NC_132782.1 (chr3), 9,05–9,12 Mb (взвешенное FST = 0,92), не перекрываются с мегарегионами агрессиво-специфичных OHB. Это подтверждает, что набор таких OHB, выделенный с помощью частотного фильтра, и общегеномный анализ FST отражают в значительной степени неперекрывающиеся аспекты дифференциации линий.

Сканирование окон FST позволяет явно показать взаимосвязь между двумя типами сигналов. Регионы с наиболее высокими значениями FST и окна, содержащие ОНВ, специфичные для агрессивной линии, перекрываются лишь избирательно: большинство регионов FST специфично для отдельных сравнений, а ни один из регионов для сравнения ручной и неселектированной линий не попадает в мегарегион ОНВ. Это ожидаемо, если регионы FST отражают совокупную дифференциацию, обусловленную генетическим дрейфом, эффектами основателей и отбором в любой из линий, тогда как фильтр DNP 80/10/10 выделяет специфичную для агрессивной линии составляющую этой дифференциации. Наибольшую уверенность обеспечивают участки, в которых результаты двух подходов совпадают.

Наиболее выраженное совпадение наблюдается в области гена *ATRNL1* на chr15, около 103,5 Мб: одно окно с высоким значением FST содержит 134 ОНВ, специфичных для агрессивной линии, и расположено внутри крупнейшего мегарегиона ОНВ. Второй важный участок – область гена *GRIN2B* на chr8, около 32,3 Мб, для которого получены независимые подтверждения двумя методами: выявлен гаплотип, специфичный для агрессивной линии (80% в агрессивной линии, 0% в обеих контрольных линиях; FDR = $4{,}0 \times 10^{-9}$), а также окно, входящее в топ-50 по FST для сравнения агрессивной и неселектированной линий. Ген *GRIN2B* кодирует субъединицу GluN2B NMDA-рецептора – ключевого регулятора синаптической пластичности и возбуждающей синаптической передачи [159]. Белок GPC4, компонент синаптического перицеллюлярного матрикса на chr17, около 100,8 Мб, содержащий 9 ОНВ, демонстрирует аналогичное совпадение результатов.

Единственный остров FST, строго специфичный для агрессивной линии, представляет дополнительный интерес. Интервал chr8 31,54–31,69 Mb (гены *ATF7IP*, *PLBD1*) характеризуется выраженной дифференциацией в обоих сравнениях с участием агрессивной линии, но не содержит ОНВ, специфичных для этой линии. Следовательно, дифференциация в данном участке связана с гаплотипным фоном, не сопровождаясь фиксацией отдельных вариантов на частотах, специфичных для агрессивной линии. ATF7IP является кофактором, стабилизирующим гистонметилтрансферазу SETDB1, и необходим для формирования гетерохроматина, опосредованного H3K9me3 [160].

Наконец, значение FST для сравнения ручной и агрессивной линий превышает значения для обоих сравнений с неселектированной линией (0,153 против 0,134 и 0,075). Наиболее аккуратным объяснением является разнонаправленная дивергенция: во многих локусах ручная и агрессивная линии отклонялись от неселекционного фона в противоположных направлениях, что увеличивало их взаимное FST по сравнению с расстоянием каждой линии от фона, близкого к исходному генофонду основателей. Острова FST, специфичные для сравнения ручной и

агрессивной линий (например, chr12 62–90 Мб и chr3 9,05–9,12 Мб с взвешенным FST = 0,92), являются кандидатными областями такого разнонаправленного отбора.

3.1.12. Поиск ОНВ в районе гена *MAOA*, специфичных для агрессивной линии

Дополнительно, учитывая выявленные ранее в опубликованных данных различия в уровнях моноаминовых нейромедиаторов у двух животных моделей, селектированных на поведения (лис и норок) [161, 162], а также ассоциации вариантов гена *MAOA* с агрессивным поведением у других видов, было решено провести исследование генетических вариантов этого гена в геномах лис и норок. Для этого использовали данные геномного секвенирования американской норки *Neogale vison*, в которых проводили поиск генетических вариантов относительно референсного генома самки американской норки (GCF_020171115.1). В общей сложности, по сравнению с референсным геномом, в исследованной выборке норок в области гена *MAOA* и его промоторном регионе было выявлено 277 генетических вариантов, из них ни один не попал в кодирующую область гена, 271 вариантов были локализованы в интронах и 6 попали в промоторную область. Анализ ассоциации выявленных вариантов с поведением, проведенный с помощью точного теста Фишера в программе PLINK, не выявил ни одного статистически значимого, с учетом поправки на множественные тестирование, генетического варианта. В геномных данных лис проводили аналогичный поиск генетических вариантов, всего в теле гена *MAOA* было выявлено 59 вариантов, 3 из них попадали в аннотированные экзоны. Однако, статистически значимой ассоциации с агрессивным поведением ни для одного из этих вариантов выявлено не было. Таким образом, в исследованных выборках вариабельность гена *MAOA* не подтверждается как значимый генетический фактор, ассоциированный с агрессивным поведением у лис и норок.

3.2. Транскриптомный анализ фронтальной коры лисиц разных линий

3.2.1. Сходство транскриптомных профилей агрессивных и неселектированных лисиц

Анализ главных компонент (PCA), проведенный на 1000 генов с наиболее вариабельной экспрессией во фронтальной коре (ФК) не демонстрирует различий между исследуемыми

линиями (Рисунок 3.8а). Образцы равномерно распределены относительно компонент. Однако при применении метода PLS-DA, с указанием параметра поведения, линия ручных лисиц образует компактный, хорошо обособленный кластер, тогда как агрессивная и неселектированная линии образуют смешанный кластер и не разделяются (Рисунок 3.8б). Как и в случае фронтальной коры, для миндалевидного тела (МНД) по результатам PCA не было выявлено формирования кластеров исследуемых линий (Рисунок 3.8бв). Однако, метод PLS-DA позволяет увидеть четкие различия между всеми тремя группами, формирующими четкие обособленные кластеры.

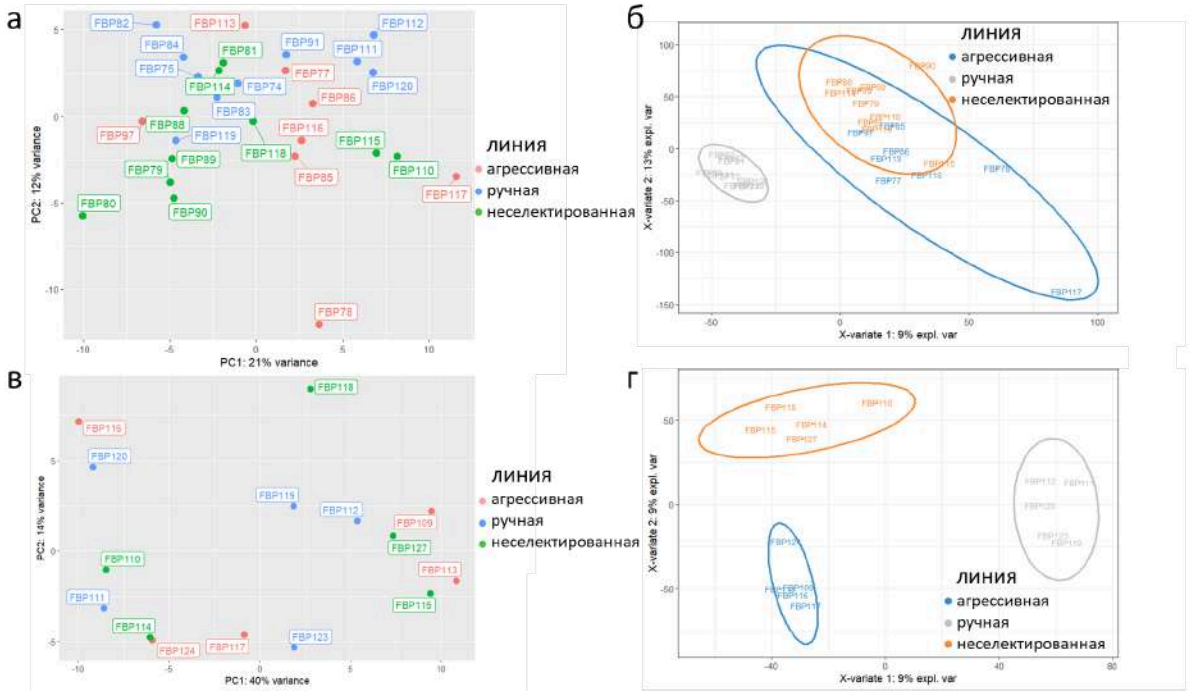


Рисунок 3.8 – Популяционная структура транскриптомов исследуемых линий в виде главных компонент для ФК (а,б) и МНД (в,г) с использованием методов снижения размерности: Анализ главных компонент (PCA) (а,в); Дискриминантный анализ методом частичных наименьших квадратов счетов (PLS-DA) (б,г). Названия исследуемых линий обозначены цветом; слева: красный – агрессивная, зеленый – неселектированная, синий – ручная; справа: синий – агрессивная, оранжевый – неселектированная, серый – ручная.

Данное наблюдение позволяет предположить, что на уровне дифференциальной экспрессии во фронтальной коре наиболее существенным (хоть и немногочисленным, учитывая значение латентной компоненты X-variate=9%) будет различие между ручными особями и агрессивными/неселектированными на поведение. Для миндалевидного тела, на основании полученной картины (первые две латентные компоненты объясняют по 9% общей вариации

экспрессии генов), можно предположить равнозначные различия между группами в определенных многомерных направлениях экспрессионного профиля.

Проведенный анализ дифференциальной экспрессии действительно позволил выявить асимметрию в различиях между исследуемыми линиями. В обеих тканях наименьшее количество дифференциально экспрессирующихся генов наблюдалось между селектированными на агрессию и неселектированными лисицами (30 генов во фронтальной коре и 20 генов в миндалевидном теле), тогда как генов, дифференциально экспрессирующихся между селектированными на дружелюбие и неселектированными лисицами было выявлено практически в 4 раза больше (111 генов во фронтальной коре и 46 генов в миндалевидном теле). Это наблюдение указывает на то, что на уровне экспрессии генов в мозге состояние дружелюбия является более видоизмененным, тогда как агрессивная линия сохраняет профиль экспрессии, близкий к исходному состоянию предковой формы, не подвергавшейся отбору. Этот результат соответствует ранее опубликованным геномным данным о том, что популяция ручных лисиц в большей степени приобрела отличие от неселектированной, чем агрессивная, что объясняется как более длительным (на десятилетие) отбором ручной линии, так и вторичной интрогрессией особей неселектированной популяции в агрессивную в 1990-х годах [67].

3.2.2. Дифференциально экспрессирующиеся гены у агрессивных лисиц

Классификация дифференциально экспрессирующихся генов между агрессивными и ручными, а также между агрессивными и неселектированными лисицами во фронтальной коре и миндалевидном теле по характеру значимости позволила выделить блок из 14 и 12 генов, соответственно, характеризующихся специфичным для агрессивной линии транскрипционным профилем (согласованное направление изменения, padj < 0,05 в попарных сравнениях агрессивных особей с ручными и с неселектированными) (Таблица 3.2).

В рамках менее строгого статистического фильтра (padj < 0,05 хотя бы для одного сравнения, p-value<0,05 для второго сравнения) выявлены блоки из 100 генов во фронтальной коре и 34 генов в миндалевидном теле. Число генов, специфичных для ручной линии, составило 133 (во фронтальной коре) и 50 (в миндалевидном теле), тогда как гены, общие для отбора на ручное и агрессивное поведение, единичны: 1 во фронтальной коре и 2 в миндалевидном теле, что указывает на то, что два направления селекции привели к преимущественно неперекрывающимся изменениям экспрессии, а не к единому «селекционному» паттерну.

Таблица 3.2 – Дифференциально экспрессируемые гены во фронтальной коре и миндалевидном теле. Значения logFC (логарифм изменения кратности), p-value и padj указаны для попарных сравнений уровня экспрессии у агрессивных особей с ручными (AT) и у агрессивных с неселектированными (AN)

ген	**$logFC_{AT}$**	**$p\text{-}value_{AT}$**	**$padj_{AT}$**	**$logFC_{AN}$**	**$p\text{-}value_{AN}$**	**$padj_{AN}$**
CTXN3	-0,89	8,42E-15	6,85E-11	-0,75	7,54E-11	2,45E-07
NLRP12	1,80	9,69E-12	5,25E-08	1,94	5,41E-13	8,79E-09
LOC140599525	-2,36	4,49E-11	1,83E-07	-1,76	9,97E-07	1,47E-03
LOC140600068	2,58	1,24E-10	3,62E-07	1,57	4,55E-05	3,03E-02
LOC112912849	1,42	1,79E-09	3,24E-06	1,64	5,24E-12	2,84E-08
CCDC192	4,95	1,17E-08	1,73E-05	2,51	2,22E-05	1,73E-02
BST1	-0,89	3,24E-08	3,22E-05	-1,08	2,59E-11	1,05E-07
LOC140599791	-2,45	4,10E-07	2,47E-04	-2,39	7,62E-07	1,24E-03
ZSCAN23	-0,52	6,03E-07	3,38E-04	-0,41	8,30E-05	4,50E-02
LOC112931556	-1,78	4,46E-06	1,76E-03	-1,89	1,11E-06	1,51E-03
LOC140599796	-2,84	4,82E-06	1,78E-03	-2,73	1,12E-05	1,07E-02
MAN1B1	0,39	8,22E-05	1,39E-02	0,69	2,57E-12	2,09E-08
CA2	-0,27	1,73E-04	2,24E-02	-0,35	6,93E-07	1,24E-03
RCSD1	-0,36	5,28E-04	4,95E-02	-0,42	4,84E-05	3,03E-02
ген	**$logFC_{AT}$**	**$p\text{-}value_{AT}$**	**$padj_{AT}$**	**$logFC_{AN}$**	**$p\text{-}value_{AN}$**	**$padj_{AN}$**
LOC112912849	2,65	6,48E-23	1,06E-18	2,14	2,24E-16	1,80E-12
LOC112907532	-1,51	2,58E-10	1,34E-06	-1,89	6,40E-16	3,43E-12
LOC112929743	-0,82	5,15E-10	1,68E-06	-0,66	5,14E-07	1,03E-03
LOC140599791	-3,30	1,64E-09	4,49E-06	-2,32	2,35E-05	2,51E-02
PPP1R17	0,93	1,13E-08	2,06E-05	1,25	3,40E-14	1,37E-10
LOC140595121	-2,34	1,86E-07	2,17E-04	-2,21	8,30E-07	1,48E-03
LOC112928444	2,20	8,57E-07	8,77E-04	2,05	4,29E-06	5,29E-03
LOC140598352	1,23	1,30E-06	1,01E-03	1,04	2,71E-05	2,56E-02
LOC112907522	-0,79	1,26E-06	1,01E-03	-1,10	9,18E-12	2,46E-08
MFSD8	0,56	5,92E-06	3,73E-03	0,68	2,58E-08	5,93E-05
LOC140599278	-1,57	1,18E-05	6,24E-03	-1,73	1,33E-06	1,78E-03
HAS1	0,93	7,87E-05	2,71E-02	1,81	3,39E-13	1,09E-09

Среди девяти генов фронтальной коры, экспрессия которых была снижена у агрессивных животных, – *CTXN3*, кодирующий нейрон-специфический мембранный белок, обогащенный в коре головного мозга [163], локус которого ассоциирован с риском шизофрении в исследовании генов-кандидатов [164]; *BST1*, кодирующий ADP-рибозилциклазу, генерирующую циклическую ADP-рибозу, участвующую в регуляции секреции окситоцина и в формировании социального узнавания и социального поведения у грызунов [165, 166]; *CA2* (карбоангидраза II), ключевой фермент буферизации pH в мозге, также сопряженный с лактат/H^+-транспортом в астроцитах [167, 168]; ген *ZSCAN23*, транскрипционный регулятор со SCAN-доменом и цинковыми пальцами [169]. Среди генов фронтальной коры с повышенной у агрессивных лисиц экспрессией – ген *NLRP12*, кодирующий противовоспалительный NOD-подобный рецептор, ослабляющий активацию микроглии и нейровоспаление в ЦНС [170]; *MAN1B1*, α-1,2-маннозидаза эндоплазматического ретикулума, мутации потери функции которой вызывают аутосомно-рецессивную умственную отсталость (MAN1B1-СДГ) [171].

В миндалевидном теле у агрессивных животных для всех трех генов с установленной аннотацией наблюдалось повышение экспрессии: : для гена *PPP1R17* (G-субстрат), ингибитора протеинфосфатазы 1, высокообогащенногой в клетках Пуркинье мозжечка и связанного с цГМФ-зависимой синаптической сигнализацией [172]; гена *MFSD8* (*CLN7*), лизосомальногоый транспортера, мутации которого вызывают нейрональный цероидный липофусциноз типа 7 (– нейродегенеративное заболевание детского возраста) [173]; и гена *HAS1*, кодируюшего (гиалуронансинтазуа 1), синтезирующую гиалуронан во внеклеточном матриксе мозга – каркасе, модулирующем нейровоспаление и поведение нервных клеток [174]. В совокупности наблюдаемые различия в уровне экспрессии связывают специфический для агрессивной линии транскрипционный профиль с социально-поведенческой нейромодуляцией (*BST1*), нейроиммунным тонусом (*NLRP12*, *HAS1*), гомеостазом лизосом и гликозилирования (*MFSD8*, *MAN1B1*) и синаптической сигнализацией (*PPP1R17*, *CA2*).

Два гена, прошедших строгий фильтр (padj < 0.05 при попарном сравнении агрессивных особей с ручными и неселектированными), были специфичны для агрессивной линии в обеих тканях – их число превышает ожидаемое случайным образом (гипергеометрический тест, $p = 4{,}7 \times 10^{-5}$ относительно общего набора протестированных генов): *LOC112912849*, экспрессия которого повышена у агрессивных животных в обеих тканях, и локус *LOC140599791*, подобный *ADCY1*, экспрессия которого в обеих тканях снижена (ФК: $logFC_{AT}$=-2,4, $logFC_{AN}$=-2,39; МНД: $logFC_{AT}$=-3,30, $logFC_{AN}$=-2,32). Сигнал, связанный с *ADCY1*-подобным геном, заслуживает особого внимания, поскольку аденилатциклаза 1, активируемая ионами Ca^{2+}, необходима для синаптической пластичности и формирования памяти о страхе, зависящей от опыта [175], а ее фармакологическое ингибирование усиливает угасание страха [176], что может иметь

непосредственное отношение к нейронным связям, опосредующим страх и агрессию и зависящим от миндалевидного тела.

При использовании мягкого фильтра – вместо значимости в обоих сравнениях с агрессивной линией требовалась значимость хотя бы в одном таком сравнении для каждой ткани (padj < 0,05 в A vs. T или A vs. N в обеих тканях) и номинальное p-value<0.05 в другом – были выявлены 27 генов с дифференциальной экспрессией и в ФК, и в МНД. Из них 24 гена демонстрировали полное совпадение направленности изменений экспрессии между тканями в обоих сравнениях с агрессивной линией (11 генов с повышенной и 13 — со сниженной экспрессией у агрессивных животных. К аннотированным генам относились *BPHL*, *BST1*, *CCDC134*, *DENND10*, *HAS1*, *IDNK*, *PELO*, *SIGLECL1* и *TUBB2B*, а также 15 локусов LOC. Эти 24 гена с совпадающей направленностью изменений экспрессии образуют ядро дифференциальной экспрессии, общее для обеих тканей.

Стоит отметить, что ранее в работе Wang et al. было выявлено 146 генов с дифференциальной экспрессией между ручными и агрессивными лисицами в префронтальной коре [68]. Это сопоставимо по масштабу с полученными нами 175 генами в ФК в сравнении агрессивных особей с ручными, несмотря на различия в используемых референсной сборке, алгоритме анализа и когортах образцов. С использованием дополнительной – неселектированной – линии мы выделяем именно специфичные для агрессивных особей различия.

3.2.3. Анализ коэкспрессии и направленных регуляторных сетей

Для поиска потенциальных паттернов экспрессии был использован алгоритм анализа WGCNA с помощью “оболочной” программы BioNERO. Конечные матрицы экспрессии, использованные для анализа обоих отделов мозга содержали 8000 генов с наибольшей вариабельностью экспрессии. Кластеризация и фильтрация образцов позволила оставить в анализе 29 образцов для фронтальной коры и 14 образцов для миндалевидного тела. Для проведения анализа WGCNA рекомендуемое число образцов ≥15 (при меньшем значении может быть проблема с подбором параметра power). Однако, так как для этого отдела мозга были достигнуты оптимальные значения R^2 и mean connectivity при значении power=11, лежащем в допустимом диапазоне, было решено пренебречь недостаточным размером выборки.

На рисунке 2.1 раздела Материалы и методы представлены результаты оценки соответствия модели топологии безмасштабной сети для каждого обоих отделов мозга. Эти

результаты позволяют дополнительно снизить уровень шума при идентификации модулей после построения сети. Как видно из рисунка, для изучаемых наборов данных значение параметра мягкого порогового преобразования (soft threshold power), равное, соответственно, девяти и одиннадцати, обеспечивало индекс соответствия топологии безмасштабной сети выше 0.80. При этом средняя связность оставалась значительно ниже 100. Таким образом, оба наборы данных были признаны пригодными для дальнейшего анализа.

Для выявления модулей алгоритм WGCNA использует алгоритм динамического отсечения дендрограммы (dynamic tree cut). Приложение Д суммирует результаты идентификации модулей для наборов данных обоих отделов мозга. Для наглядности на рисунках представлены как исходные, еще не объединенные модули, так и модули после объединения. Анализ каждой дендрограммы показал, что ни в одном из условий ветви, соответствующие генам, не достигали высоты 0.50. Поскольку высота дендрограммы отражает степень различия между ее конечными узлами, отсутствие групп генов, достигающих уровня 0.50, указывает на сравнительно низкую согласованность профилей экспрессии отдельных генов в образцах каждого условия. Это также отражается в количестве генов, входящих в модули в каждом условии.

На рисунке 3.9 представлена сгруппированная гистограмма, обобщающая количество генов в каждом модуле. Цвет столбцов соответствует цветовому обозначению соответствующих модулей.

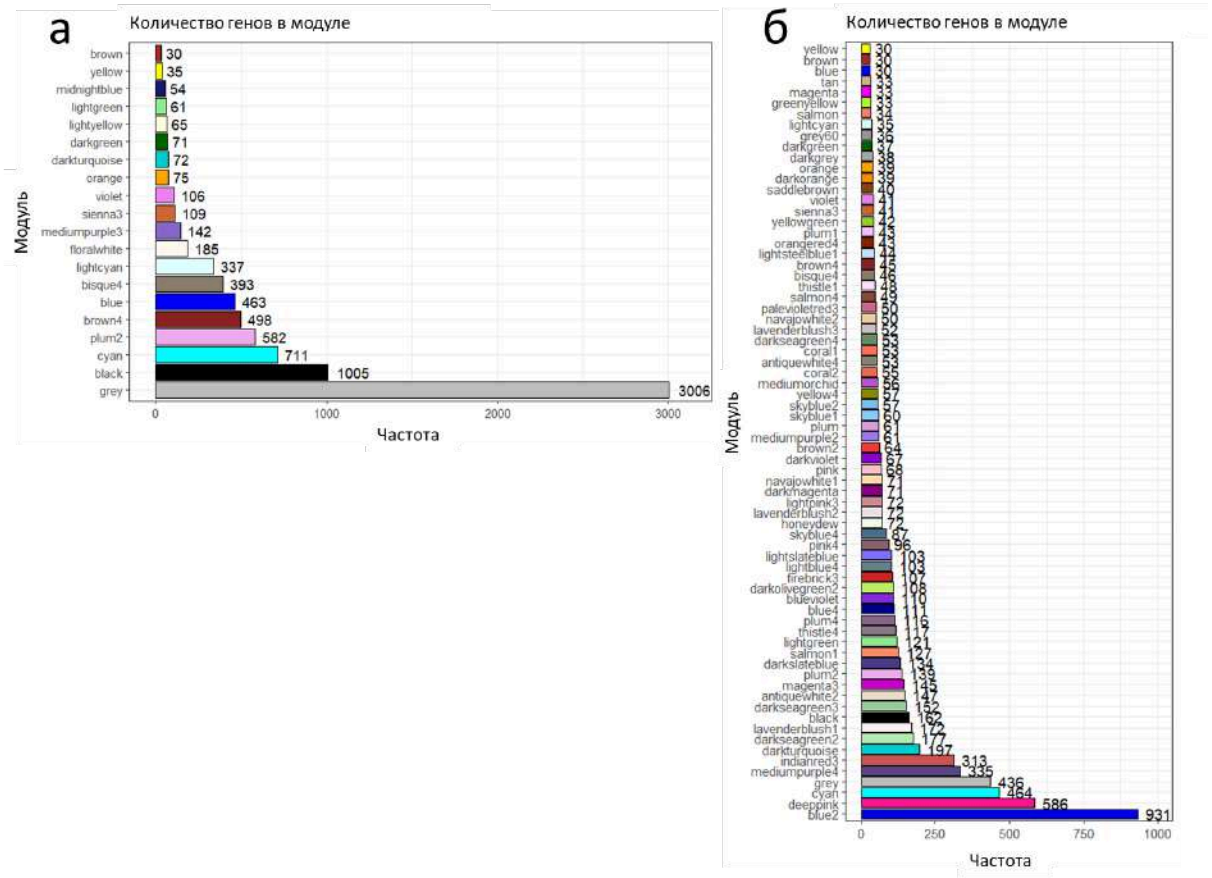


Рисунок 3.9 – Количество генов в каждом модуле коэкспрессии для (а) фронтальной коры и (б) миндалевидного тела.

Далее рассматривали только модули со специфичным для группы агрессивных животных паттерном экспрессии (p-value<0.05).

3.2.3.1. Модули коэкспрессии во фронтальной коре

Во фронтальной коре относительно агрессивного фенотипа были обнаружены корреляции, как позитивные, так и негативные. Как показано на рисунке 3.10, умеренные (|r|≥ 0.4) положительные корреляции были обнаружены для светлого-желтого (*lightyellow*) и песочного (*tan*) модулей, а также отрицательная для зеленого (*green*) модуля. Для каждого из модулей был определен генный состав (Приложение Е) с помощью функции get_edge_list.

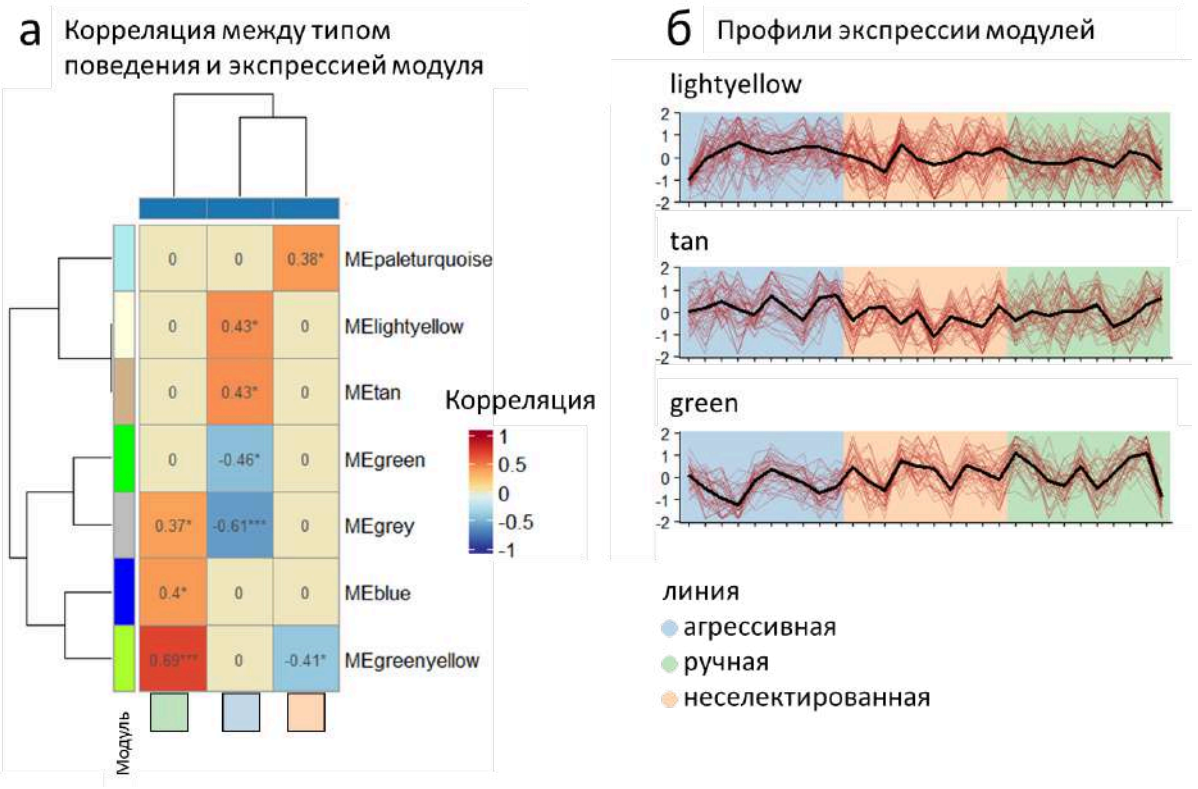


Рисунок 3.10 – а) Корреляция выявленных модулей коэкспрессии генов с изучаемыми поведенческими группами для фронтальной коры. Исследуемые группы обозначены цветом: ручная – зеленым, агрессивная – голубым, неселектированная – оранжевым. б) Профили экспрессии трех модулей с наибольшим по модулю коэффициентом корреляции.

Наибольший интерес представляет модуль *green*, включающий в себя 35 генов, объединяет гены с известными функциями в развитии нейронов, формировании нейронных связей и клеточной сигнализации.. Функции генов модуля связаны с несколькими клеточными процессами. Так, ген *SHTN1*, идентифицированный как “хаб”-ген (Рисунок 3.11), кодирует белок shootin-1, участвующий в поляризации нейронов, росте аксонов и направленном движении их конусов роста, что было показано в экспериментах на культивируемых нейронах и животных моделях [175, 176].

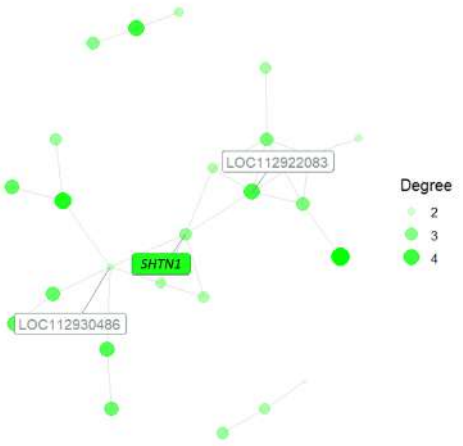


Рисунок 3.11 – Модуль *green* коэкспрессии генов во фронтальной коре. В центре обозначен “хаб”-ген *SHTN1*.

Ген *SEMA3E* кодирует сигнальную молекулу семафорин-3E, участвующую во взаимодействиях, направляющих развитие нервных связей; роль сигнального пути SEMA3E–PLXND1 в формировании проекций гиппокампа была ранее показана на мышах [179]. Ген *SLC8A1* кодирует натрий-кальциевый канал NCX1, участвующий в регуляции внутриклеточной концентрации кальция [180]. Ген *KREMEN1* кодирует рецептор, участвующий в регуляции сигнального пути Wnt через взаимодействие с DKK1 и комплексом LRP6 [181]. Ген *SCAI* кодирует белок, участвующий в регуляции транскрипции, опосредованной MRTF, и подавлении инвазии клеток. В исследовании на крысах белок Scai также обнаружили в нейронах; эксперименты на клетках нейронального происхождения показали, что варианты SCAI различаются по влиянию на рост нейритов, однако функция гена *SCAI* в нервной системе изучена не полностью [182]. Ген *DRD4* кодирует дофаминовый рецептор D4, сопряженный с G-белками и участвующий в передаче дофаминового сигнала [183]. Ген *SCN11A* кодирует потенциал-зависимый натриевый канал Nav1.9, экспрессируемый преимущественно в ноцицепторах – сенсорных нейронах, участвующих в передаче сигналов боли. Функция канала подтверждена исследованиями мутаций у людей и соответствующими моделями на мышах [184]. Наконец, в контексте поведения особенно особенно интересен ген *GRM3*, кодирующий метаботропный глутаматный рецептор: в ранее опубликованном исследовании геномных различий между линиями серебристых лис, отбиравшихся по ручному и агрессивному поведению, был описан однонуклеотидный вариант в этом гене и различия его частот между линиями [68]. Таким образом, функциональные характеристики значительной части генов позволяют предположить, что выявленный модуль отражает скоординированные различия в

экспрессионной организации фронтальной коры, затрагивающие глутаматергическую сигнализацию, кальциевый гомеостаз и процессы нейрональной организации.

3.2.3.2 Модули коэкспрессии в миндалевидном теле

В миндалевидном теле также относительно агрессивного фенотипа были обнаружены корреляции, как позитивные, так и негативные. Как показано на рисунке 3.12, умеренные (|r|≥ 0.4) положительные корреляции были обнаружены для кремово-песочного (*navajowhite1*) и светло-синего (*lightblue4*) модулей, а также отрицательная для синего (*blue4*) модуля. Для каждого из модулей был определен генный состав (Приложение Ж) с помощью функции get_edge_list.

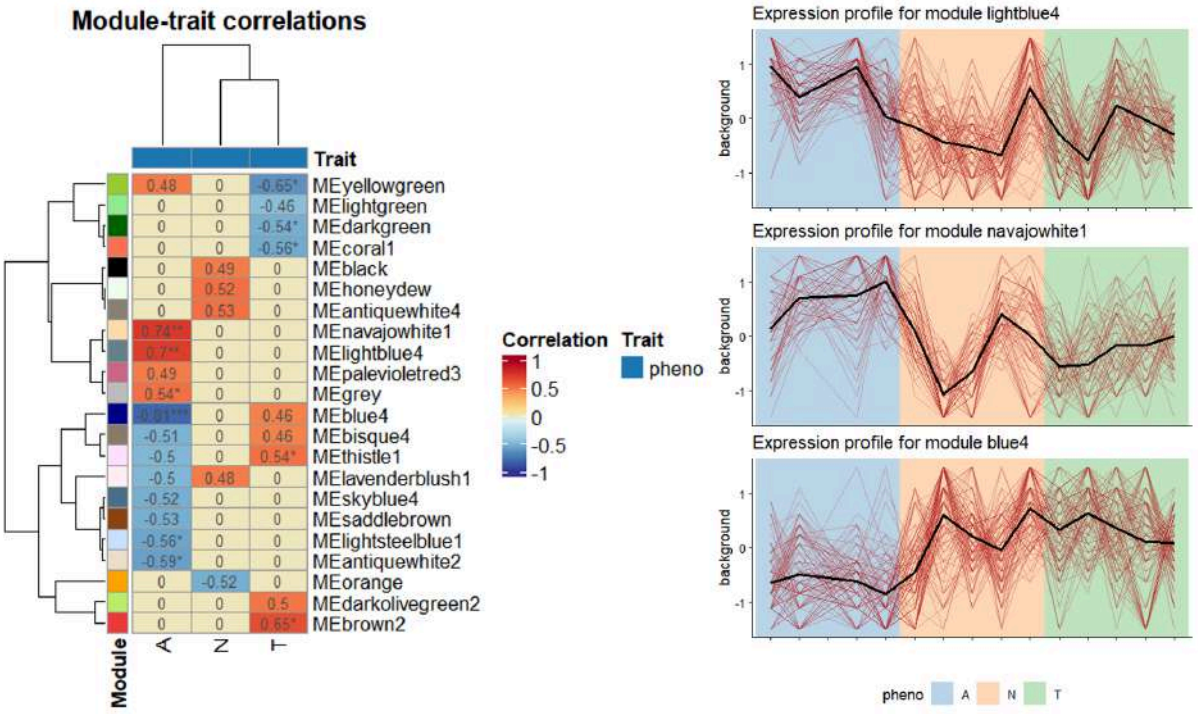


Рисунок 3.12 – а) Корреляция выявленных модулей коэкспрессии генов с изучаемыми поведенческими группами для миндалевидного тела. Обозначенные фенотипы: T – ручной, A – агрессивный, N – неселектированный. б) Профили экспрессии трех модулей с наибольшим по модулю коэффициентом корреляции.

Интерес представляет модуль *navajowhite1*, положительно коррелирующий с агрессивным фенотипом (r=0.74). Этот модуль включает в себя 69 генов, связанных с иммунной

регуляцией, внеклеточным матриксом, клеточной сигнализацией, метаболизмом и поддержанием клеточных органелл. В качестве “хаб”-гена был идентифицирован дифференциально экспрессируемый ген *PPP1R17* (Рисунок 3.13), описанный ранее в разделе 2.2 как ингибитор протеинфосфатазы 1, высокообогащенный в клетках Пуркинье мозжечка и связанный с цГМФ-зависимой синаптической сигнализацией.

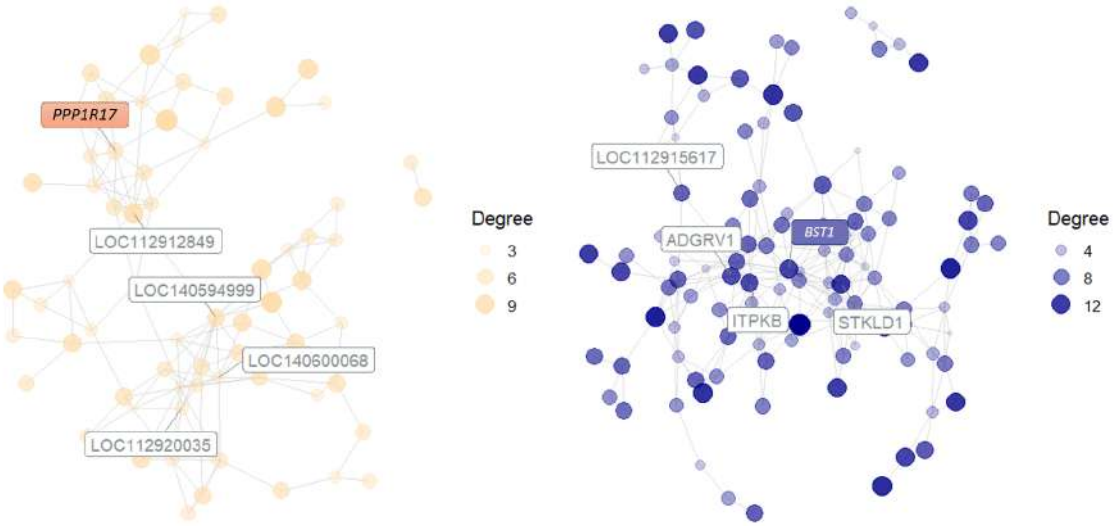


Рисунок 3.13 – Модули (а) *navajowhite1* и (б) *blue4* коэкспрессии генов в миндалевидном теле. Обозначены “хаб”-гены *PPP1R17* и *BST1*, соответственно.

Ген *FN1* кодирует фибронектин – компонент внеклеточного матрикса, участвующий в клеточной адгезии и сборке матрикса [185]. Дифференциально экспрессируемый ген *HAS1* участвует в синтезе гиалуронана, другого компонента внеклеточного матрикса; его роль исследовали, в частности, в контексте изменений тканей при болезни Альцгеймера [186]. Ген *BMP2* кодирует сигнальный белок семейства BMP, участвующий в регуляции клеточных процессов; экспериментальные исследования связывают BMP2-сигнализацию с изменениями цитоскелета и фокальных контактов в ответ на механическую нагрузку [187].

Другая группа представлена генами иммунной регуляции. Дифференциально экспрессируемый ген *NLRP12* участвует в регуляции врожденного иммунного ответа, а ген *CD69* кодирует белок, экспрессия которого повышается при активации иммунных клеток [188, 189]. Ген *HAVCR2* кодирует иммунорегуляторный рецептор TIM-3, а ген *TRIM14* участвует в регуляции врожденных иммунных сигнальных путей [190, 191].

Среди генов с установленными функциями в нервной системе также выделяется ген *P2RX3*, кодирующий АТР-зависимый пуринергический рецептор, участвующий в передаче

сигналов в сенсорных нейронах, в частности связанных с ноцицепцией. Роль P2RX3/P2RX7-сигнализации изучалась также в клетках гиппокампа мыши [192]. Ген *MPZ* кодирует основной белок миелина периферической нервной системы, а дифференциально экспрессируемый ген *MFSD8* – лизосомальный мембранный белок, нарушения которого связаны с нейрональным цероидным липофусцинозом. Для гена *DRD4* известно, что рецептор D4 модулирует дофаминергическую передачу и активность нейронов префронтальной коры; его влияние на внутриклеточные сигнальные пути зависит от активности нейронов [183]. Это свидетельство функции DRD4 в нервной системе, но не прямое доказательство его роли в поведении лис или именно в амигдале.

Также, стоит выделить модуль *blue4*, отрицательно коррелирующий с агрессивным фенотипом (r=-0.81). Этот модуль включает в себя 110 генов, связанных как с функциями в нервной системе, так и с более общими метаболическими или клеточными функциями. В качестве “хаб”-гена был идентифицирован дифференциально экспрессируемый ген *BST1* (Рисунок 2.3.5б), описанный ранее в разделе 2.2 как ADP-рибозилциклаза, генерирующая циклическую ADP-рибозу, участвующую в регуляции секреции окситоцина и в формировании социального узнавания и социального поведения у грызунов.

Среди генов, участвующих в передаче сигнала между нейронами и возбудимости можно выделить гены *ACHE*, *GRIK3*, *KCNH2* и *GNAL*. Ген *ACHE* кодирует ацетилхолинэстеразу, фермент, расщепляющий ацетилхолин и прекращающий его действие в холинергических синапсах [193]. Ген *GRIK3* кодирует субъединицу GluK3 ионотропного глутаматного рецептора каинатного типа, участвующего в возбуждающей нейротрансмиссии [194]. Ген *KCNH2* кодирует потенциал-зависимый калиевый канал Kv11.1 [195]. Ген *GNAL* кодирует G-белок Gαolf, участвующий в передаче сигнала от определенных рецепторов, особенно хорошо изученный в обонятельной системе и стриатуме [196].

Среди генов, участвующих в процессе формирования и перестройки нервных связей особенно релевантны гены *FGF22* и *RGMA*. В экспериментах на мышах ген *FGF22* способствует формированию возбуждающих пресинаптических окончаний в гиппокампе; нарушения этого механизма сопровождаются изменениями синаптической организации и восприимчивости к судорогам [197]. Ген *RGMA* участвует в направленном росте аксонов и формировании нервных связей. Его функции показаны в экспериментах на развивающемся мозге и в моделях гиппокампа [198]. Ген *SEMA3C* также относится к семафоринам – сигнальным молекулам, регулирующим направление роста аксонов; экспериментальная роль *SEMA3C* описана, например, при развитии глазодвигательных нервов [199].

К группе генов, участвующих в нейрональной дифференцировке и метаболизме нейромедиаторов, относятся гены *ALDH1A1* и *SOX2*. Ген *ALDH1A1* кодирует

альдегиддегидрогеназу, участвующую в метаболизме альдегидов и ретиноевой кислоты. В работе Kim et al. на мышах было показано, что продукт гена *ALDH1A1* участвует в альтернативном пути синтеза ГАМК в дофаминергических нейронах среднего мозга [200]. Однако более позднее исследование показало, что ГАМК-сигнализация в этих нейронах может обеспечиваться захватом ГАМК из внеклеточной среды, а не ее синтезом посредством *ALDH1A1 [201]*. Поэтому связь гена *ALDH1A1* с ГАМК-ергической передачей нельзя считать однозначно установленной. Ген *SOX2* кодирует транскрипционный фактор, необходимый для поддержания нейральных стволовых клеток; в экспериментах на мышах его принудительная экспрессия позволяла перепрограммировать астроциты в нейральные предшественники и нейроны [202].

3.3. Эпигеномный анализ

3.3.1. Первичная оценка данных секвенирования иммунопреципитированного хроматина

Для получения профилей гистоновых меток H3K4me3 и H3K27ac в клетках фронтальной коры мозга лисицы были использованы образцы тканей фронтальной коры головного мозга от 18 особей (по 6 особей для каждой фенотипической группы. Для каждой особи были приготовлены по две библиотеки – из ДНК, иммунопреципитированной с антителами к гистоновым меткам H3K4me3 и H3K27ac, соответственно. В качестве контроля, для одного образца (FBP85) также была приготовлена библиотека из неиммунопреципитированной ДНК. В результате глубокого секвенирования было получено в среднем по 31 млн прочтений для каждой из 36 библиотек из иммунопреципитированной ДНК и 50 млн прочтений для библиотеки из неиммунопреципитированной ДНК (Приложение 3). В среднем, 96,6% полученных прочтений было картировано на референсный геном лисицы. Анализ картированных прочтений для обеих гистоновых меток выявил их неравномерное распределение в геноме лисицы, характерное для данных ChIP-seq. Так, для всех образцов картированные прочтения сконцентрированы в меньшей доле геномных участков, чем в контрольной библиотеке, что говорит об обогащении конкретных участков генома прочтениями, а не о случайном их распределении (Рисунок 3.14).

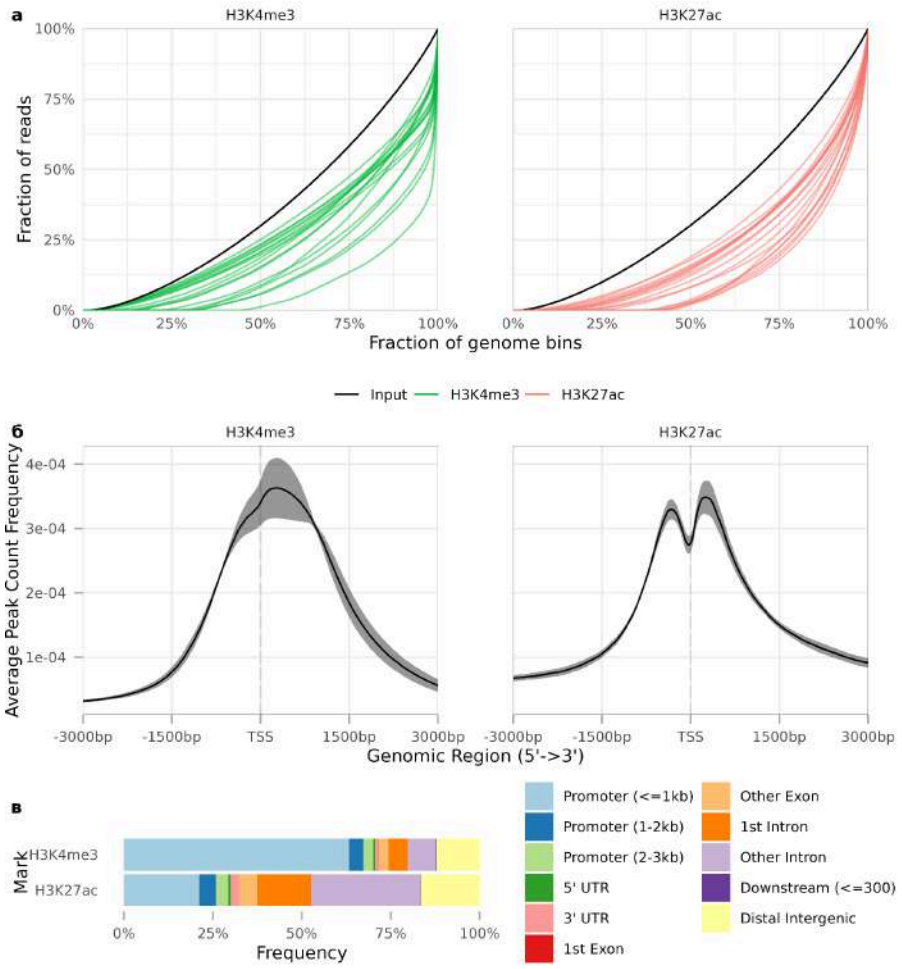


Рисунок 3.14 – Результаты качественной оценки данных секвенирования. а, распределение глубины покрытия в геноме; черным выделена кривая для контрольного образца неиммунопреципитированной ДНК, зеленым и красным цветами – кривые для образцов H3K4me3 и H3K27a, соответственно. б, распределение покрытия пиками обогащения в промоторных регионах генов (± 3000 пар оснований относительно сайта начала транскрипции). в, Распределение аннотации пиков H3K4me3 и H3K27ac относительно генов.

Для каждого образца были определены генетические регионы повышенной плотности прочтений (пики) по сравнению с контрольной библиотекой и геномным шумом. Для метки H3K4me3 было определено в среднем по 25 тыс. таких участков в каждом образце, а для метки H3K27ac – по 64 тыс. Распределение покрытия обнаруженными пиками в промоторных регионах (3000 пар оснований в обе стороны от точки начала транскрипции) свидетельствует о повышенной плотности прочтений в этих регионах генома по сравнению с другими регионами генома (Рисунок 3.14). Сравнение этих участков с генными локусами в геноме лисицы показало, что регионы, обогащенные меткой H3K4me3, и в самом деле расположены, в основном, рядом с промоторными участками: в среднем, 60% пиков находится на расстоянии не более 3000 пар оснований от точки начала транскрипции (Рисунок 3.14). Для метки H3K27ac, в среднем только

25% пиков находится рядом с промоторами, а большая часть пиков – в интронных или межгенных регионах генома (в среднем, 70%) (Рисунок 3.14). Эти наблюдения соответствуют ожиданиям: H3K4me3 маркирует в основном промоторные регионы, а H3K27ac – и промоторные, и энхансерные участки. Таким образом, впервые были получены эпигенетические профили и карты регуляторных участков клеток фронтальной коры мозга лисиц *Vulpes vulpes*, в том числе для животных, характеризующихся разным типом поведения.

3.3.2. Попарное сравнение профилей иммунопреципитированного хроматина с синхронным сигналом от обеих гистоновых меток

Так как обе исследованные эпигенетические метки (H3K4me3 и H3K27ac) маркируют промоторные регионы (H3K27ac маркирует и энхансерные регионы тоже), их совместный анализ позволяет достоверно выявить изменения в состоянии хроматина, связанного с активными промоторами генов. Мы провели поиск участков генома со статистически значимым дифференциальным обогащением как при сравнении агрессивных и неселектированных животных, так и при сравнении агрессивных животных с ручными ($p\text{-value} < 0.05$). При этом, учитывались только участки с одинаковым направлением изменения: плотность прочтений каждой из эпигенетических меток выше или ниже у агрессивных животных по сравнению с двумя другими группами. Таким образом всего было выявлено 478 пиков H3K4me3 и 4634 пиков H3K27ac (Рисунок 3.15).

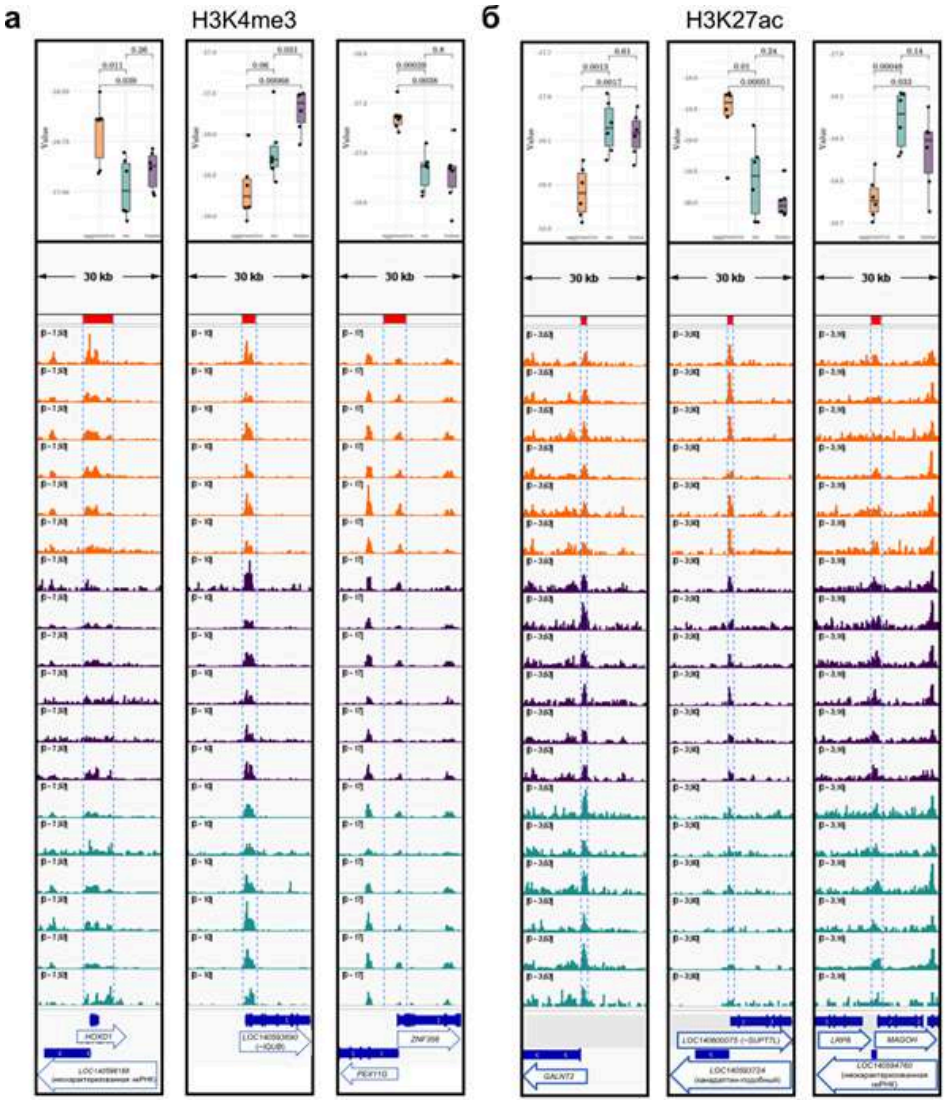


Рисунок 3.15 – Примеры эпигенетических измененией в клетках фронтальной коры у лисиц с агрессивным поведением для меток H3K4me3 (а) и H3K27ac (б) в генах *HOXD1*, *IQUB*, *PEX11G*, *GALNT2*, *SUPTL7*, *MAGOH*. На каждой панели вверху приведены графики нормализованных размеров пиков в трех поведенческих группах. Над скобками обозначены p-value для попарных сравнений. На каждой панели внизу приведены графики нормализованного сигнала меток H3K4me3 (а) и H3K27ac (б) в локусе пика (выделен пунктирными синими линиями), для которого наблюдается разница в эпигенетическом статусе между поведенческими группами. На всех панелях цвет обозначает поведенческую группу (оранжевый – агрессивные, бирюзовый – ручные, фиолетовый – неселектированные животные).

Из них 73 геномных локуса демонстрировали изменения одновременно для обеих меток в одном направлении у агрессивных животных по сравнению с ручными и неселектированными. Среди этих 73 локусов 48 соответствуют промоторам аннотированных генов. Из них 35 кодируют белки, а остальные являются генами некодирующей РНК и псевдогенами (Приложение И). При этом бóльшая часть совместно дифференцированных локусов (~59 %) характеризуется повышенным уровнем как H3K4me3, так и H3K27ac у

агрессивных животных (logFC > 0) по сравнению с ручными и неселектированными. Оставшаяся часть (~41 %) демонстрирует пониженный уровень обеих меток у агрессивных особей по сравнению с двумя другими группами. Среди всех локусов с повышенным уровнем одновременно двух меток H3K4me3 и H3K27ac выделяются несколько функциональных групп (Таблица 3.3).

Таблица 3.3 – Функциональная характеристика генов, в промоторе которых наблюдается пониженный или повышенный уровень триметилирования четвертого лизина и, одновременно, ацетилирования двадцать седьмого лизина гистона H3 у агрессивных животных по сравнению с ручными и неселектированными.

Ген	logFC	Продукт гена	Ссылка
Миелиновые и цитоскелетные белки			
KLK6	+0.41	калликреин-6, экспрессируемый олигодендроцитами	[203]
TUBB4A	+0.77	кодирующий мозгоспецифический бетта-тубулин класса A	[204]
CLDN11	+0.38	клаудин-11/олигодендроцит специфичный белок	[205]
Элементы транскрипционного регуляторного механизма			
*LOC140600075 (*гомолог *SUPT7L)*	+1.38	ядерная гамма-подобная субъединица комплекса STAGA 65	[206]
TAF5L	+0.29	субъединица комплекса TFIID/SAGA	[207]
HOXD1	+0.42	гомеобоксный транскрипционный фактор, участвует в дифференцировке клеток нервного гребня	[208]
ZNF358	+0.99	цинковый палец	[209]
Гены метаболизма			
CPOX	+0.30	копропорфириноген оксидаза в биосинтезе гема	[210]
PEX11G	+0.99	гамма-фактор 11 пероксисомального биогенеза	[211]
DDHD2	+0.29	фосфолипаза, участвующую в гидролизе липидов	[212]
Процессы сигналлинга и длинных некодирующих РНК			

Продолжение таблицы 3.3

COMMD9	+0.40	белок, содержащий домен COMM 9, участвует в регуляции сигналлинга NOTCH	[213]
LGI3	+0.29	белок из семейства LGI, который секретируется олигодендроцитами и участвует в кластеризации калиевых каналов Kv1 через взаимодействие с рецептором ADAM23	[214]
KCNIP2	−0.72	вспомогательная субъединица калиевых каналов Kv4 типа A, участвует в процессе возбудимости нейронов	[215]
RUNX1T1	−1.05	транскрипционный корепрессор ETO (Eight Twenty-One), участвующий в подавлении транскрипции	[216]
Элементы внеклеточного матрикса			
EGFL7	−0.65	содержащий EGF-подобный домен белок, участвующий в регуляции ангиогенеза и Notch сигналлинге	[217, 218]
ANGPT2	−0.32	ангиопоэтин-2, оказывает нейропротектирующий эффект	[219]
MFAP2	−0.97	микрофибрилл-ассоциированный белок 2, экспрессируется в возбуждающих и тормозных нейронах	[220]

Гены с пониженным уровнем одновременно двух эпигенетических меток кодируют, например, элементы внеклеточного матрикса и транскрипционные регуляторы . Среди всех этих локусов с пониженным уровнем одновременно H3K4me3 и H3K27ac, наибольшее снижение эпигенетического сигнала у агрессивных животных по сравнению с ручными и неселектированными наблюдалось в промоторе гена неохарактеризованной некодирующей РНК *LOC140594598* (logFC=-2.47). Среди продуктов генов с повышенным уровнем H3K4me3 и H3K27ac, можно выделить миелиновые и цитоскелетные белки (в частности, к этой группе относятся гены *KLK6* и *TUBB4A*, Рисунок 3.16), элементы транскрипционного регуляторного механизма, а также белки, связанные с процессами нейронального сигналинга.

Оба гена *KLK6* и *TUBB4A* экспрессируются в центральной нервной системе, а их продукты вовлечены в метаболизм структурных белков и поддержание целостности миелиновых оболочек нейронов [203, 204]. Нарушения миелинизации и цитоскелетной организации в префронтальной коре ассоциированы с изменениями скорости проведения

импульсов, что в свою очередь может влиять на поведенческие реакции и стресс-реактивность, тогда как повышенный эпигенетический сигнал в локусах *KLK6* и *TUBB4A* у агрессивных животных может указывать на активацию процессов, необходимых для компенсации нейрональных изменений, происходящих в мозге агрессивных животных.

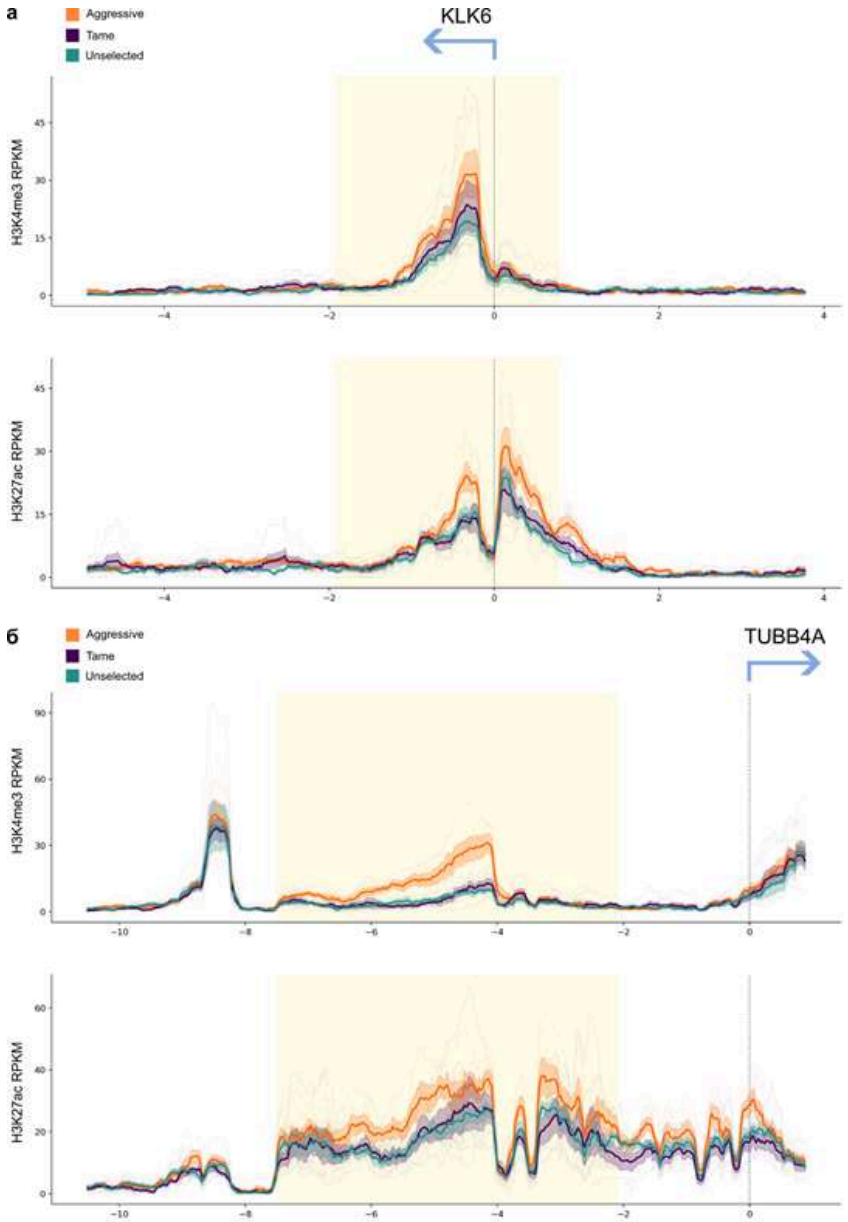


Рисунок 3.16 – Увеличение уровней H3K4me3 и H3K27a в районе промоторов генов *KLK6* (а) и *TUBB4A* (б) в клетках фронтальной коры у животных с агрессивным поведением по сравнению с ручными и неселектированными. Нулевая точка на оси абсцисс соответствует сайту начала транскрипции, направление транскрипции указано стрелкой. По оси ординат обозначена нормализованная глубина прочтений RPKM).

3.3.3. Попарное сравнение профилей иммунопреципитированного хроматина с сигналом от гистоновой метки H3K4me3

Помимо участков, в которых наблюдались согласованные изменения обеих гистоновых меток, мы выявили регионы, дифференциально регулируемые исключительно на уровне H3K4me3 ($p < 0,05$, $|logFC| > 0,05$, плотность прочтений выше или ниже у агрессивных особей по сравнению с обеими контрольными группами), тогда как уровень H3K27ac в этих локусах значимо не отличался между группами ($p > 0,05$). Всего было идентифицировано 298 таких локусов, из которых 65 участков соответствовали промоторам аннотированных генов, причем 39 из них кодируют белки. Гены, для которых мы выявили пониженный уровень H3K4me3 у агрессивных лис по сравнению с ручными и неселектированными, вовлечены в ключевые нейробиологические процессы (Таблица 3.4), такие как образование синапсов, возбудимость нейронов, метаболизм нейромедиаторов, гены, обеспечивающие структурную целостность клетки и матрикса, а также транскрипционную регуляцию.

Таблица 3.4 – Функциональная характеристика генов, в промоторе которых наблюдается пониженный или повышенный уровень триметилирования четвертого лизина гистона H3 у агрессивных животных по сравнению с ручными и неселектированными.

Ген	logFC	Продукт гена	Ссылка
Гены синаптической организации, адгезии и нейрональной возбудимости			
IGSF9	–0,84	белок иммуноглобулинового суперсемейства 9, участвующий в организации тормозных синапсов	[221]
DTNB	-0,915	белок дистробревин бета, локализующийся в постсинаптическом пространстве	[222]
KCNA5	-0,93	альфа-субъединица калиевого канала Kv1.5, определяющая возбудимость мембраны нейрона и параметры реполяризации потенциала действия	[223]
Гены метаболизма нейромедиаторов и клеточной сигнализации			
ABAT	−0.54	ГАМК-трансаминаза, катализирующая деградацию ГАМК	[224]
AGPAT3	−0,97	фермент 1-ацилглицерол-3-фосфат О-ацилтрансфераза 3, участвующая в биосинтезе фосфолипидов и ремоделировании мембран нейронов	[225]
Гены посттранскрипционной регуляции			

Продолжение таблицы 3.4

LOC112914685 (гомолог гена *AGO4*)	−0,9	ключевой белок RISC-комплекса, эффекторный компонент РНК-интерференции	[226]
NCBP3	−0,94	ядерный кэп-связывающий белок 3, участвующий в процессинге и экспорте мРНК из ядра в цитоплазму	[227]
RIOX2	+0.69	фермент рибосомная оксигеназа 2, участвующий в посттрансляционной модификации и процессинге рибосомной РНК, а также в стабилизации варианта гистона H2A.Z, ассоциирован с регуляцией транскрипционной активности, в частности экспрессии генов пролиферации и дифференцировки, что приводит к когнитивным нарушениям	[228, 229]
Гены, кодирующие структурные белки клетки и внеклеточного матрикса			
KRT8	+0.675	белок промежуточных филаментов, который обеспечивает структурную целостность эпителиальных клеток	[230]
ABI3BP	+0.605	связывающий белок семейства ABI 3, вовлеченный в регуляцию клеточной адгезии и ремоделирование внеклеточного матрикса	[231]

Кроме того, следует отдельно отметить среди выявленных локусов ген *LOC11291363* (гомолог гена *C4A)*, кодирующий компонент комплемента 4A. Белки классического пути комплемента экспрессируются в нейронах и играют важнейшую роль в элиминации избыточных синапсов в ходе развития (синаптический прунинг) [232], следовательно, обнаруженное нами снижение уровня H3K4me3 в промоторе *C4A* у агрессивных животных позволяет выдвинуть гипотезу о гипофункции системы комплемента в мозге агрессивных особей.

Среди генов с повышенным уровнем H3K4me3 без значимого изменения уровня H3K27ac у агрессивных животных по сравнению с ручными и неселектированными функциональный интерес представляет ген *PTPN20*, в промоторе которого у агрессивных особей наблюдается повышенный уровень H3K4me3 (logFC=+1.86). Кодируемая им тирозин-специфичная протеин-тирозинфосфатаза, обеспечивает структурную целостность систем удаления метаболитов мозга [233], что позволяет рассматривать данный локус как

потенциально вовлеченный в нейрометаболические процессы, которые могут быть ассоциированные с поведенческим фенотипом.

3.3.4. Попарное сравнение профилей иммунопреципитированного хроматина с сигналом от гистоновой метки H3K27ac

Наконец, для выявления локусов с дифференциальным уровнем эпигенетической метки H3K27ac, маркирующей активные энхансеры и промоторы, мы провели сравнительный анализ между агрессивными животными и двумя другими группами (ручными и неселекционированными). Критерием отбора были p-value<0.05, абсолютное значение logFC > 0.5, а также совпадающий знак logFC в обоих сравнениях (агрессивные особи против ручных и агрессивные против неселектированных). Поскольку гистоновая метка H3K27ac маркирует как промоторы, так и энхансеры, для выделения именно энхансерных участков, расположенных в промоторных регионах генов, мы рассматривали дифференциально обогащенные локусы H3K27ac, удалив регионы, перекрывающиеся с пиками H3K4me3. Всего в геноме было идентифицировано 3070 таких локусов, из которых 145 участков соответствуют промоторам аннотированных генов, из которых 73 кодируют белки. Выявленные таким образом гены принадлежат к различным функциональным группам, ключевые из которых представлены в Таблице 3.5.

Таблица 3.5 – Функциональная характеристика генов, в промоторе которых наблюдается пониженный или повышенный уровень ацетилирования двадцать седьмого лизина гистона H3 у агрессивных животных по сравнению с ручными и неселектированными.

Синаптическая пластичность и нейрональная сигнализация			
Ген	logFC	Продукт	Ссылка
CAMK2A	-1.22	Кальций/кальмодулин-зависимая протеинкиназа II альфа, которая является главным регулятором синаптической пластичности, долговременной потенциации, а также процессов обучения и памяти. снижение уровня этой киназы в префронтальной коре (ПФК) агрессивных животных согласуется с предыдущими данными транскриптомики лисиц, демонстрирующими различия в глутаматном пути	[234, 235]
CHRNA5	-1.30	Субъединица α5 нейронального никотинового ацетилхолинового рецептора, экспрессирующаяся во многих областях мозга, включая префронтальную кору. Варианты этого гена ассоциированы с агрессией и эмоциональной регуляцией	[236]
DUSP1	-1.00	Двойная специфическая фосфатаза 1 (MKP-1), негативный регулятор MAPK-каскада, инактивирующий сигнальные пути p38 и JNK. Стресс-индуцированная экспрессия *DUSP1* в гиппокампе связана с депрессивоподобным поведением	[237]
RGS10	-0.93	негативный регулятор G-белкового сигналинга 10, высокий уровень экспрессии в микроглии, подавляет провоспалительную сигнализацию. Снижение уровня RGS10 ассоциирована с усилением экспрессии цитокинов и нейродегенерацией	[238]
PTPRE	-2.50	Рецепторная тирозинпротеинфосфатаза эпсилон, высокий уровень экспрессии в мозге. Дефицит этого белка приводит к гипомиелинизации и повышению активности потенциал-зависимых K^+-каналов	[239]
AGER	-0.71	Рецептор конечных продуктов гликирования (RAGE), негативно регулирует долговременную синаптическую потенциацию	[240]
Сигнализация нейромедиаторов и гормонов			

Продолжение таблицы 3.5

PPARD	+0.85	Ядерный рецептор, активируемый пероксисомным пролифератором дельта, обладает нейропротекторным эффектом, а его активация подавляет депрессивоподобное поведение и стимулирует нейрогенез в гиппокампе	[86]
RARA	-1.46	Рецептор ретиноевой кислоты альфа, контролирует синаптическую пластичность через нетранскрипционные эффекты, влияя на трансляцию GluA1 (субъединицы AMPA-рецептора)	[241]
Транскрипционная и эпигенетическая регуляция			
ELK4	-0.91	ETS-транскрипционный фактор, член подсемейства TCF, экспрессируется в мозге и участвует в MAPK/ERK сигнальном пути. Подавляет активацию микроглии, ослабляя нейровоспаление и когнитивную дисфункцию	[242]
HMGN2	-0.89	Белок группы высокомобильности, связывающий нуклеосомы, высокий уровень экспрессии в нейральных стволовых/прогениторных клетках развивающегося и взрослого мозга. Потеря Hmgn2 в плюрипотентных клетках приводит к спонтанной нейрональной дифференцировке.	[243]
BACH1	-0.54	Транскрипционный гем-связывающий репрессор, регулирует окислительный стресс. Играет критическую роль в защите тканей от окислительного стресса в мозге	[244]
RECK	+1.09	Мембранно-связанный белок, регулятор гематоэнцефалического барьера через Wnt-сигналинг	[245]
Клеточная адгезия, внеклеточный матрикс и цитоскелет			
TNS4	-1.22	Тензин 4 (белок фокальных контактов)	[246]
SHKBP1	-1.34	SH3KBP1-связывающий белок 1, предотвращает деградацию рецептора эпидермального фактора роста (EGFR), участвующего в различных биологических процессах	[247]
LAMB2	-1.03	Субъединица бета 2 ламинина, локализуется в области гематоэнцефалического барьера	[248]

В частности, гены с пониженным у агрессивных особей уровнем ацетилирования в промоторных регионах, относятся к функциональным группам генов, связанных с обеспечением синаптической пластичности и нейрональной сигнализации (гены *CAMK2A*, *CHRNA5*), а также транскрипционной и эпигенетической регуляции (гены *ELK4*, *BACH1*).

3.3.5. Сопоставление с ранее опубликованными данными по экспрессии

Дополнительно, для сопоставления обнаруженных эпигенетических изменений уровню экспрессии соответствующих генов мы соотнесли наши результаты анализа гистоновых меток H3K4me3 и H3K27ac с опубликованными ранее результатами оценки уровня экспрессии генов в клетках префронтальной коры агрессивных и ручных лисиц [68]. Мы обнаружили, что для 5 белок-кодирующих генов (*ABI3BP*, *RARA*, *DUSP1*, *CD82*, *PTGFRN*) направление изменений триметилирования четвертого лизина или ацетилирования двадцать седьмого лизина гистона H3 совпало с направлением изменений экспрессии мРНК у агрессивных животных по сравнению с ручными (Таблица 3.6).

Таблица 3.6 – Сопоставление результатов дифференциального обогащения с ранее опубликованными данными дифференциальной экспрессии. ChIP logFC – сравнение агрессивных особей с ручными и неселектированными на основе нашего исследования, RNA logFC – сравнение ручных особей с агрессивными по ранее опубликованным данным.

ген	тип гена	цепь	ChIP logFC	гистоновая метка	RNA logFC
ABI3BP	белок кодирующий	-	0,605	H3K4me3	-1,41
RARA	белок кодирующий	-	-1,46	H3K27ac	0.54
DUSP1	белок кодирующий	+	-0,995	H3K27ac	0.66
CD82	белок кодирующий	+	-0,86	H3K27ac	0.79
PTGFRN	белок кодирующий	-	-1,525	H3K27ac	0,46

Эти гены представляют различные функциональные категории, потенциально значимые для поведенческих различий. Так, например, ранее было показана ассоциация вариантов в гене *ABI3BP* с суицидальным поведением у людей [249], а уровень метилирования ДНК этого гена

статистически значимо различается в группах пациентов с суицидальными мыслями и без них [250], что в целом подчеркивает важную роль данного гена в формировании определенных типов поведения. Ген *DUSP1* является ключевым негативным регулятором сигнального каскада MAPK, который играет важнейшую роль в механизмах поддержания нейрональной пластичности и функционировании мозга. Нарушение его экспрессии выявлено у пациентов с депрессивным поведением, а также на животных моделях депрессий (грызунах) [237].

Кроме того, мы проанализировали промотор гена *MAOA*. Этот ген кодирует фермент моноаминоксидазу А, которая катализирует деградацию ключевых нейромедиаторов, а низкая активность *MAOA*, связанная с мутациями в промоторной области или гиперметилированием, ассоциирована с повышенной агрессивностью и импульсивностью как у животных, так и у человека [251–253]. При сравнении агрессивных особей с неселектированными мы не выявили статистически значимых различий в уровнях H3K4me3 или H3K27ac в промоторе гена MAOA (p-value > 0,05). Однако, мы обнаружили статистически значимое снижение уровней обеих меток у агрессивных особей по сравнению с ручными (p-value = 0,03).

Таким образом, в настоящей работе мы впервые применили технологию иммунопреципитации хроматина с последующим секвенированием для картирования регуляторных элементов в геноме лисицы – вида, для которого до сих пор отсутствовала какая-либо аннотация регуляторных элементов. Мы показали, что полученные для лисицы эпигенетические профили меток H3K4me3 и H3K27ac колокализуются с областями, соответствующими потенциальным координатам промоторов генов. Это позволило не только впервые идентифицировать регуляторные элементы в геноме лисицы, но и выявить эпигенетические различия между группами лисиц с разными типами поведения. При сравнении эпигенетических профилей агрессивных лисиц с ручными и неселектированными нами были выявлены эпигенетические изменения в генах трех функциональных кластеров, вовлеченных в процессы развития и функционирование нервной системы: миелинизации и олигодендроцитов (*TUBB4A*, *CLDN11*, *KLK6*, *LGI3*), нейрогенеза (*NEUROD2*, *ASCL1*), а также синаптической и дофаминергической сигнализации (*CAMK2A*, *DRD1*, *AXIN2*). Кроме того, для генов *ABI3BP*, *RARA*, *DUSP1*, *CD82* и *PTGFRN* выявлена согласованность направлений эпигенетических и транскрипционных изменений.

Следует отметить, что выявленное нами снижение эпигенетического сигнала в некоторых маркерных участках (*HMGN2*, *KCNIP2*, *KCNA5*, *MFAP2*) не обязательно означает подавление экспрессии данных генов, а скорее отражает сужение или ослабление активного хроматинового домена в ряде нейрональных генов у агрессивных лисиц. Напротив, ручные лисицы, по-видимому, приобрели в ходе одомашнивания более широкие, доступные для транскрипции хроматиновые домены, что согласуется с моделью «широкого домена» регуляции

генов, кодирующих критически важные для базового метаболизма, поддержания жизнедеятельности и клеточной идентичности белки [103, 254].

3.4. Интегративный анализ

В предыдущих разделах настоящей диссертации были получены и описаны различия между исследуемыми линиями на разных генетических уровнях (популяционная структура, однонуклеотидные варианты (3.1), экспрессия мРНК (3.2), активность регуляторных элементов (3.3)). Далее было решено сосредоточиться на выявлении корреляции между этими уровнями (3.4.1) и ее интерпретацией (3.4.2).

3.4.1.1. Специфичные для агрессивной линии ОНВ, совпадают с участками регуляторного хроматина с дифференциальным обогащением

Из 3 582 ОНВ, специфичных для агрессивной линии, подавляющее большинство расположено в некодирующих областях: 49% – в интронах, 39% – в межгенных областях, 11% – в генах нкРНК; миссенс-вариантами являются лишь 10. Это позволяет предположить, что отбор преимущественно действовал через регуляторную изменчивость. Двенадцать вариантов непосредственно подтверждают эту гипотезу: они совпадают с пиками гистоновых меток, обогащение которыми также специфично для агрессивной линии по сравнению с ручной и неселектированной (Приложение К).

В одном из выявленных в разделе 3.1.2. мегарегионе на хромосоме 1 (13–18 Мб) расположен кластер генов семейства калликреинов. В его пределах один ОНВ, специфичный для агрессивной линии (NC_132780.1:16,928,678 C>T; частота 0.80 в агрессивной линии, 0.067 в ручной и 0.00 в контрольной линии), находится внутри пика, расположенного вблизи TSS (16,928,353–16,931,054). Этот пик охватывает промотор и первый интрон гена *KLK6*. Ген *KLK6* расположен на минус-цепи, поэтому пик включает область непосредственно перед геном и прилегающий к ней участок его 5'-конца. Обогащение обоими типами меток в этом пике статистически различается для агрессивной (A) линии в обоих сравнениях (с ручной (T) и с неселектированной (N)): logFC для H3K4me3 составляет +0.38 (AvsT) и +0.52 (AvsN), а для H3K27ac – +0.41 (AvsT) и +0.31 (AvsN) (Рисунок 3.17).

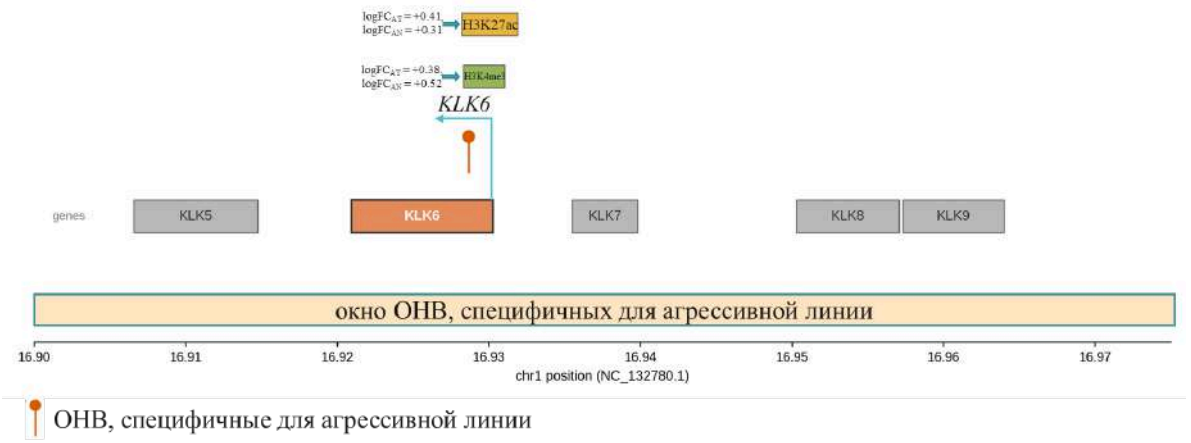


Рисунок 3.17 – Схематическое изображение специфичных для агрессивной линии: OHB в промоторе гена *KLK6*; локуса дифференцированного обогащения по гистоновым меткам H3K4me3 и H3K27ac; окна из 76 OHB. Локусы обогащения показаны желтым (H3K27ac) и зеленым (H3K4me3) цветами.

Одновременное увеличение H3K4me3 и H3K27ac в промоторе представляет собой характерную хроматиновую сигнатуру активированного гена, находящегося в состоянии транскрипционной готовности. При этом сам *KLK6* не демонстрирует дифференциальной экспрессии в ткани в целом: $|\log FC| < 0.02$, $padj > 0.6$ во всех четырех сравнениях.

В пределах тела гена *HIF3A* находятся пять OHB, специфичных для агрессивной линии (13,333,670–13,343,897 на NC_132780.1; частоты 0.80–0.92 в агрессивной линии и 0.00–0.08 в обеих контрольных линиях). Один из них совпадает с пиком H3K27ac, для которого выявлено дифференциальное обогащение с отрицательным logFC в обоих сравнениях AvsT и AvsN (-0.72 (AvsT) и -0.71 (AvsN)) (Рисунок 3.18).

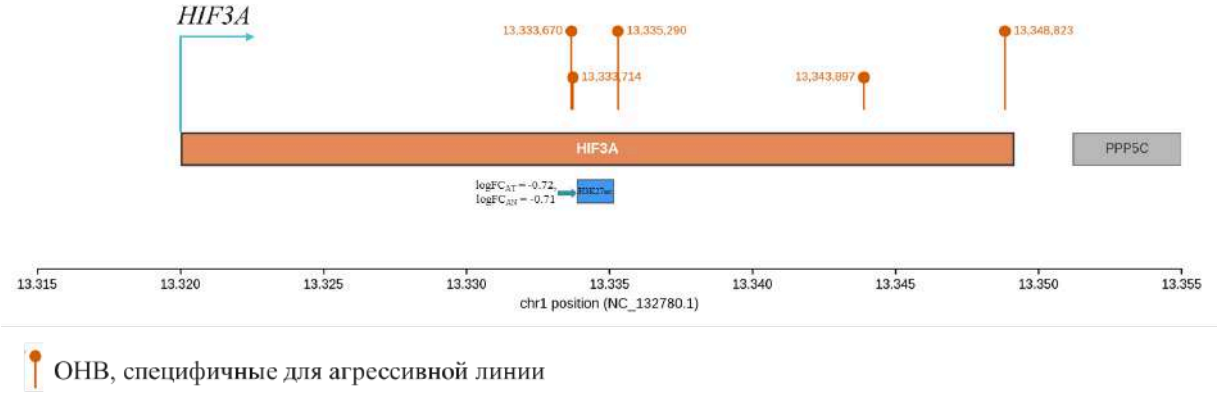


Рисунок 3.18 – Схематическое изображение специфичных для агрессивной линии: OHB в теле гена *HIF3A*; локуса дифференцированного обогащения по гистоновой метке H3K27ac. Локус обогащения показан синим (H3K27ac) цветом.

Дифференциальная экспрессия *HIF3A* в ткани в целом отсутствует (padj > 0.86 во всех сравнениях).

Таким образом, оба локуса демонстрируют одну и ту же закономерность: аллель, характерный для селектированной на агрессивное поведение линии, находится в активном регуляторном хроматине, состояние обогащения которого различается у агрессивных и неселектированных животных, тогда как уровень мРНК в ткани в целом остается неизменным. В данном случае было выдвинуто предположение, что регуляторный вариант действует специфично для определенного типа клеток или в зависимости от контекста, и его эффект не выявляется при анализе РНК из целой ткани.

3.4.1.2. Дифференциально экспрессирующиеся гены внутри окон OHB

Пересечение 242 окон OHB (57.5 Mb) с координатами генов и списками дифференциально экспрессирующихся генов выявило 23 дифференциально экспрессируемых гена, расположенных внутри окон, специфичных для агрессивной линии, среди аннотированных: гены *SIGLECL1, HAS1, BST1, RGS9BP, DHDH, CTU1, TNNI3, MFSD8, FANCI, NLRP12, CCDC192, ATP2C2, RLBP1, RBM20* и *AFAP1L2*. Кроме того, выявлено восемь некодирующих генов, представленных LOC-моделями. Среди них – ген *LOC112912849*,

происходящий от транспозируемого элемента и демонстрирующий наиболее выраженное совпадение генотипических и экспрессионных различий во всем наборе данных: повышенная экспрессия в обеих тканях, специфичность для агрессивной линии и 21 ОНВ в соответствующем окне.

Два гена из этого списка объединяют наиболее убедительные доказательства.

Так, ген *MFSD8* имеет наибольшее количество ОНВ в пределах тела гена среди всех генов в общегеномном сравнении: 52 ОНВ, специфичных для агрессивной линии, все расположены в интронах. В миндалевидном теле агрессивных животных экспрессия гена *MFSD8* повышена (AvsT: logFC +0.46, padj=0.0037; AvsN: logFC +0.68, padj=5.6e-5). Кроме того, мегарегион хромосомы 4 (13.2–15.8 Mb), в котором находится ген *MFSD8*, – единственный мегарегион ОНВ со статистически значимым по результатам перестановочного теста обогащением DE-генами: наблюдался 1 DE-ген при ожидаемом значении 0.034, p = 0.033.

Ген *BST1* расположен в регионе, содержащем восемь ОНВ, но сам не несет кодирующих вариантов. У агрессивных животных экспрессия гена *BST1* снижена во фронтальной коре (AvsT: logFC −0.82, padj=3.3e-5; AvsN: logFC −1.05, padj=1.1e-7), а также в миндалевидном теле (AvsN: logFC −0.39, padj=0.047).

3.4.1.3. Взаимно противоположные изменения экспрессии в локусе *CTXN3–CCDC192*

На хромосоме 12 окно ОНВ, содержащее ген *CCDC192* (57,375,237–57,764,496; девять ОНВ, все расположены в интронах гена *CCDC192*), заканчивается за 2.6 Кб до 3’ конца соседнего гена *CTXN3* (Рисунок 3.19).

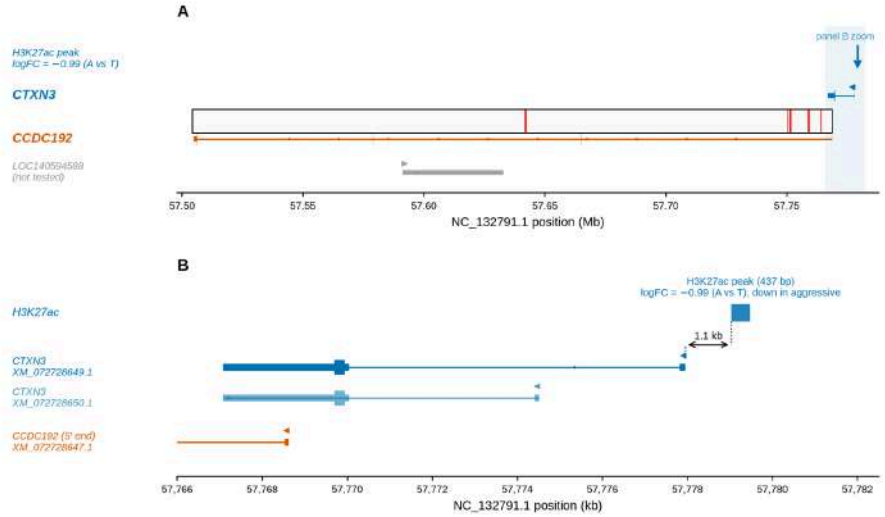


Рисунок 3.19 – Схематичное изображение генов *CCDC192* и *CTXN3,* ОНВ в регионе гена *CCDC192* и пика дифференциального ацетилирования.

Во фронтальной коре агрессивных животных экспрессия этих двух генов изменяется в противоположных направлениях: экспрессия гена *CCDC192* значительно повышена (AvsT: logFC +4.00, padj=2.1e-6; AvsN: logFC +1.11, padj=0.017), тогда как экспрессия гена *CTXN3* снижена (AvsT: logFC −0.86, padj=7.1e-11; AvsN logFC −0.73, padj=2.5e-7). На расстоянии 1Кб от TSS гена *CTXN3* расположен пик H3K27ac с дифференциальным обогащением (AvsT: logFC -0,82; AvsN: logFC -1,17). Это согласуется с возможным существованием общего регуляторного элемента, который совместно регулирует два нейрональных гена с противоположной направленностью транскрипции либо обеспечивает конкурирующую регуляцию между ними.

3.4.1.4. Модули коэкспрессии объединяют гены-кандидаты в сети, коррелирующие с агрессивным поведением

Анализ взвешенных сетей коэкспрессии генов (WGCNA), выполненный для обеих тканей, выявил модули, коррелирующие с агрессивным поведением. Их хаб-гены совпадают с генами, выявленными при анализе ОНВ и экспрессии.

Во фронтальной коре выявлен модуль с отрицательной корреляцией с агрессией (r = −0.46), хаб-геном которого является ген *SHTN1*. В теле гена *SHTN1* расположено 20 ОНВ, специфичных для агрессивной линии, включая миссенс-вариант V524I (K524I в нумерации аминокислот человеческого белка; SIFT deleterious 0.000; лизин консервативен у всех таксонов,

включенных в выравнивание), описанный в разделе 3.1.4. При этом дифференциальная экспрессия гена *SHTN1* в ткани в целом не выявлена.

В миндалевидном теле выявлен модуль с положительной корреляцией с агрессией (r = +0.74), хаб-геном которого является ген *PPP1R17*. В состав модуля также входят гены *HAS1*, *NLRP12*, *MFSD8* и *DRD4*. Три из пяти генов этого модуля – *MFSD8*, *HAS1* и *NLRP12* – сами являются дифференциально экспрессируемыми генами, расположенными внутри окон ОНВ. Таким образом, этот модуль независимо связывает локусы, содержащие варианты, с различиями в экспрессии. Также в миндалевидном теле был выявлен модуль с отрицательной корреляцией с агрессией (r = −0.81), представляющий собой наиболее сильную корреляцию между модулем и исследуемым фенотипом. Его хаб-ген – *BST1*, экспрессия которого снижена в лобной коре агрессивных животных и который расположен внутри окна ОНВ.

3.4.2. Регуляторная архитектура поведенческого отбора

Интеграция трех уровней данных – ОНВ, специфичных для агрессивной линии, различий в обогащении гистоновых меток H3K4me3/H3K27ac, дифференциальной экспрессии в двух тканях мозга (и модулей коэкспрессии) – указывает на регуляторную модель поведенческой дивергенции. Из 3582 ОНВ только 10 затрагивают кодирующую последовательность; оба варианта, расположенные в участках активного хроматина с дифференциальным обогащением (промотор гена *KLK6* и энхансер в пределах тела гена *HIF3A*), являются некодирующими. Это согласуется с результатами эксперимента по селекции крыс, в котором локусы, связанные с ручным поведением, картировались в регуляторных областях, а контроль признака имел полигенную основу [34], а также с исследованием генома лис, где поведенческие QTL были обогащены регуляторными генами и генами нейроразвития, а не участками с фиксированными изменениями кодирующей последовательности [67].

Статистически значимое обогащение дифференциально экспрессируемыми генами (ДЭГ) среди генов, содержавших ОНВ (в 5.8 раза в миндалине и в 17.1 раза для ДЭГ, специфичных для агрессивной линии, в лобной коре; BH p = 0.0057), а также корреляция между количеством ОНВ и выраженностью различий в экспрессии количественно характеризуют связь генотипа и экспрессии на уровне всего генома.

3.4.2.1. Протеолитическое ремоделирование нейронных цепей

ОНВ в промоторе гена *KLK6* представляет собой наиболее интересного кандидата на роль регуляторного варианта из-за специфичности для агрессивной линии и расположения в промоторном регионе, в котором обогащение обеими метками – H3K4me3 и H3K27ac – увеличивается у агрессивных животных в обоих сравнениях. Ген *KLK6* кодирует сериновую протеазу, экспрессирующуюся в ЦНС и секретируемую олигодендроцитами и нейронами. Этот фермент расщепляет белковые компоненты миелина и внеклеточного матрикса и связан с рассеянным склерозом, болезнями Альцгеймера и Паркинсона, а также психическими расстройствами [255]. Белок *KLK6*, секретируемый олигодендроцитами, способствует прогрессированию экспериментального аутоиммунного энцефаломиелита [203], а генетическое ингибирование KLK6 или его специфическое подавление в олигодендроцитах ускоряет миелинизацию в ходе развития. При этом при демиелинизирующих заболеваниях ЦНС тканеспецифичная экспрессия гена *KLK6* регулируется различными промоторами [256].

Увеличение доли активного промоторного хроматина в локусе гена *KLK6* у агрессивных животных позволило нам выдвинуть гипотезу: протеазозависимое ремоделирование миелина и внеклеточного окружения различается между исследуемыми линиями. Отсутствие дифференциальной экспрессии в ткани в целом ожидаемо, если вариант действует в ограниченной популяции клеток: олигодендроциты составляют лишь небольшую долю клеток, представленных в транскриптоме тканевого гомогената.

3.4.2.2. Гипоксическая и метаболическая сигнализация

ОНВ в пределах тела гена *HIF3A* находится также в локусе со снижением обогащения H3K27ac у агрессивных животных. Белок HIF3A является ингибирующим представителем семейства транскрипционных факторов, участвующих в ответе на гипоксию. Метилирование и экспрессия гена *HIF3A* служат специфичными для отдельных жировых депо маркерами дисфункции жировой ткани [257] и вариаций массы тела [258]. Варианты Hif3a также различают мышиные селектированные линии, характеризующиеся жировым и худощавым фенотипами [259].

Наличие варианта, характерного для определенной селектированной линии, в энхансере гена *HIF3A*, отмеченном гистоновой меткой H3K27ac, может связывать поведенческий отбор с

метаболическими процессами и реакцией на гипоксию. Эту взаимосвязь, возможно, имеет смысл рассматривать в контексте различий между агрессивной и ручной линией в контексте энергозатрат и физиологии стресса. Однако отсутствие соответствующих изменений экспрессии в ткани в целом означает, что пока это наблюдение ограничивается уровнем хроматина.

3.4.2.3. Ось «лизосомы – нейровоспаление» в миндалевидном теле: гены *MFSD8, NLRP12, BST1, PPP1R17* и *DRD4*

Положительно коррелирующий с агрессией (r = +0.74) модуль коэкспрессии генов в миндалевидном теле объединяет пять генов с согласующейся функциональной направленностью. Ген *MFSD8*/*CLN7* кодирует лизосомальный мембранный транспортер; мутации этого гена вызывают нейрональный цероидный липофусциноз (болезнь Баттена) [173, 260] – нейродегенеративное заболевание, сопровождающееся вторичным нейровоспалением. Экспрессия гена *MFSD8* повышена в миндалевидном теле агрессивных животных, а сам ген несет наибольшее в геноме количество ОНВ в пределах тела гена – 52 варианта.

Белок NLRP12 с инфламмасомой образует белковый комплекс, функция которого зависит от контекста, а также представляет собой отрицательный регулятор сигнального пути NF-kappaB [261, 262]. В экспериментальных моделях аутоиммунного энцефаломиелита NLRP12 проявляет противовоспалительные свойства [170]. В лобной коре агрессивных животных экспрессия гена *NLRP12* повышена; ген расположен внутри окна ОНВ.

Ген *BST1*/*CD157* кодирует ADP-рибозилциклазу и представляет собой локус GWAS, неоднократно связывавшийся с болезнью Паркинсона [263, 264]. Полиморфизмы в гене *BST1* также ассоциированы с нейропсихиатрическими фенотипами. У мышей с нокаутом *CD157* наблюдаются поведенческие признаки, напоминающие тревожность и депрессию [265]. Ген *BST1* является хаб-геном модуля с наиболее сильной корреляцией с агрессивным поведением (r = −0.81), при этом его экспрессия в лобной коре агрессивных животных снижена.

Ген *PPP1R17* служит маркером отдельной популяции нейронов гипоталамуса, участвующих в регуляции потребления пищи и коммуникации между гипоталамусом и жировой тканью [266, 267], и влияет на сигнальные процессы с участием PP1 и актина [268].

Ген *DRD4* – один из классических генов поведенческой генетики, кодирующий дофаминовый рецептор. Взаимодействие DRD4 с факторами среды позволяет прогнозировать экстернализующее поведение у детей [183]. Гены дофаминовых рецепторов демонстрируют

признаки отбора при одомашнивании бойцовых петухов [269], а среди генов-кандидатов, связанных с агрессивностью собак, присутствуют гены дофаминергической системы [270, 271].

В совокупности этот модуль позволяет предположить, что функциональная активность лизосомально-аутофагической системы миндалины, нейроиммунная сигнализация и дофаминергическая чувствительность сопряжены с агрессивным поведением.

3.4.2.4. Внеклеточный матрикс и перинейрональные сети

Ген *HAS1* участвует в синтезе гиалуронана – структурной основы перинейрональных сетей (PNN) [272]. Известно, что PNN в миндалевидном теле регулируют пластичность механизмов формирования страха: их образование защищает воспоминания о страхе от угасания и закрывает критический период, в течение которого травматические воспоминания могут быть устранены [273]. Воздействие на внеклеточный матрикс также изменяет извлечение воспоминаний о страхе [274].

Экспрессия гена *HAS1* повышена в миндалине агрессивных животных по сравнению с ручными и неселектированными. Ген расположен внутри окна ОНВ и входит в модуль с корреляцией r = +0.74. В качестве гипотезы можно предположить, что отбор на агрессивность стабилизирует структуру внеклеточного матрикса и PNN в миндалине, тем самым закрепляя защитные поведенческие реакции, тогда как у ручных животных сохраняется более пластичное состояние PNN.

3.4.2.5. Нейрональная полярность и кортикальная идентичность

Хаб-геном модуля коэкспрессии фронтальной коры (r = −0.46) является *SHTN1*/shootin1 – ключевой организатор нейрональной поляризации и инициации роста аксонов [108, 110], необходимый для роста нейритов [107]. Ген *SHTN1* – единственный ген, сочетающий три характеристики: (1) кодирующий вариант, который по прогнозу SIFT оказывает повреждающее воздействие и затрагивает полностью консервативный аминокислотный остаток (V524I); (2) статус хаб-гена в модуле, коррелирующем с агрессивностью; (3) расположение в мегарегионе ОНВ. При этом изменения экспрессии в исследуемых тканях отсутствуют. Такой профиль позволяет предположить действие варианта на изоформы, специфичные для определенного типа клеток, или на регуляцию белка, а не на общее количество транскрипта.

В соседнем локусе *CTXN3–CCDC192* взаимно противоположные изменения генной экспрессии – повышение *CCDC192* и снижение *CTXN3*, оба специфичные для агрессивной линии, – в сочетании с пиком H3K27ac перед геном *CTXN3* и окончанием окна OHB за 2.6 kb до гена *CTXN3* указывают на возможную дивергентную регуляцию со стороны общего регуляторного элемента.

Ген *CTXN3* кодирует кортексин – специфичный для нейронов мембранный белок коры головного мозга [142], впервые клонированный как консервативный для мозга ген [141] и связанный с шизофренией у человека [140]. Ген *CCDC192* кодирует малоизученный белок с coiled-coil-доменом о функции которого нет достаточной информации [143].

3.4.2.6. Связь с геномикой одомашнивания лис

Представленные нами результаты расширяют программу исследований геномики селекции лис в трех направлениях.

Во-первых, они воспроизводят с более высоким разрешением основной вывод, полученный в исследованиях генома лис: поведенческие QTL обогащены генами, связанными с синаптической функцией, неврологическими заболеваниями и поведением . Кроме того, результаты связывают этот вывод с исследованиями регуляторных полигенных локусов ручного поведения у крыс [80, 81] и транскрипционным ответом на селекцию в тех же линиях лис [83, 88].

Во-вторых, впервые представлен уровень анализа хроматина: OHB, специфичные для агрессивной линии, расположены в участках с дифференциальным обогащением активных гистоновых меток в локусах генов *KLK6* и *HIF3A*. Это обеспечивает первые прямые связи между вариантами и энхансерами в данной модели.

В-третьих, модули коэкспрессии WGCNA позволяют перейти от списков генов к сетям, причем хаб-гены *SHTN1*, *PPP1R17* и *BST1* – именно те гены, для которых имеются независимые свидетельства наличия вариантов или различий в экспрессии. Такой результат позволяет предположить, что структура коэкспрессии отражает отобранный фенотип, а не технические ковариаты.

ЗАКЛЮЧЕНИЕ

Данная работа посвящена исследованию молекулярных механизмов агрессивного поведения на модели лисицы *Vulpes vulpes*. Впервые для данной модели данные получены и проанализированы одновременно на геномном, транскриптомном и эпигеномном уровнях.

На основе анализа геномных последовательностей лисиц трех линий (селектированные на агрессивное или ручное поведение по отношению к человеку и неселектированные) был выявлен общий гаплотип, специфичный для животных агрессивной линии, включающий локус гена *GRIN2B* и хромосомный регион, содержащий гены *ATF7IP*/*PLBD1*. По результатам анализа выявлено 3582 однонуклеотидных варианта, специфичных для особей агрессивной линии, на основе информации о консервативности, предсказаниях алгоритмов и структурных характеристиках, отобраны наиболее вероятные гены кандидаты, варианты в которых приводят к изменению белковых последовательностей и потенциально могут быть связаны с агрессивным поведением: *ACAN*, *SHTN1* и *NOL8*.

В результате анализ уровня экспрессии генов во фронтальной коре и миндалевидном теле выявлены гены и модули коэкспрессии генов, уровень экспрессии которых различается у животных агрессивной и ручной и неселектированной линий. Выявленные гены с дифференциальной экспрессией участвуют в лизосомных процессах, процессах связанных с организацией синаптической архитектуры и процессах нейровоспаления, что указывает на их роль в формировании агрессивного поведения.

По результатам анализа эпигенетических маркеров (гистоновых меток H3K4me3 и H3K27ac) сформирована карта регуляторных элементов (промоторы, энхансеры) фронтальной коры лисицы *Vulpes vulpes*. Сравнительный анализ локусов со статистически значимыми различиями эпигенетических профилей в разных линиях выявил различия в эпигенетических профилях. При сравнении агрессивных лисиц с ручными и неселектированными животными обнаружено 73 локуса с согласованными изменениями обеих гистоновых меток (H3K4me3 и H3K27ac), причем примерно в 59% из них уровень меток повышался, а в 41% – снижался. В локусах генов, связанных с формированием и поддержанием миелиновой оболочки и цитоскелетом выявлен повышенный эпигенетический сигнал у агрессивных животных, указывающий на повышенную активность этих генов. Пониженный эпигенетический сигнал, указывающий на снижение активности, выявлен в генах, связанных с внеклеточным матриксом и регуляцией транскрипции. Эти данные свидетельствуют о разнонаправленной эпигенетической регуляции при формировании агрессивного поведения.

Интеграция трех уровней данных позволила объединить три независимых перечня генов в единую модель: ОНВ, специфичные для агрессивной линии, преимущественно расположены в некодирующих областях и статистически связаны с различиями в экспрессии генов. Гены, содержащие ОНВ, обогащены среди дифференциально экспрессируемых генов. Мегарегион на chr4 (содержащий ген *MFSD8*) – единственный, для которого выявлено статистически значимое по результатам перестановочного теста обогащение генами с дифференциальной экспрессией ($p = 0,033$). Обнаружена корреляция между генетическими вариантами и состоянием хроматина: ОНВ в промоторе гена *KLK6* расположен в локусе с повышенным обогащением гистоновых меток H3K4me3 и H3K27ac в агрессивной линии, а ОНВ в теле гена *HIF3A* – в локусе со сниженным обогащением H3K27ac. При этом в обоих случаях изменения экспрессии на уровне ткани в целом отсутствуют, что соответствует паттерну регуляторной изменчивости, специфичной для отдельных типов клеток. Хаб-гены WGCNA (*SHTN1*, *PPP1R17*, *BST1*) совпадают с генами, содержащими ОНВ или дифференциально экспрессируемыми.

ВЫВОДЫ

1. Выявлены геномные особенности агрессивной линии лисиц (*Vulpes vulpes*), включая 3582 агрессивно-специфичных нуклеотидных варианта (большинство расположено в некодирующих областях), объединенных в 242 геномных окна и три крупных мегарегиона, а также 10 миссенс-вариантов.
2. Установлены различия уровня экспрессии генов во фронтальной коре и миндалевидном теле между животными агрессивной, ручной и неселектированной линий и модули коэкспрессии, коррелирующие с поведенческим фенотипом, связанные с процессами синаптической пластичности, нейрональной сигнализации, нейровоспаления и клеточного метаболизма.
3. Впервые охарактеризованы эпигенетические профили H3K4me3 и H3K27ac во фронтальной коре лисиц и выявлены различия состояния активного хроматина между линиями. Изменения затрагивают гены, связанные с миелинизацией, нейрогенезом, синаптической и дофаминергической сигнализацией, причем значительная часть различий локализована в энхансерных областях.
4. На основе интеграции геномных, транскриптомных и эпигенетических данных были определены функциональные особенности генов *ACAN*, *SHTN1*, *MFSD8*, *BST1* и регуляторной области в локусе *KLK6*, характерные для лисиц с агрессивным поведением.
5. Выявленные отличия, специфичные для агрессивной линии, согласовываются с ранее установленной полигенной моделью регуляции поведения, а геномная локализация показывает существенный вклад некодирующих регуляторных вариантов и изменения хроматинового состояния.

Список литературы

1. *Miles D.R.*, *Carey G.* Genetic and environmental architecture on human aggression. *J Pers Soc Psychol*. 1997;72(1):207-217. doi 10.1037/0022-3514.72.1.207

2. *Rhee S.H.*, *Waldman I.D.* Genetic and environmental influences on antisocial behavior: a meta-analysis of twin and adoption studies. *Psychol Bull*. 2002;128(3):490-529. doi 10.1037/0033-2909.128.3.490

3. *Tuvblad C.*, *Baker L.A.* Human aggression across the lifespan. In: 2011:171-214. doi 10.1016/B978-0-12-380858-5.00007-1

4. *Porsch R.M., Middeldorp C.M., Cherny S.S., et al.* Longitudinal heritability of childhood aggression. *Am J Med Genet B Neuropsychiatr Genet*. 2016;171(5):697-707. doi 10.1002/ajmg.b.32420

5. *Poulin F.*, *Boivin M.* Reactive and proactive aggression: evidence of a two-factor model. *Psychol Assess*. 2000;12(2):115-122. doi 10.1037/1040-3590.12.2.115

6. *Brendgen M., Vitaro F., Boivin M., et al.* Examining genetic and environmental effects on reactive versus proactive aggression. *Dev Psychol*. 2006;42(6):1299-1312. doi 10.1037/0012-1649.42.6.1299

7. *Baker L.A.*, *Raine A.*, *Liu J.*, *Jacobson K.C.* Differential genetic and environmental influences on reactive and proactive aggression in children. *J Abnorm Child Psychol*. 2008;36(8):1265-1278. doi 10.1007/s10802-008-9249-1

8. *Burt S.A.* Are there meaningful etiological differences within antisocial behavior? results of a meta-analysis. *Clin Psychol Rev*. 2009;29(2):163-178. doi 10.1016/j.cpr.2008.12.004

9. *Burt S.A.* Do etiological influences on aggression overlap with those on rule breaking? a meta-analysis. *Psychol Med*. 2013;43(9):1801-1812. doi 10.1017/S0033291712001894

10. *Fernàndez‐Castillo N.*, *Cormand B.* Aggressive behavior in humans: genes and pathways identified through association studies. *Am J Med Genet B Neuropsychiatr Genet*. 2016;171(5):676-696. doi 10.1002/ajmg.b.32419

11. *Ficks C.A.*, *Waldman I.D.* Candidate genes for aggression and antisocial behavior: a meta-analysis of association studies of the 5httlpr and maoa-uvntr. *Behav Genet*. 2014;44(5):427-444. doi 10.1007/s10519-014-9661-y

12. *Vassos E.*, *Collier D.A.*, *Fazel S.* Systematic meta-analyses and field synopsis of genetic association studies of violence and aggression. *Mol Psychiatry*. 2014;19(4):471-477. doi 10.1038/mp.2013.31

13. *Koyama E., Kant T., Takata A., et al.* Genetics of child aggression, a systematic review. *Transl Psychiatry*. 2024;14(1):252. doi 10.1038/s41398-024-02870-7

14. *Zhang‐James Y.*, *Faraone S.V.* Genetic architecture for human aggression: a study of gene–phenotype relationship in omim. *Am J Med Genet B Neuropsychiatr Genet*. 2016;171(5):641-649. doi 10.1002/ajmg.b.32363

15. *Cases O., Seif I., Grimsby J., et al.* Aggressive behavior and altered amounts of brain serotonin and norepinephrine in mice lacking maoa. *Science*. 1995;268(5218):1763-1766. doi 10.1126/science.7792602

16. *Odintsova V.V., Roetman P.J., Ip H.F., et al.* Genomics of human aggression: current state of genome-wide studies and an automated systematic review tool. *Psychiatr Genet*. 2019;29(5):170-190. doi 10.1097/YPG.0000000000000239

17. *Tielbeek J.J., Johansson A., Polderman T.J.C., et al.* Genome-wide association studies of a broad spectrum of antisocial behavior. *JAMA Psychiatry*. 2017;74(12):1242. doi 10.1001/jamapsychiatry.2017.3069

18. *Pappa I., St Pourcain B., Benke K., et al.* A genome‐wide approach to children's aggressive behavior: the eagle consortium. *Am J Med Genet B Neuropsychiatr Genet*. 2016;171(5):562-572. doi 10.1002/ajmg.b.32333

19. *Ip H.F., Van Der Laan C.M., Krapohl E.M.L., et al.* Genetic association study of childhood aggression across raters, instruments, and age. *Transl Psychiatry*. 2021;11(1):413. doi 10.1038/s41398-021-01480-x

20. *Tielbeek J.J., Uffelmann E., Williams B.S., et al.* Uncovering the genetic architecture of broad antisocial behavior through a genome-wide association study meta-analysis. *Mol Psychiatry*. 2022;27(11):4453-4463. doi 10.1038/s41380-022-01793-3

21. *Zhang-James Y., Fernàndez-Castillo N., Hess J.L., et al.* An integrated analysis of genes and functional pathways for aggression in human and rodent models. *Mol Psychiatry*. 2019;24(11):1655-1667. doi 10.1038/s41380-018-0068-7

22. *Piras I.S.*, *Braccagni G.*, *Huentelman M.J.*, *Bortolato M.* A preliminary transcriptomic analysis of the orbitofrontal cortex of antisocial individuals. *CNS Neurosci Ther*. 2023;29(11):3173-3182. doi 10.1111/cns.14283

23. *Punzi G., Ursini G., Viscanti G., et al.* Association of a noncoding rna postmortem with suicide by violent means and in vivo with aggressive phenotypes. *Biol Psychiatry*. 2019;85(5):417-424. doi 10.1016/j.biopsych.2018.11.002

24. *Coccaro E.F., Irwin M., Arevalo J.M.G., et al.* Gene expression in peripheral blood mononuclear cells in impulsive aggression: intermittent explosive disorder compared with non-aggressive healthy and psychiatric controls. *Psychoneuroendocrinology*. 2022;136:105453. doi 10.1016/j.psyneuen.2021.105453

25. *Malki K., Du Rietz E., Crusio W.E., et al.* Transcriptome analysis of genes and gene networks involved in aggressive behavior in mouse and zebrafish. *Am J Med Genet B Neuropsychiatr Genet*. 2016;171(6):827-838. doi 10.1002/ajmg.b.32451

26. *Reichmann F.*, *Pilic J.*, *Trajanoski S.*, *Norton W.H.J.* Transcriptomic underpinnings of high and low mirror aggression zebrafish behaviours. *BMC Biol*. 2022;20(1):97. doi 10.1186/s12915-022-01298-z

27. *Redina O., Babenko V., Smagin D., et al.* Gene expression changes in the ventral tegmental area of male mice with alternative social behavior experience in chronic agonistic interactions. *Int J Mol Sci*. 2020;21(18):6599. doi 10.3390/ijms21186599

28. *Rawat R.S., Bhambri A., Pal M., et al.* Early life stressful experiences escalate aggressive behavior in adulthood via changes in transthyretin expression and function. *eLife*. 2022;11:e77968. doi 10.7554/eLife.77968

29. *McCann K.E., Sinkiewicz D.M., Rosenhauer A.M., et al.* Transcriptomic analysis reveals sex-dependent expression patterns in the basolateral amygdala of dominant and subordinate animals after acute social conflict. *Mol Neurobiol*. 2019;56(5):3768-3779. doi 10.1007/s12035-018-1339-7

30. *Malki K., Pain O., Du Rietz E., et al.* Genes and gene networks implicated in aggression related behaviour. *Neurogenetics*. 2014;15(4):255-266. doi 10.1007/s10048-014-0417-x

31. *Brodkin E.S., Goforth S.A., Keene A.H., et al.* Identification of quantitative trait loci that affect aggressive behavior in mice. *J Neurosci*. 2002;22(3):1165-1170. doi 10.1523/JNEUROSCI.22-03-01165.2002

32. *Nehrenberg D.L., Wang S., Buus R.J., et al.* Genomic mapping of social behavior traits in a f2 cross derived from mice selectively bred for high aggression. *BMC Genet*. 2010;11(1):113. doi 10.1186/1471-2156-11-113

33. *Delprato A., Bonheur B., Algéo M. ‑P., et al.* A quantitative trait locus on chromosome 1 modulates intermale aggression in mice. *Genes Brain Behav*. 2018;17(7):e12469. doi 10.1111/gbb.12469

34. *Albert F.W., Carlborg Ö., Plyusnina I., et al.* Genetic architecture of tameness in a rat model of animal domestication. *Genetics*. 2009;182(2):541-554. doi 10.1534/genetics.109.102186

35. *Edwards A.C.*, *Rollmann S.M.*, *Morgan T.J.*, *Mackay T.F.C.* Quantitative genomics of aggressive behavior in drosophila melanogaster. Frankel WN, ed. *PLoS Genet*. 2006;2(9):e154. doi 10.1371/journal.pgen.0020154

36. *Shorter J., Couch C., Huang W., et al.* Genetic architecture of natural variation in drosophila melanogaster aggressive behavior. *Proc Natl Acad Sci*. 2015;112(27). doi 10.1073/pnas.1510104112

37. *Hoyer S.C., Eckart A., Herrel A., et al.* Octopamine in male aggression of drosophila. *Curr Biol*. 2008;18(3):159-167. doi 10.1016/j.cub.2007.12.052

38. *Alekseyenko O.V.*, *Chan Y.B.*, *Li R.*, *Kravitz E.A.* Single dopaminergic neurons that modulate aggression in drosophila. *Proc Natl Acad Sci*. 2013;110(15):6151-6156. doi 10.1073/pnas.1303446110

39. *Baran N.M.*, *Streelman J.T.* Ecotype differences in aggression, neural activity and behaviorally relevant gene expression in cichlid fish. *Genes Brain Behav*. 2020;19(6):e12657. doi 10.1111/gbb.12657

40. *Merritt J.R., Grogan K.E., Zinzow-Kramer W.M., et al.* A supergene-linked estrogen receptor drives alternative phenotypes in a polymorphic songbird. *Proc Natl Acad Sci*. 2020;117(35):21673-21680. doi 10.1073/pnas.2011347117

41. *Zapata I.*, *Serpell J.A.*, *Alvarez C.E.* Genetic mapping of canine fear and aggression. *BMC Genomics*. 2016;17(1):572. doi 10.1186/s12864-016-2936-3

42. *Battivelli D.*, *Fan Z.*, *Hu H.*, *Gross C.T.* How can ethology inform the neuroscience of fear, aggression and dominance? *Nat Rev Neurosci*. 2024;25(12):809-819. doi 10.1038/s41583-024-00858-2

43. *Heyne H.O., Lautenschläger S., Nelson R., et al.* Genetic influences on brain gene expression in rats selected for tameness and aggression. *Genetics*. 2014;198(3):1277-1290. doi 10.1534/genetics.114.168948

44. *Alekseyenko O.V.*, *Chan Y.B.*, *Li R.*, *Kravitz E.A.* Single dopaminergic neurons that modulate aggression in *drosophila*. *Proc Natl Acad Sci*. 2013;110(15):6151-6156. doi 10.1073/pnas.1303446110

45. *Bell A.M.*, *Bukhari S.A.*, *Sanogo Y.O.* Natural variation in brain gene expression profiles of aggressive and nonaggressive individual sticklebacks. *Behaviour*. 2016;153(13-14):1723-1743. doi 10.1163/1568539X-00003393

46. *Larson G.*, *Fuller D.Q.* The evolution of animal domestication. *Annu Rev Ecol Evol Syst*. 2014;45(1):115-136. doi 10.1146/annurev-ecolsys-110512-135813

47. *Zeder M.A.* Core questions in domestication research. *Proc Natl Acad Sci*. 2015;112(11):3191-3198. doi 10.1073/pnas.1501711112

48. *Frantz L.A.F.*, *Bradley D.G.*, *Larson G.*, *Orlando L.* Animal domestication in the era of ancient genomics. *Nat Rev Genet*. 2020;21(8):449-460. doi 10.1038/s41576-020-0225-0

49. *Jensen P.* Behavior genetics and the domestication of animals. *Annu Rev Anim Biosci*. 2014;2(1):85-104. doi 10.1146/annurev-animal-022513-114135

50. *Trut L.* Early canid domestication: the farm-fox experiment. *Am Sci*. 1999;87(2):160. doi 10.1511/1999.2.160

51. *Trut L.*, *Oskina I.*, *Kharlamova A.* Animal evolution during domestication: the domesticated fox as a model. *BioEssays*. 2009;31(3):349-360. doi 10.1002/bies.200800070

52. *Statham M.J., Trut L.N., Sacks B.N., et al.* On the origin of a domesticated species: identifying the parent population of russian silver foxes (vulpes vulpes): the origin of russian silver foxes. *Biol J Linn Soc*. 2011;103(1):168-175. doi 10.1111/j.1095-8312.2011.01629.x

53. *Belyaev D.K.* Destabilizing selection as a factor in domestication. *J Hered*. 1979;70(5):301-308. doi 10.1093/oxfordjournals.jhered.a109263

54. *Lord K.A.*, *Larson G.*, *Coppinger R.P.*, *Karlsson E.K.* The history of farm foxes undermines the animal domestication syndrome. *Trends Ecol Evol*. 2020;35(2):125-136. doi 10.1016/j.tree.2019.10.011

55. *Hare B., Plyusnina I., Ignacio N., et al.* Social cognitive evolution in captive foxes is a correlated by-product of experimental domestication. *Curr Biol*. 2005;15(3):226-230. doi 10.1016/j.cub.2005.01.040

56. *Hansen S.W.* Selection for behavioural traits in farm mink. *Appl Anim Behav Sci*. 1996;49(2):137-148. doi 10.1016/0168-1591(96)01045-3

57. *Malmkvist J.*, *Hansen S.W.* Generalization of fear in farm mink, mustela vison, genetically selected for behaviour towards humans. *Anim Behav*. 2002;64(3):487-501. doi 10.1006/anbe.2002.3058

58. *Thirstrup J.P.*, *Villumsen T.M.*, *Malmkvist J.*, *Lund M.S.* Selection for temperament has no negative consequences on important production traits in farmed mink1. *J Anim Sci*. 2019;97(5):1987-1995. doi 10.1093/jas/skz089

59. *Hansen B.K.*, *Jeppesen L.L.*, *Berg P.* Stereotypic behaviour in farm mink (*neovison vison*) can be reduced by selection. *J Anim Breed Genet*. 2010;127(1):64-73. doi 10.1111/j.1439-0388.2009.00814.x

60. *Alemu S., Bijma P., Møller S., et al.* Indirect genetic effects contribute substantially to heritable variation in aggression-related traits in group-housed mink (neovison vison). *Genet Sel Evol*. 2014;46(1):30. doi 10.1186/1297-9686-46-30

61. *Klotchkov D.V.*, *Trapezov O.V.*, *Kharlamova A.V.* Folliculogenesis, onset of puberty and fecundity of mink () selectively bred for docility or aggressiveness. *Theriogenology*. 1998;49(8):1545-1553. doi 10.1016/S0093-691X(98)00100-9

62. *Trapezov O.V.*, *Trapezova L.I.*, *Sergeev E.G.* [Effect of coat color mutations on behavioral polymorphism in farm populations of american minks (mustela vison schreber, 1777) and sables (martes zibellina linnaeus, 1758)]. *Genetika*. 2008;44(4):516-523.

63. *Kukekova A.V., Trut L.N., Chase K., et al.* Measurement of segregating behaviors in experimental silver fox pedigrees. *Behav Genet*. 2008;38(2):185-194. doi 10.1007/s10519-007-9180-1

64. *Kukekova A.V., Trut L.N., Chase K., et al.* Mapping loci for fox domestication: deconstruction/reconstruction of a behavioral phenotype. *Behav Genet*. 2011;41(4):593-606. doi 10.1007/s10519-010-9418-1

65. *Kukekova A.V., Temnykh S.V., Johnson J.L., et al.* Genetics of behavior in the silver fox. *Mamm Genome*. 2012;23(1-2):164-177. doi 10.1007/s00335-011-9373-z

66. *Johnson J.L., Wittgenstein H., Mitchell S.E., et al.* Genotyping-by-sequencing (gbs) detects genetic structure and confirms behavioral qtl in tame and aggressive foxes (vulpes vulpes). Murphy WJ, ed. *PLOS ONE*. 2015;10(6):e0127013. doi 10.1371/journal.pone.0127013

67. *Kukekova A.V., Johnson J.L., Xiang X., et al.* Red fox genome assembly identifies genomic regions associated with tame and aggressive behaviours. *Nat Ecol Evol*. 2018;2(9):1479-1491. doi 10.1038/s41559-018-0611-6

68. *Wang X., Pipes L., Trut L.N., et al.* Genomic responses to selection for tame/aggressive behaviors in the silver fox (vulpes vulpes). *Proc Natl Acad Sci*. 2018;115(41):10398-10403. doi 10.1073/pnas.1800889115

69. *Rosenfeld C.S., Hekman J.P., Johnson J.L., et al.* Hypothalamic transcriptome of tame and aggressive silver foxes (vulpes vulpes) identifies gene expression differences shared across brain regions. *Genes Brain Behav*. 2020;19(1):1-22. doi 10.1111/gbb.12614

70. *Alexandrovich Yu.V., Antonov E.V., Shikhevich S.G., et al.* The expression profile of genes associated with behavior, stress, and adult neurogenesis along the hippocampal dorsoventral axis in tame and aggressive foxes. *Vavilov J Genet Breed*. 2023;27(6):651-661. doi 10.18699/VJGB-23-76

71. *Popova N.K.*, *Voitenko N.N.*, *Kulikov A.V.*, *Avgustinovich D.F.* Evidence for the involvement of central serotonin in mechanism of domestication of silver foxes. *Pharmacol Biochem Behav*. 1991;40(4):751-756. doi 10.1016/0091-3057(91)90080-L

72. *Gulevich R.G., Oskina I.N., Shikhevich S.G., et al.* Effect of selection for behavior on pituitary-adrenal axis and proopiomelanocortin gene expression in silver foxes (vulpes vulpes). *Physiol Behav*. 2004;82(2-3):513-518. doi 10.1016/j.physbeh.2004.04.062

73. *Hecht E.E., Kukekova A.V., Gutman D.A., et al.* Neuromorphological changes following selection for tameness and aggression in the russian farm-fox experiment. *J Neurosci*. 2021;41(28):6144-6156. doi 10.1523/JNEUROSCI.3114-20.2021

74. *Rogers Flattery C.N., Abdulla M., Barton S.A., et al.* The brain of the silver fox (vulpes vulpes): a neuroanatomical reference of cell-stained histological and mri images. *Brain Struct Funct*. 2023;228(5):1177-1189. doi 10.1007/s00429-023-02648-5

75. *Waltes R.*, *Chiocchetti A.G.*, *Freitag C.M.* The neurobiological basis of human aggression: a review on genetic and epigenetic mechanisms. *Am J Med Genet B Neuropsychiatr Genet*. 2016;171(5):650-675. doi 10.1002/ajmg.b.32388

76. *Duclot F.*, *Kabbaj M.* Epigenetics of aggression. In: 2021:283-310. doi 10.1007/7854_2021_252

77. Genetics and epigenetics of human aggression. In: *Brain and Crime*. V 197. Handbook of Clinical Neurology. Elsevier; 2023:13-44. doi 10.1016/B978-0-12-821375-9.00005-0

78. *Provençal N.*, *Booij L.*, *Tremblay R.E.* The developmental origins of chronic physical aggression: biological pathways triggered by early life adversity. Hoppeler HH, ed. *J Exp Biol*. 2015;218(1):123-133. doi 10.1242/jeb.111401

79. *Van Dongen J., Hagenbeek F.A., Suderman M., et al.* DNA methylation signatures of aggression and closely related constructs: a meta-analysis of epigenome-wide studies across the lifespan. *Mol Psychiatry*. 2021;26(6):2148-2162. doi 10.1038/s41380-020-00987-x

80. *Caspi A., McClay J., Moffitt T.E., et al.* Role of genotype in the cycle of violence in maltreated children. *Science*. 2002;297(5582):851-854. doi 10.1126/science.1072290

81. *Weaver I.C.G., Cervoni N., Champagne F.A., et al.* Epigenetic programming by maternal behavior. *Nat Neurosci*. 2004;7(8):847-854. doi 10.1038/nn1276

82. *Roth T.L.*, *Lubin F.D.*, *Funk A.J.*, *Sweatt J.D.* Lasting epigenetic influence of early-life adversity on the bdnf gene. *Biol Psychiatry*. 2009;65(9):760-769. doi 10.1016/j.biopsych.2008.11.028

83. *Suderman M., McGowan P.O., Sasaki A., et al.* Conserved epigenetic sensitivity to early life experience in the rat and human hippocampus. *Proc Natl Acad Sci*. 2012;109(supplement_2):17266-17272. doi 10.1073/pnas.1121260109

84. *Márquez C., Poirier G.L., Cordero M.I., et al.* Peripuberty stress leads to abnormal aggression, altered amygdala and orbitofrontal reactivity and increased prefrontal maoa gene expression. *Transl Psychiatry*. 2013;3(1):e216-e216. doi 10.1038/tp.2012.144

85. *Konar A.*, *Rastogi M.*, *Bhambri A.* Brain region specific methylation and sirt1 binding changes in maoa promoter is associated with sexual dimorphism in early life stress induced aggressive behavior. *Neurochem Int*. 2019;129:104510. doi 10.1016/j.neuint.2019.104510

86. *Matrisciano F.*, *Pinna G.* PPAR-α hypermethylation in the hippocampus of mice exposed to social isolation stress is associated with enhanced neuroinflammation and aggressive behavior. *Int J Mol Sci*. 2021;22(19):10678. doi 10.3390/ijms221910678

87. *Labonté B., Abdallah K., Maussion G., et al.* Regulation of impulsive and aggressive behaviours by a novel lncrna. *Mol Psychiatry*. 2021;26(8):3751-3764. doi 10.1038/s41380-019-0637-4

88. *Jensen K.P., Covault J., Conner T.S., et al.* A common polymorphism in serotonin receptor 1b mrna moderates regulation by mir-96 and associates with aggressive human behaviors. *Mol Psychiatry*. 2009;14(4):381-389. doi 10.1038/mp.2008.15

89. *González-Giraldo Y., Camargo A., López-León S., et al.* A functional snp in mir124-1, a brain expressed mirna gene, is associated with aggressiveness in a colombian sample. *Eur Psychiatry*. 2015;30(4):499-503. doi 10.1016/j.eurpsy.2015.03.002

90. *Prichard M.R., Grogan K.E., Merritt J.R., et al.* Allele-specific cis-regulatory methylation of the gene for vasoactive intestinal peptide in white-throated sparrows. *Genes Brain Behav*. 2022;21(8):e12831. doi 10.1111/gbb.12831

91. *Bentz A.B., George E.M., Wolf S.E., et al.* Experimental competition induces immediate and lasting effects on the neurogenome in free-living female birds. *Proc Natl Acad Sci*. 2021;118(13):e2016154118. doi 10.1073/pnas.2016154118

92. *Gupta T., Morgan H.R., Andrews J.C., et al.* Methyl-cpg binding domain proteins inhibit interspecies courtship and promote aggression in drosophila. *Sci Rep*. 2017;7(1):5420. doi 10.1038/s41598-017-05844-6

93. *Bresnahan S.T., Lee E., Clark L., et al.* Examining parent-of-origin effects on transcription and rna methylation in mediating aggressive behavior in honey bees (apis mellifera). *BMC Genomics*. 2023;24(1):315. doi 10.1186/s12864-023-09411-4

94. *Allis C.D.*, *Jenuwein T.* The molecular hallmarks of epigenetic control. *Nat Rev Genet*. 2016;17(8):487-500. doi 10.1038/nrg.2016.59

95. *Kouzarides T.* Chromatin modifications and their function. *Cell*. 2007;128(4):693-705. doi 10.1016/j.cell.2007.02.005

96. *Barski A., Cuddapah S., Cui K., et al.* High-resolution profiling of histone methylations in the human genome. *Cell*. 2007;129(4):823-837. doi 10.1016/j.cell.2007.05.009

97. *Wang Z., Zang C., Rosenfeld J.A., et al.* Combinatorial patterns of histone acetylations and methylations in the human genome. *Nat Genet*. 2008;40(7):897-903. doi 10.1038/ng.154

98. *Heintzman N.D., Stuart R.K., Hon G., et al.* Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat Genet*. 2007;39(3):311-318. doi 10.1038/ng1966

99. *Creyghton M.P., Cheng A.W., Welstead G.G., et al.* Histone h3k27ac separates active from poised enhancers and predicts developmental state. *Proc Natl Acad Sci*. 2010;107(50):21931-21936. doi 10.1073/pnas.1016071107

100. *Rada-Iglesias A., Bajpai R., Swigut T., et al.* A unique chromatin signature uncovers early developmental enhancers in humans. *Nature*. 2011;470(7333):279-283. doi 10.1038/nature09692

101. *Calo E.*, *Wysocka J.* Modification of enhancer chromatin: what, how, and why? *Mol Cell*. 2013;49(5):825-837. doi 10.1016/j.molcel.2013.01.038

102. *Howe F.S.*, *Fischl H.*, *Murray S.C.*, *Mellor J.* Is h3k4me3 instructive for transcription activation? *BioEssays*. 2017;39(1):1-12. doi 10.1002/bies.201600095

103. *Dincer A., Gavin D.P., Xu K., et al.* Deciphering h3k4me3 broad domains associated with gene-regulatory networks and conserved epigenomic landscapes in the human brain. *Transl Psychiatry*. 2015;5(11):e679-e679. doi 10.1038/tp.2015.169

104. *Girdhar K., Hoffman G.E., Jiang Y., et al.* Cell-specific histone modification maps in the human frontal lobe link schizophrenia risk to the neuronal epigenome. *Nat Neurosci*. 2018;21(8):8. doi 10.1038/s41593-018-0187-0

105. *Kim T.K., Hemberg M., Gray J.M., et al.* Widespread transcription at neuronal activity-regulated enhancers. *Nature*. 2010;465(7295):182-187. doi 10.1038/nature09033

106. *Malik A.N., Vierbuchen T., Hemberg M., et al.* Genome-wide identification and characterization of functional neuronal activity–dependent enhancers. *Nat Neurosci*. 2014;17(10):1330-1339. doi 10.1038/nn.3808

107. *Hernandez Carballo L.G.*, *Li P.*, *Senek R.*, *Yan Z.* Systemic histone deacetylase inhibition ameliorates the aberrant responses to acute stress in socially isolated male mice. *J Physiol*. 2024;602(9):2047-2060. doi 10.1113/JP285875

108. *Zhang F., Rein B., Zhong P., et al.* Synergistic inhibition of histone modifiers produces therapeutic effects in adult shank3-deficient mice. *Transl Psychiatry*. 2021;11(1):99. doi 10.1038/s41398-021-01233-w

109. *Quah Y.C.R.*, *Li D.* Epigenetic modification of fruitless in the protocerebrum influences male drosophila courtship behaviour. In: Guo H, Ren H, Wang V, et al., eds. *IRC-SET 2021*. Springer Nature Singapore; 2022:569-582. doi 10.1007/978-981-16-9869-9_45

110. *Roadmap Epigenomics Consortium, Kundaje A., Meuleman W., et al.* Integrative analysis of 111 reference human epigenomes. *Nature*. 2015;518(7539):317-330. doi 10.1038/nature14248

111. *The ENCODE Project Consortium, Abascal F., Acosta R., et al.* Expanded encyclopaedias of dna elements in the human and mouse genomes. *Nature*. 2020;583(7818):699-710. doi 10.1038/s41586-020-2493-4

112. *The GTEx Consortium, Aguet F., Anand S., et al.* The gtex consortium atlas of genetic regulatory effects across human tissues. *Science*. 2020;369(6509):1318-1330. doi 10.1126/science.aaz1776

113. *Li M., Santpere G., Imamura Kawasawa Y., et al.* Integrative functional genomic analysis of human brain development and neuropsychiatric risks. *Science*. 2018;362(6420):eaat7615. doi 10.1126/science.aat7615

114. *Wang D., Liu S., Warrell J., et al.* Comprehensive functional genomic resource and integrative model for the human brain. *Science*. 2018;362(6420):eaat8464. doi 10.1126/science.aat8464

115. *Giambartolomei C., Vukcevic D., Schadt E.E., et al.* Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. Williams SM, ed. *PLoS Genet*. 2014;10(5):e1004383. doi 10.1371/journal.pgen.1004383

116. *Gusev A., Ko A., Shi H., et al.* Integrative approaches for large-scale transcriptome-wide association studies. *Nat Genet*. 2016;48(3):245-252. doi 10.1038/ng.3506

117. *Fulco C.P., Nasser J., Jones T.R., et al.* Activity-by-contact model of enhancer–promoter regulation from thousands of crispr perturbations. *Nat Genet*. 2019;51(12):1664-1669. doi 10.1038/s41588-019-0538-0

118. *Nasser J., Bergman D.T., Fulco C.P., et al.* Genome-wide enhancer maps link risk variants to disease genes. *Nature*. 2021;593(7858):238-243. doi 10.1038/s41586-021-03446-x

119. *Langfelder P.*, *Horvath S.* WGCNA: an r package for weighted correlation network analysis. *BMC Bioinformatics*. 2008;9(1):559. doi 10.1186/1471-2105-9-559

120. *Argelaguet R., Velten B., Arnol D., et al.* Multi‐Omics factor analysis—a framework for unsupervised integration of multi‐omics data sets. *Mol Syst Biol*. 2018;14(6):MSB178124. doi 10.15252/msb.20178124

121. Amygdala connectivity and aggression. In: *Brain and Crime*. V 197. Handbook of Clinical Neurology. Elsevier; 2023:87-106. doi 10.1016/B978-0-12-821375-9.00002-5

122. *Fanti K.A., Zacharaki G., Theodotou P., et al.* Neurobiology of violence. In: *Handbook of Gun Violence*. Elsevier; 2025:119-131. doi 10.1016/B978-0-323-95272-9.00024-3

123. *Rando H.M., Alexander E.P., Preckler-Quisquater S., et al.* Missing history of a modern domesticate: historical demographics and genetic diversity in farm-bred red fox populations. vonHoldt B, ed. *J Hered*. 2024;115(4):411-423. doi 10.1093/jhered/esae022

124. *Li H.*, *Durbin R.* Fast and accurate short read alignment with burrows-wheeler transform. *Bioinforma Oxf Engl*. 2009;25(14):14. doi 10.1093/bioinformatics/btp324

125. Picard toolkit. Published online 2019. https://broadinstitute.github.io/picard/

126. *McKenna A., Hanna M., Banks E., et al.* The genome analysis toolkit: a mapreduce framework for analyzing next-generation dna sequencing data. *Genome Res*. 2010;20(9):1297-1303. doi 10.1101/gr.107524.110

127. *DePristo M.A., Banks E., Poplin R., et al.* A framework for variation discovery and genotyping using next-generation dna sequencing data. *Nat Genet*. 2011;43(5):491-498. doi 10.1038/ng.806

128. *Poplin R., Ruano-Rubio V., DePristo M.A., et al.* Scaling accurate genetic variant discovery to tens of thousands of samples. *Genomics*. Preprint posted online November 14, 2017. doi 10.1101/201178

129. *Li H.* A statistical framework for snp calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics*. 2011;27(21):2987-2993. doi 10.1093/bioinformatics/btr509

130. *Danecek P., Bonfield J.K., Liddle J., et al.* Twelve years of samtools and bcftools. *GigaScience*. 2021;10(2):giab008. doi 10.1093/gigascience/giab008

131. *Delaneau O.*, *Zagury J.F.*, *Marchini J.* Improved whole-chromosome phasing for disease and population genetic studies. *Nat Methods*. 2013;10(1):5-6. doi 10.1038/nmeth.2307

132. *Chang C.C., Chow C.C., Tellier L.C., et al.* Second-generation plink: rising to the challenge of larger and richer datasets. *Gigascience*. 2015;4(1):s13742-015-0047-0048. doi 10.1186/s13742-015-0047-8

133. *Patterson N.*, *Price A.L.*, *Reich D.* Population structure and eigenanalysis. *PLoS Genet*. 2006;2(12):e190. doi 10.1371/journal.pgen.0020190

134. *Privé F., Luu K., Blum M.G.B., et al.* Efficient toolkit implementing best practices for principal component analysis of population genetic data. Schwartz R, ed. *Bioinformatics*. 2020;36(16):4449-4457. doi 10.1093/bioinformatics/btaa520

135. *Alexander D.H.*, *Novembre J.*, *Lange K.* Fast model-based estimation of ancestry in unrelated individuals. *Genome Res*. 2009;19(9):1655-1664. doi 10.1101/gr.094052.109

136. Https://ggplot2.tidyverse.org.

137. Https://cran.r-project.org/package=cmplot.

138. *Manakhov A.D., Andreeva T.V., Trapezov O.V., et al.* Genome analysis identifies the mutant genes for common industrial silverblue and hedlund white coat colours in american mink. *Sci Rep*. 2019;9(1):4581. doi 10.1038/s41598-019-40918-7

139. *Manakhov A.D., Mintseva M.Yu., Andreev I.A., et al.* Genome analysis of american minks reveals link of mutations in ras-related protein-38 gene to moyle brown coat phenotype. *Sci Rep*. 2020;10(1):15876. doi 10.1038/s41598-020-72239-5

140. *Manakhov A.D., Mintseva M.Yu., Andreeva T.V., et al.* Shadow coat colour in american mink associated with a missense mutation in the *kit* gene. *Anim Genet*. 2022;53(4):522-525. doi 10.1111/age.13202

141. *Li H.* Aligning sequence reads, clone sequences and assembly contigs with bwa-mem. Published online 2013. doi 10.48550/ARXIV.1303.3997

142. *McKenna A., Hanna M., Banks E., et al.* The genome analysis toolkit: a mapreduce framework for analyzing next-generation dna sequencing data. *Genome Res*. 2010;20(9):1297-1303. doi 10.1101/gr.107524.110

143. *Purcell S., Neale B., Todd-Brown K., et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007;81(3):559-575. doi 10.1086/519795

144. *Kim D., Paggi J.M., Park C., et al.* Graph-based genome alignment and genotyping with hisat2 and hisat-genotype. *Nat Biotechnol*. 2019;37(8):8. doi 10.1038/s41587-019-0201-4

145. *Liao Y.*, *Smyth G.K.*, *Shi W.* FeatureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinforma Oxf Engl*. 2014;30(7):7. doi 10.1093/bioinformatics/btt656

146. *Love M.I.*, *Huber W.*, *Anders S.* Moderated estimation of fold change and dispersion for rna-seq data with deseq2. *Genome Biol*. 2014;15(12):12. doi 10.1186/s13059-014-0550-8

147. *Rohart F.*, *Gautier B.*, *Singh A.*, *Lê Cao K.A.* MixOmics: an r package for ‘omics feature selection and multiple data integration. Schneidman D, ed. *PLOS Comput Biol*. 2017;13(11):e1005752. doi 10.1371/journal.pcbi.1005752

148. *Almeida-Silva F.*, *Venancio T.M.* BioNERO: an all-in-one r/bioconductor package for comprehensive and easy biological network reconstruction. *Funct Integr Genomics*. 2022;22(1):131-136. doi 10.1007/s10142-021-00821-9

149. *Oldham M.C.*, *Langfelder P.*, *Horvath S.* Network methods for describing sample relationships in genomic datasets: application to huntington's disease. *BMC Syst Biol*. 2012;6(1):63. doi 10.1186/1752-0509-6-63

150. *Florence C.S., Bergen G., Atherly A., et al.* Practical guidelines for high-resolution epigenomic profiling of nucleosomal histones in postmortem human brain tissue. *Physiol Behav*. 2018;66(4):693-698. doi 10.1016/j.biopsych.2016.03.1048.Practical

151. *Zhang Y., Liu T., Meyer C.A., et al.* Model-based analysis of chip-seq (macs). *Genome Biol*. 2008;9(9):R137. doi 10.1186/gb-2008-9-9-r137

152. *Rainer J.*, *Gatto L.*, *Weichenberger C.X.* Ensembldb: an r package to create and use ensembl-based annotation resources. Hancock J, ed. *Bioinformatics*. 2019;35(17):3151-3153. doi 10.1093/bioinformatics/btz031

153. *Yu G.*, *Wang L.G.*, *He Q.Y.* ChIPseeker: an r/bioconductor package for chip peak annotation, comparison and visualization. *Bioinformatics*. 2015;31(14):2382-2383. doi 10.1093/bioinformatics/btv145

154. *Hurst L.D., Sachenkova O., Daub C., et al.* A simple metric of promoter architecture robustly predicts expression breadth of human genes suggesting that most transcription factors are positive regulators. *Genome Biol*. 2014;15(7):413. doi 10.1186/s13059-014-0413-3

155. *Lawrence M.*, *Gentleman R.*, *Carey V.* Rtracklayer: an r package for interfacing with genome browsers. *Bioinformatics*. 2009;25(14):1841-1842. doi 10.1093/bioinformatics/btp328

156. *Lun A.T.L.*, *Smyth G.K.* Csaw: a bioconductor package for differential binding analysis of chip-seq data using sliding windows. *Nucleic Acids Res*. 2016;44(5):5. doi 10.1093/nar/gkv1191

157. *Chen Y., Chen L., Lun A.T.L., et al.* EdgeR v4: powerful differential analysis of sequencing data with expanded functionality and improved support for small counts and larger datasets. *Nucleic Acids Res*. 2025;53(2):gkaf018. doi 10.1093/nar/gkaf018

158. *Ignatiadis N.*, *Klaus B.*, *Zaugg J.B.*, *Huber W.* Data-driven hypothesis weighting increases detection power in genome-scale multiple testing. *Nat Methods*. 2016;13(7):577-580. doi 10.1038/nmeth.3885

159. *Paoletti P.*, *Bellone C.*, *Zhou Q.* NMDA receptor subunit diversity: impact on receptor properties, synaptic plasticity and disease. *Nat Rev Neurosci*. 2013;14(6):383-400. doi 10.1038/nrn3504

160. *Timms R.T., Tchasovnikarova I.A., Antrobus R., et al.* ATF7IP-mediated stabilization of the histone methyltransferase setdb1 is essential for heterochromatin formation by the hush complex. *Cell Rep*. 2016;17(3):653-659. doi 10.1016/j.celrep.2016.09.050

161. *Rouvinen K.*, *Archbold S.*, *Laffin S.*, *Harri M.* Long-term effects of tryptophan on behavioural response and growing-furring performance in silver fox (vulpes vulpes). *Appl Anim Behav Sci*. 1999;63(1):65-77. doi 10.1016/S0168-1591(98)00241-X

162. *Kulikov A.V., Bazhenova E.Y., Kulikova E.A., et al.* Interplay between aggression, brain monoamines and fur color mutation in the american mink. *Genes Brain Behav*. 2016;15(8):733-740. doi 10.1111/gbb.12313

163. *Coulter P.M.*, *Bautista E.A.*, *Margulies J.E.*, *Watson J.B.* Identification of cortexin: a novel, neuron‐specific, 82‐residue membrane protein enriched in rodent cerebral cortex. *J Neurochem*. 1993;61(2):756-759. doi 10.1111/j.1471-4159.1993.tb02183.x

164. *Šerý O., Lochman J., Povová J., et al.* Association between 5q23.2-located polymorphism of ctxn3 gene (cortexin 3) and schizophrenia in european-caucasian males; implications for the aetiology of schizophrenia. *Behav Brain Funct*. 2015;11(1):10. doi 10.1186/s12993-015-0057-9

165. *Higashida H., Hashii M., Tanaka Y., et al.* CD38, cd157, and rage as molecular determinants for social behavior. *Cells*. 2019;9(1):62. doi 10.3390/cells9010062

166. *Higashida H., Liang M., Yoshihara T., et al.* An immunohistochemical, enzymatic, and behavioral study of cd157/bst-1 as a neuroregulator. *BMC Neurosci*. 2017;18(1):35. doi 10.1186/s12868-017-0350-7

167. *Deitmer J.W., Theparambil S.M., Ruminot I., et al.* Energy dynamics in the brain: contributions of astrocytes to metabolism and ph homeostasis. *Front Neurosci*. 2019;13:1301. doi 10.3389/fnins.2019.01301

168. *Chesler M.* Regulation and modulation of ph in the brain. *Physiol Rev*. 2003;83(4):1183-1221. doi 10.1152/physrev.00010.2003

169. *Du Q., Zhang M., Gao A., et al.* Epigenetic silencing *zscan23* promotes pancreatic cancer growth by activating wnt signaling. *Cancer Biol Ther*. 2024;25(1):2302924. doi 10.1080/15384047.2024.2302924

170. *Gharagozloo M., Mahvelati T.M., Imbeault E., et al.* The nod-like receptor, nlrp12, plays an anti-inflammatory role in experimental autoimmune encephalomyelitis. *J Neuroinflammation*. 2015;12(1):198. doi 10.1186/s12974-015-0414-5

171. *Rafiq M.A., Kuss A.W., Puettmann L., et al.* Mutations in the alpha 1,2-mannosidase gene, man1b1, cause autosomal-recessive intellectual disability. *Am J Hum Genet*. 2011;89(1):176-182. doi 10.1016/j.ajhg.2011.06.006

172. *Detre J.*, *Nairn A.*, *Aswad D.*, *Greengard P.* Localization in mammalian brain of g-substrate, a specific substrate for guanosine 3',5'-cyclic monophosphate-dependent protein kinase. *J Neurosci*. 1984;4(11):2843-2849. doi 10.1523/JNEUROSCI.04-11-02843.1984

173. *Siintola E., Topcu M., Aula N., et al.* The novel neuronal ceroid lipofuscinosis gene mfsd8 encodes a putative lysosomal transporter. *Am J Hum Genet*. 2007;81(1):136-146. doi 10.1086/518902

174. *Peters A.*, *Sherman L.S.* Diverse roles for hyaluronan and hyaluronan receptors in the developing and adult nervous system. *Int J Mol Sci*. 2020;21(17):5988. doi 10.3390/ijms21175988

175. *Wieczorek L., Majumdar D., Wills T.A., et al.* Absence of ca2+-stimulated adenylyl cyclases leads to reduced synaptic plasticity and impaired experience-dependent fear memory. *Transl Psychiatry*. 2012;2(5):e126-e126. doi 10.1038/tp.2012.50

176. *Shi W., Chen Q.Y., Ma Y., et al.* Selective enhancement of fear extinction by inhibiting neuronal adenylyl cyclase 1 (ac1) in aged mice. *Mol Brain*. 2024;17(1):11. doi 10.1186/s13041-024-01083-9

177. *Baba K., Yoshida W., Toriyama M., et al.* Gradient-reading and mechano-effector machinery for netrin-1-induced axon guidance. *eLife*. 2018;7:e34593. doi 10.7554/eLife.34593

178. *Toriyama M., Shimada T., Kim K.B., et al.* Shootin1: a protein involved in the organization of an asymmetric signal for neuronal polarization. *J Cell Biol*. 2006;175(1):147-157. doi 10.1083/jcb.200604160

179. *Baba K., Yoshida W., Toriyama M., et al.* Gradient-reading and mechano-effector machinery for netrin-1-induced axon guidance. *eLife*. 2018;7:e34593. doi 10.7554/eLife.34593

180. *Philipson K.D.*, *Quednau B.D.*, *Nicoll D.A.* The sodium/calcium exchanger family?slc8. *Pfl�gers Arch Eur J Physiol*. 2004;447(5):543-548. doi 10.1007/s00424-003-1065-4

181. *Mao B., Wu W., Davidson G., et al.* Kremen proteins are dickkopf receptors that regulate wnt/β-catenin signalling. *Nature*. 2002;417(6889):664-667. doi 10.1038/nature756

182. *Mizukoshi M., Nozawa A., Oomizo S., et al.* Differential localization and roles of splice variants of rat suppressor of cancer cell invasion (scai) in neuronal cells. *Biochem Biophys Res Commun*. 2020;529(3):615-621. doi 10.1016/j.bbrc.2020.06.064

183. *Bakermans-Kranenburg M.J.*, *Van IJzendoorn M.H.* Gene-environment interaction of the dopamine d4 receptor (drd4) and observed maternal insensitivity predicting externalizing behavior in preschoolers. *Dev Psychobiol*. 2006;48(5):406-409. doi 10.1002/dev.20152

184. *Leipold E., Liebmann L., Korenke G.C., et al.* A de novo gain-of-function mutation in scn11a causes loss of pain perception. *Nat Genet*. 2013;45(11):1399-1404. doi 10.1038/ng.2767

185. *Mao Y.*, *Schwarzbauer J.E.* Fibronectin fibrillogenesis, a cell-mediated matrix assembly process. *Matrix Biol*. 2005;24(6):389-399. doi 10.1016/j.matbio.2005.06.008

186. *Zhang Y.H., Sun X.T., Guo R.F., et al.* AβPP-tau-has1 axis trigger has1-related nuclear speckles and gene transcription in alzheimer's disease. *Matrix Biol*. 2024;129:29-43. doi 10.1016/j.matbio.2024.03.003

187. *Görlitz S., Brauer E., Günther R., et al.* Temporal regulation of bmp2 growth factor signaling in response to mechanical loading is linked to cytoskeletal and focal adhesion remodeling. *Commun Biol*. 2024;7(1):1064. doi 10.1038/s42003-024-06753-x

188. *Wei B., Billman Z.P., Nozaki K., et al.* NLRP3, nlrp6, and nlrp12 are inflammasomes with distinct expression patterns. *Front Immunol*. 2024;15:1418290. doi 10.3389/fimmu.2024.1418290

189. *Cortés J.R., Sánchez-Díaz R., Bovolenta E.R., et al.* Maintenance of immune tolerance by foxp3+ regulatory t cells requires cd69 expression. *J Autoimmun*. 2014;55:51-62. doi 10.1016/j.jaut.2014.05.007

190. *Gayden T., Sepulveda F.E., Khuong-Quang D.A., et al.* Germline havcr2 mutations altering tim-3 characterize subcutaneous panniculitis-like t cell lymphomas with hemophagocytic lymphohistiocytic syndrome. *Nat Genet*. 2018;50(12):1650-1657. doi 10.1038/s41588-018-0251-4

191. *Li X., Zhou F., Niu K., et al.* Emerging discoveries on the role of trim14: from diseases to immune regulation. *Cell Death Discov*. 2024;10(1):513. doi 10.1038/s41420-024-02276-w

192. *Gong M., Ye S., Li W.X., et al.* Regulatory function of praja ring finger ubiquitin ligase 2 mediated by the *p2rx3/p2rx7* axis in mouse hippocampal neuronal cells. *Am J Physiol-Cell Physiol*. 2020;318(6):C1123-C1135. doi 10.1152/ajpcell.00070.2019

193. *McHardy S.F.*, *Wang H.Y.L.*, *McCowen S.V.*, *Valdez M.C.* Recent advances in acetylcholinesterase inhibitors and reactivators: an update on the patent literature (2012-2015). *Expert Opin Ther Pat*. 2017;27(4):455-476. doi 10.1080/13543776.2017.1272571

194. *Bay Y., Venskutonytė R., Frantsen S.M., et al.* Small‐molecule positive allosteric modulation of homomeric kainate receptors GluK1‐3: development of screening assays and insight into GluK3 structure. *FEBS J*. 2024;291(7):1506-1529. doi 10.1111/febs.17046

195. *Gong Q.*, *Stump M.R.*, *Zhou Z.* Regulation of kv11.1 potassium channel c-terminal isoform expression by the rna-binding proteins hur and hud. *J Biol Chem*. 2018;293(51):19624-19632. doi 10.1074/jbc.RA118.003720

196. *Laurin N., Ickowicz A., Pathare T., et al.* Investigation of the g protein subunit gαolf gene (gnal) in attention deficit/hyperactivity disorder. *J Psychiatr Res*. 2008;42(2):117-124. doi 10.1016/j.jpsychires.2006.10.010

197. *Terauchi A., Johnson-Venkatesh E.M., Toth A.B., et al.* Distinct fgfs promote differentiation of excitatory and inhibitory synapses. *Nature*. 2010;465(7299):783-787. doi 10.1038/nature09041

198. *Brinks H., Conrad S., Vogt J., et al.* The repulsive guidance molecule rgma is involved in the formation of afferent connections in the dentate gyrus. *J Neurosci*. 2004;24(15):3862-3869. doi 10.1523/JNEUROSCI.5296-03.2004

199. *Ferrario J.E., Baskaran P., Clark C., et al.* Axon guidance in the developing ocular motor system and duane retraction syndrome depends on semaphorin signaling via alpha2-chimaerin. *Proc Natl Acad Sci U S A*. 2012;109(36):14669-14674. doi 10.1073/pnas.1116481109

200. *Kim J.I., Ganesan S., Luo S.X., et al.* Aldehyde dehydrogenase 1a1 mediates a gaba synthesis pathway in midbrain dopaminergic neurons. *Science*. 2015;350(6256):102-106. doi 10.1126/science.aac4690

201. *Melani R.*, *Tritsch N.X.* Inhibitory co-transmission from midbrain dopamine neurons relies on presynaptic gaba uptake. *Cell Rep*. 2022;39(3):110716. doi 10.1016/j.celrep.2022.110716

202. *Niu W., Zang T., Smith D.K., et al.* SOX2 reprograms resident astrocytes into neural progenitors in the adult brain. *Stem Cell Rep*. 2015;4(5):780-794. doi 10.1016/j.stemcr.2015.03.006

203. *Bando Y., Hagiwara Y., Suzuki Y., et al.* Kallikrein 6 secreted by oligodendrocytes regulates the progression of experimental autoimmune encephalomyelitis. *Glia*. 2018;66(2):359-378. doi 10.1002/glia.23249

204. *Leandro‐García L.J., Leskelä S., Landa I., et al.* Tumoral and tissue‐specific expression of the major human β‐tubulin isotypes. *Cytoskeleton*. 2010;67(4):214-223. doi 10.1002/cm.20436

205. *Gjervan S.C., Ozgoren O.K., Gow A., et al.* Claudin-11 in health and disease: implications for myelin disorders, hearing, and fertility. *Front Cell Neurosci*. 2024;17:1344090. doi 10.3389/fncel.2023.1344090

206. *Kopp J., Koch L.A., Lyubenova H., et al.* Loss-of-function variants affecting the staga complex component supt7l cause a developmental disorder with generalized lipodystrophy. *Hum Genet*. 2024;143(5):683-694. doi 10.1007/s00439-024-02669-y

207. *Chistiakov D.A., Chernisheva A., Savost'anov K.V., et al.* The taf5l gene on chromosome 1q42 is associated with type 1 diabetes in russian affected patients. *Autoimmunity*. 2005;38(4):283-293. doi 10.1080/08916930500128594

208. *Zhang C., Zhang S., Xu H., et al.* HOXD1 regulates neural crest cells differentiation and polycerate development in sheep. *Sci Rep*. 2025;15(1):13140. doi 10.1038/s41598-025-97865-9

209. *Wride M.A., Mansergh F.C., Somani J.M., et al.* Characterization and in silico mapping of a novel murine zinc finger transcription factor. *Gene*. 2002;289(1-2):49-59. doi 10.1016/S0378-1119(02)00473-0

210. *Lash T.D., Lamm T.R., Schaber J.A., et al.* Normal and abnormal heme biosynthesis. part 7. synthesis and metabolism of coproporphyrinogen-iii analogues with acetate or butyrate side chains on rings c and d. development of a modified model for the active site of coproporphyrinogen oxidase. *Bioorg Med Chem*. 2011;19(4):1492-1504. doi 10.1016/j.bmc.2010.12.053

211. *Li X., Baumgart E., Dong G.X., et al.* PEX11α is required for peroxisome proliferation in response to 4-phenylbutyrate but is dispensable for peroxisome proliferator-activated receptor alpha-mediated peroxisome proliferation. *Mol Cell Biol*. 2002;22(23):8226-8240. doi 10.1128/MCB.22.23.8226-8240.2002

212. *Inoue H., Baba T., Sato S., et al.* Roles of sam and ddhd domains in mammalian intracellular phospholipase a1 kiaa0725p. *Biochim Biophys Acta BBA - Mol Cell Res*. 2012;1823(4):930-939. doi 10.1016/j.bbamcr.2012.02.002

213. *Short B.* Notch commandeers commd9 for endosomal sorting. *J Cell Biol*. 2015;211(3):486-486. doi 10.1083/jcb.2113iti3

214. *Miyazaki Y., Otsuka T., Yamagata Y., et al.* Oligodendrocyte-derived lgi3 and its receptor adam23 organize juxtaparanodal kv1 channel clustering for short-term synaptic plasticity. *Cell Rep*. 2024;43(1):113634. doi 10.1016/j.celrep.2023.113634

215. *Wang H.G., He X.P., Li Q., et al.* The auxiliary subunit kchip2 is an essential regulator of homeostatic excitability. *J Biol Chem*. 2013;288(19):13258-13268. doi 10.1074/jbc.M112.434548

216. *Hildebrand D., Tiefenbach J., Heinzel T., et al.* Multiple regions of eto cooperate in transcriptional repression. *J Biol Chem*. 2001;276(13):9889-9895. doi 10.1074/jbc.M010582200

217. *Nichol D.*, *Stuhlmann H.* EGFL7: a unique angiogenic signaling factor in vascular development and disease. *Blood*. 2012;119(6):1345-1352. doi 10.1182/blood-2011-10-322446

218. *Bill M., Pathmanathan A., Karunasiri M., et al.* EGFL7 antagonizes notch signaling and represents a novel therapeutic target in acute myeloid leukemia. *Clin Cancer Res*. 2020;26(3):669-678. doi 10.1158/1078-0432.CCR-19-2479

219. *Lv L.L., Du Y.T., Chen X., et al.* Neuroprotective effect of angiopoietin2 is associated with angiogenesis in mouse brain following ischemic stroke. *Brain Sci*. 2022;12(11):1428. doi 10.3390/brainsci12111428

220. *Xu W., Geng R., Zhao Y., et al.* Microfibrillar-associated protein 2 is a prognostic marker that correlates with the immune microenvironment in glioma. *Front Genet*. 2022;13:989521. doi 10.3389/fgene.2022.989521

221. *Clarin J.D.*, *Reddy N.*, *Alexandropoulos C.*, *Gao W.J.* The role of cell adhesion molecule igsf9b at the inhibitory synapse and psychiatric disease. *Neurosci Biobehav Rev*. 2024;156:105476. doi 10.1016/j.neubiorev.2023.105476

222. *Jantrapirom S., Nimlamool W., Temviriyanukul P., et al.* Dystrobrevin is required postsynaptically for homeostatic potentiation at the drosophila nmj. *Biochim Biophys Acta BBA - Mol Basis Dis*. 2019;1865(6):1579-1591. doi 10.1016/j.bbadis.2019.03.008

223. *Ayon R.J.*, *Yuan J.X.J.* Revisiting the role of kcna5 in pulmonary arterial hypertension. *Am J Respir Cell Mol Biol*. 2023;69(2):123-125. doi 10.1165/rcmb.2023-0119ED

224. *Besse A., Wu P., Bruni F., et al.* The gaba transaminase, abat, is essential for mitochondrial nucleoside metabolism. *Cell Metab*. 2015;21(3):417-427. doi 10.1016/j.cmet.2015.02.008

225. *Malik M.A., Saqib M.A.N., Mientjes E., et al.* A loss of function variant in agpat3 underlies intellectual disability and retinitis pigmentosa (idrp) syndrome. *Eur J Hum Genet*. 2023;31(12):1447-1454. doi 10.1038/s41431-023-01475-w

226. *Chalertpet K., Pin-on P., Aporntewan C., et al.* Argonaute 4 as an effector protein in rna-directed dna methylation in human cells. *Front Genet*. 2019;10:645. doi 10.3389/fgene.2019.00645

227. *Gebhardt A., Habjan M., Benda C., et al.* MRNA export through an additional cap-binding complex consisting of ncbp1 and ncbp3. *Nat Commun*. 2015;6(1):8192. doi 10.1038/ncomms9192

228. *Zhou L., Zhao X., Sun J., et al.* Mina53 demethylates histone h4 arginine 3 asymmetric dimethylation to regulate neural stem/progenitor cell identity. *Nat Commun*. 2024;15(1):10227. doi 10.1038/s41467-024-54680-6

229. *Bräuer K.E., Brockers K., Moneer J., et al.* Phylogenetic and genomic analyses of the ribosomal oxygenases riox1 (no66) and riox2 (mina53) provide new insights into their evolution. *BMC Evol Biol*. 2018;18(1):96. doi 10.1186/s12862-018-1215-0

230. *Miao Q., Xu Y., Yin H., et al.* KRT8 phosphorylation regulates the epithelial‐mesenchymal transition in retinal pigment epithelial cells through autophagy modulation. *J Cell Mol Med*. 2020;24(5):3217-3228. doi 10.1111/jcmm.14998

231. *Delfín D.A.*, *DeAguero J.L.*, *McKown E.N.* The extracellular matrix protein abi3bp in cardiovascular health and disease. *Front Cardiovasc Med*. 2019;6:23. doi 10.3389/fcvm.2019.00023

232. *Presumey J.*, *Bialas A.R.*, *Carroll M.C.* Complement system in neural synapse elimination in development and disease. In: *Advances in Immunology*. V 135. Elsevier; 2017:53-79. doi 10.1016/bs.ai.2017.06.004

233. *Young K.A.*, *Biggins L.*, *Sharpe H.J.* Protein tyrosine phosphatases in cell adhesion. *Biochem J*. 2021;12;478(5):1061-1083. doi 10.1042/BCJ20200511.

234. *Easton A.C., Lourdusamy A., Loth E., et al.* CAMK2A polymorphisms predict working memory performance in humans. *Mol Psychiatry*. 2013;18(8):850-852. doi 10.1038/mp.2012.114

235. *Yasuda R.*, *Hayashi Y.*, *Hell J.W.* CaMKII: a central molecular organizer of synaptic plasticity, learning and memory. *Nat Rev Neurosci*. 2022;23(11):666-682. doi 10.1038/s41583-022-00624-2

236. *De Chaumont F., Icick R., Gorwood P., et al.* The influence of chrna5 and d398n missense variation on social and emotional behaviors in rodents and humans. *Transl Psychiatry*. 2025;15(1):507. doi 10.1038/s41398-025-03725-5

237. *Duric V., Banasr M., Licznerski P., et al.* A negative regulator of map kinase causes depressive behavior. *Nat Med*. 2010;16(11):1328-1332. doi 10.1038/nm.2219

238. *Talele S., Gonzalez S., Trudeau J., et al.* A phenotypic high-throughput screen identifies small molecule modulators of endogenous rgs10 in bv-2 cells. *J Med Chem*. 2024;67(22):20343-20352. doi 10.1021/acs.jmedchem.4c01738

239. *Peretz A., Gil‐Henn H., Sobko A., et al.* Hypomyelination and increased activity of voltage‐gated k+ channels in mice lacking protein tyrosine phosphatase ϵ. *EMBO J*. 2000;19(15):4036-4045. doi 10.1093/emboj/19.15.4036

240. *Li W., Chen Q., Peng C., et al.* Roles of the receptor for advanced glycation end products and its ligands in the pathogenesis of alzheimer's disease. *Int J Mol Sci*. 2025;26(1):403. doi 10.3390/ijms26010403

241. *Duester G.* Insufficient support for retinoic acid receptor control of synaptic plasticity through a non-genomic mechanism. *Front Neuroendocrinol*. 2023;71:101099. doi 10.1016/j.yfrne.2023.101099

242. *Yang H., Yuan Y., Yang K., et al.* ELK4 ameliorates cognitive impairment and neuroinflammation induced by obstructive sleep apnea. *Brain Res Bull*. 2024;216:111054. doi 10.1016/j.brainresbull.2024.111054

243. *Garza-Manero S., Sindi A.A.A., Mohan G., et al.* Maintenance of active chromatin states by hmgn2 is required for stem cell identity in a pluripotent stem cell model. *Epigenetics Chromatin*. 2019;12(1):73. doi 10.1186/s13072-019-0320-7

244. *Song M. lu, Sun Y. yun, Yin H. jun, et al.* P-coumaric acid alleviates neuronal damage in ischemic stroke mice by promoting bach1 nuclear export and degradation. *Acta Pharmacol Sin*. 2025;46(8):2136-2150. doi 10.1038/s41401-025-01510-0

245. *Cho C.*, *Smallwood P.M.*, *Nathans J.* Reck and gpr124 are essential receptor cofactors for wnt7a/wnt7b-specific signaling in mammalian cns angiogenesis and blood-brain barrier regulation. *Neuron*. 2017;95(5):1056-1073.e5. doi 10.1016/j.neuron.2017.07.031

246. *Seo E.Y., Jin S.P., Kim Y.K., et al.* Integrin-β4–tns4–focal adhesion kinase signaling mediates keratinocyte proliferation in human skin. *J Invest Dermatol*. 2017;137(3):763-766. doi 10.1016/j.jid.2016.10.039

247. *Feng L., Wang J., Jin H., et al.* SH3KBP1‐binding protein 1 prevents epidermal growth factor receptor degradation by the interruption of c‐cbl‐cin85 complex. *Cell Biochem Funct*. 2011;29(7):589-596. doi 10.1002/cbf.1792

248. *Bannykh K.S., Fuentes-Fayos A.C., Linesch P.W., et al.* Laminin beta 2 is localized at the sites of blood–brain barrier and its disruption is associated with increased vascular permeability, histochemical, and transcriptomic study. *J Histochem Cytochem*. 2024;72(10):641-667. doi 10.1369/00221554241281896

249. *Kimbrel N.A., Garrett M.E., Dennis M.F., et al.* A genome-wide association study of suicide attempts and suicidal ideation in u.s. military veterans. *Psychiatry Res*. 2018;269:64-69. doi 10.1016/j.psychres.2018.07.017

250. *Kari D.*, *Mijiti P.*, *Zou S.*, *Zhang P.* Study on the correlation between suicidal ideation and abi3bp gene、dpysl2 gene methylation in pediatric bipolar disorder with depressive episode. *Heliyon*. 2024;10(1):e23680. doi 10.1016/j.heliyon.2023.e23680

251. *Chester D.S., DeWall C.N., Derefinko K.J., et al.* Monoamine oxidase a (maoa) genotype predicts greater aggression through impulsive reactivity to negative affect. *Behav Brain Res*. 2015;283:97-101. doi 10.1016/j.bbr.2015.01.034

252. *Mcdermott R., Tingley D., Cowden J., et al. Monoamine Oxidase A Gene (MAOA) Predicts Behavioral Aggression Following Provocation*. 2009. www.pnas.org/cgi/content/full/

253. *Kalbitzer U., Roos C., Kopp G.H., et al.* Insights into the genetic foundation of aggression in papio and the evolution of two length-polymorphisms in the promoter regions of serotonin-related genes (5-httlpr and maoalpr) in papionini. *BMC Evol Biol*. 2016;16(1):121. doi 10.1186/s12862-016-0693-1

254. *Kent D., Marchetti L., Mikulasova A., et al.* Broad h3k4me3 domains: maintaining cellular identity and their implication in super‐enhancer hijacking. *BioEssays*. 2023;45(10):2200239. doi 10.1002/bies.202200239

255. *Mella C.*, *Figueroa C.D.*, *Otth C.*, *Ehrenfeld P.* Involvement of kallikrein-related peptidases in nervous system disorders. *Front Cell Neurosci*. 2020;14:166. doi 10.3389/fncel.2020.00166

256. *Christophi G.P., Isackson P.J., Blaber S., et al.* Distinct promoters regulate tissue‐specific and differential expression of kallikrein 6 in cns demyelinating disease. *J Neurochem*. 2004;91(6):1439-1449. doi 10.1111/j.1471-4159.2004.02826.x

257. *Pfeiffer S., Krüger J., Maierhofer A., et al.* Hypoxia-inducible factor 3a gene expression and methylation in adipose tissue is related to adipose tissue dysfunction. *Sci Rep*. 2016;6(1):27969. doi 10.1038/srep27969

258. *Dick K.J., Nelson C.P., Tsaprouni L., et al.* DNA methylation and body-mass index: a genome-wide analysis. *The Lancet*. 2014;383(9933):1990-1998. doi 10.1016/S0140-6736(13)62674-4

259. *Mikec Š., Šimon M., Morton N.M., et al.* Genetic variants of the hypoxia‐inducible factor 3 alpha subunit (hif3a) gene in the fat and lean mouse selection lines. *Mol Biol Rep*. 2022;49(6):4619-4631. doi 10.1007/s11033-022-07309-0

260. *Yap S.Q.*, *Kim W.D.*, *Huber R.J.* Mfsd8 modulates growth and the early stages of multicellular development in dictyostelium discoideum. *Front Cell Dev Biol*. 2022;10:930235. doi 10.3389/fcell.2022.930235

261. *Tuladhar S.*, *Kanneganti T.D.* NLRP12 in innate immunity and inflammation. *Mol Aspects Med*. 2020;76:100887. doi 10.1016/j.mam.2020.100887

262. *Huang L., Tao Y., Wu X., et al.* The role of nlrp12 in inflammatory diseases. *Eur J Pharmacol*. 2023;956:175995. doi 10.1016/j.ejphar.2023.175995

263. *Saad M., Lesage S., Saint-Pierre A., et al.* Genome-wide association study confirms bst1 and suggests a locus on 12q24 as the risk loci for parkinson's disease in the european population. *Hum Mol Genet*. 2011;20(3):615-627. doi 10.1093/hmg/ddq497

264. *Li J., Luo J., Liu L., et al.* The association between cd157/bst1 polymorphisms and the susceptibility of parkinson’s disease: a meta-analysis. *Neuropsychiatr Dis Treat*. 2019;Volume 15:1089-1102. doi 10.2147/NDT.S190935

265. *Lopatina O., Yoshihara T., Nishimura T., et al.* Anxiety- and depression-like behavior in mice lacking the cd157/bst1 gene, a risk factor for parkinson's disease. *Front Behav Neurosci*. 2014;8. doi 10.3389/fnbeh.2014.00133

266. *Caglar C.*, *Friedman J.* Restriction of food intake by ppp1r17-expressing neurons in the dmh. *Proc Natl Acad Sci*. 2021;118(13):e2100194118. doi 10.1073/pnas.2100194118

267. *Tokizane K.*, *Brace C.S.*, *Imai S. ichiro*. DMHPpp1r17 neurons regulate aging and lifespan in mice through hypothalamic-adipose inter-tissue communication. *Cell Metab*. 2024;36(2):377-392.e11. doi 10.1016/j.cmet.2023.12.011

268. *Allen P.B., Greenfield A.T., Svenningsson P., et al.* Phactrs 1–4: a family of protein phosphatase 1 and actin regulatory proteins. *Proc Natl Acad Sci*. 2004;101(18):7187-7192. doi 10.1073/pnas.0401673101

269. *Komiyama T., Iwama H., Osada N., et al.* Dopamine receptor genes and evolutionary differentiation in the domestication of fighting cocks and long-crowing chickens. Gilbert T, ed. *PLoS ONE*. 2014;9(7):e101778. doi 10.1371/journal.pone.0101778

270. *Cagan A.*, *Blass T.* Identification of genomic variants putatively targeted by selection during dog domestication. *BMC Evol Biol*. 2016;16(1):10. doi 10.1186/s12862-015-0579-7

271. *Proskura W.S., Frost A., Gugała L., et al.* Genetic background of aggressive behaviour in dogs. *Acta Vet Brno*. 2013;82(4):441-445. doi 10.2754/avb201382040441

272. *Fawcett J.W., Fyhn M., Jendelova P., et al.* The extracellular matrix and perineuronal nets in memory. *Mol Psychiatry*. 2022;27(8):3192-3203. doi 10.1038/s41380-022-01634-3

273. *Mueller-Buehl C., Pakusch J., Bader V., et al.* Combined loss of brevican, neurocan, tenascin-c and tenascin-r leads to impaired fear retrieval due to perineuronal net loss. *Sci Rep*. 2025;15(1):5528. doi 10.1038/s41598-025-89580-2

274. *Thompson E.H., Lensjø K.K., Wigestrand M.B., et al.* Removal of perineuronal nets disrupts recall of a remote fear memory. *Proc Natl Acad Sci*. 2018;115(3):607-612. doi 10.1073/pnas.1713530115

Приложение А

Список образцов с фенотипами, используемых в полногеномном анализе

Образец	Поведение	Образец	Поведение	Образец	Поведение
S000242	агрессивное	S000241	дружелюбное	S000293	неселектированное
S000275	агрессивное	S000290	дружелюбное	S000329	неселектированное
S000332	агрессивное	S000292	дружелюбное	S000331	неселектированное
S000339	агрессивное	S000324	дружелюбное	S000334	неселектированное
S000340	агрессивное	S000326	дружелюбное	S000354	неселектированное
S000364	агрессивное	S000327	дружелюбное	S000355	неселектированное
S000377	агрессивное	S000342	дружелюбное	S000377	неселектированное
S000436	агрессивное	S000346	дружелюбное	S000428	неселектированное
S000511	агрессивное	S000347	дружелюбное	S000439	неселектированное
S000526	агрессивное	S000348	дружелюбное	S000440	неселектированное
FBP109	агрессивное	S000352	дружелюбное	S000452	неселектированное
FBP113	агрессивное	S000427	дружелюбное	FBP110	неселектированное
FBP116	агрессивное	FBP112	дружелюбное	FBP114	неселектированное
FBP117	агрессивное	FBP119	дружелюбное	FBP115	неселектированное
FBP124	агрессивное	FBP120	дружелюбное	FBP118	неселектированное
				FBP127	неселектированное

Приложение Б

Список образцов с фенотипами, используемых в транскриптомном анализе.

Образец	Фенотип
FB17	ручной
FBP75	ручной
FBP82	ручной
FBP83	ручной
FBP84	ручной
FBP91	ручной
FBP85	агрессивный
FBP86	агрессивный
FBP76	агрессивный
FBP77	агрессивный
FBP78	агрессивный
FBP97	агрессивный
FBP79	неселектированный
FBP80	неселектированный
FBP81	неселектированный
FBP88	неселектированный
FBP89	неселектированный
FBP90	неселектированный

Приложение В
Статистика секвенирования геномных библиотек.

образец	поведение	кол-во прочтений	картировано	кол-во дупликатов	средняя глубина прочтения
S000241	ручное	1093086135	99.44%	4.62%	40,18
S000242	агрессивное	911003003	99.41%	8.24%	31,75
S000275	агрессивное	387351080	99.21%	44.07%	7,96
S000290	ручное	174938897	98.91%	25.10%	4,83
S000292	ручное	231919885	99.20%	42.58%	4,94
S000293	неселектированное	290323676	99.31%	48.50%	5,49
S000324	ручное	139904185	99.19%	18.92%	4,22
S000326	ручное	162504421	98.79%	4.39%	5,51
S000327	ручное	121256689	98.50%	4.86%	4,13
S000329	неселектированное	134730446	98.78%	5.15%	4,61
S000331	неселектированное	150294905	95.30%	4.68%	5,07
S000332	агрессивное	177336514	97.59%	4.80%	6,23
S000334	неселектированное	146765833	98.70%	2.61%	5,13
S000339	агрессивное	120212340	98.53%	1.67%	4,26
S000340	агрессивное	142094201	97.26%	0.56%	5,18
S000342	ручное	170219749	90.53%	19.54%	4,32
S000346	ручное	111632638	99.18%	0.56%	4,18
S000347	ручное	129113924	99.22%	0.94%	4,83
S000348	ручное	128225348	99.18%	1.04%	4,75
S000352	ручное	116939893	99.06%	0.94%	4,27
S000354	неселектированное	113731682	99.11%	0.64%	4,25
S000355	неселектированное	134986610	98.95%	0.73%	4,91
S000363	агрессивное	137648237	98.92%	0.74%	4,94
S000364	агрессивное	130392637	99.15%	0.82%	4,94
S000377	ручное	116434899	99.12%	0.96%	4,29

S000427	ручное	352246087	98.88%	4.20%	12,14
S000428	неселектированное	148059187	98.70%	0.86%	5,43
S000436	агрессивное	198435110	92.67%	2.10%	6,41
S000439	неселектированное	155991057	96.00%	1.80%	5,45
S000440	неселектированное	148309809	96.82%	1.41%	5,43
S000452	неселектированное	127237221	94.54%	2.43%	4,44
S000511	агрессивное	238743474	69.26%	6.25%	5,26
S000526	агрессивное	241838305	99.07%	1.07%	8,54
S003257	агрессивное	345231067	98.74%	6.28%	17,49
S003258	неселектированное	273558620	99.02%	6.46%	14,03
S003260	ручное	352751301	99.09%	6.43%	18,20
S003261	агрессивное	367812307	99.11%	7.31%	18,82
S003262	неселектированное	318133438	99.07%	6.49%	16,40
S003263	неселектированное	340055691	99.04%	7.11%	17,19
S003264	агрессивное	359886555	99.06%	6.58%	18,43
S003265	агрессивное	435754848	99.08%	7.99%	22,06
S003266	неселектированное	308200167	99.12%	7.13%	15,69
S003267	ручное	377400297	98.99%	7.31%	19,00
S003268	ручное	411854297	99.06%	7.53%	20,69
S003381	агрессивное	322178192	99.04%	7.59%	16,31
S003382	неселектированное	294267678	99.02%	6.87%	15,04

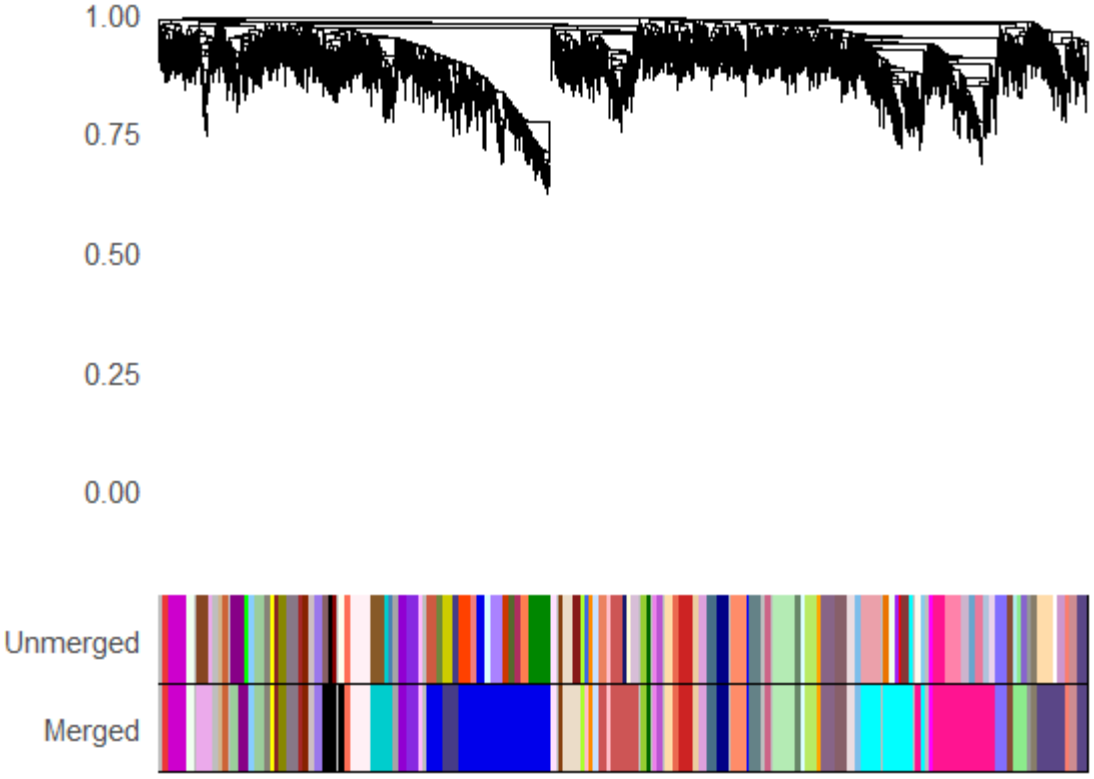
Приложение Г

Множественные выравнивания белковых последовательностей, содержащих миссенс-замены, специфичные для агрессивных линий, для 10 организмов



Приложение Д

Дендрограммы построения генных сетей для (а) фронтальной коры и (б) миндалевидного тела. По оси ординат отмечена высота дендрограмм



Приложение Е

Состав модулей коэкспрессии генов с наибольшим по модулю коэффициентом корреляции при сравнении фронтальной коры агрессивных с ручными и неселектированными животными.

Модуль *green*:

GRM3, SHTN1, SEMA3E, SLC8A1, CYP26B1, KREMEN1, NXPH3, ADAMTSL1, TLL2, LOC112922083, SCAI, KYAT3, LOC112930486, DRD4, NPPA, LOC140594757, SCN11A, LOC140594679, LOC140594696, XKR9.

Модуль tan:

STPG3, PTGS1, TMEM132C, SEMA3G, HAS1, LOC140599430, FAH, NINJ2, MATK, ALKBH2, LOC140594373, LOC140594463, ADRA2C, LOC140595508, CHST7

Модуль lightyellow:

NXNL2, SMIM22, PCMT1, ATL1, INKA2, PSMA4, ESR2, CAMKK2, WNT5B, LOC112928492, LOC140599777, LOC140594237, PKD2, C12H9orf152, RASL10A, RYBP, SIRT5, LOC112935618, CHSY3, LOC140593714, TMEM114, MLPH, LOC112918098, LOC112923029, SHISA6, LOC112931924, SPTLC3, DSTN, LOC140595820, LOC112928865

Приложение Ж

Состав модулей коэкспрессии генов с наибольшим по модулю коэффициентом корреляции при сравнении миндалевидного тела агрессивных с ручными и неселектированными животными.

Модуль lightblue4:
LOC140598592, LOC112931228, ACAT2, PACSIN1, NAT9, LOC140599376, GRAMD1A, CENPX, LOC140597728, LOC140597858, LOC140595525, LOC112929733, HAP1, LOC140598240, LOC140597976, LOC112935468, LOC140598166, DNAAF19, LOC140597956, AP4M1, LOC112918227, LOC112908325, LOC140598621, LOC112927216, SORT1, AOPEP, C4H4orf36, DYDC2, LOC112909334, WFS1, POPDC1, PHKG2, LOC112929849, ITPR3, LOC112928397, CREB3L1, MCOLN3, SYT15B, PSMD13, MSRA, RAD1, FERRY3, LOC112930687, LOC112915081, CRIM1, NEFM, LOC140598646, MGST3, BSN, LOC112920480, LOC140594651, HSPB7, FST, LOC112926766, PAK1IP1, LOC140595102, FKBPL, LOC140594532, C7H8orf48, ABRA, LOC112925892, GPR37L1, LOC112909137, FANCI, NUDT15, LOC112915452, LOC140595786, SAAL1, SPATS2, YWHAZ, LOXHD1, MIIP, IL18, RNF112, LOC112915244, LOC112919085, LOC112933727, ATP6V1G1, LOC112928791, LOC112922264, NDUFB3, RAB15, CLEC2L, LOC112930693, CCL21, LYRM4, LOC112925880, TNFRSF9, TRPM1, LGI2, LOC140595704, LOC112922524, FOXRED2, LOC112909192, LOC140596224, TRPC5, RTL4, LOC140596289, G6PD

Модуль navajowhite1:
HAS1, LOC112910787, LOC140597816, TFAP2D, MAN1B1, CCN6, NLRP12, LOC112935140, LOC140598209, MFSD8, LOC140598362, LOC140598492, LOC140593923, LOC112930987, MPZ, LOC112912849, LOC112929582, C1QTNF1, ALPL, ARAP3, LOC112907366, LOC140598369, LOC140598396, CD69, LOC140600068, PEX1, LOC112925536, PPP1R17, LOC112916370, P2RX3, LOC140600075, ESS2, MINPP1, IQCM, LOC112927615, CMTM7, DIPK2A, LOC112910881, ENG, TMC5, LOC140599693, LOC112907500, LOC112923779, IFT74, MRPL19, PTS, LOC140594999, ARSB, LOC140599886, FMO1, HAVCR2, GALE, TRIM14, ADSS2, FN1, LOC112917240, LOC140593830, LOC112929909, LOC112918751, FMO3, LOC140595388, BMP2, CTXND1, LOC112920035, ZNF485, LOC140595689, LOC112928703, LOC112922585, LOC112922871.

Модуль blue4:
CTU1, TNNI3, ALDH1A1, LOC112913606, GAA, STKLD1, CFAP77, KIT, LOC140597999, LOC112929743, FBXO2, CYTIP, SP5, LOC112926703, ACHE, ZNF771, ATP2A1, LOC140598085, BCAR3, TBC1D9, LOC140598506, ADH4, FGFR4, GYPC, KYNU, TTC9, VSX2, TSHR, LOC112930351, LOC112912930, LOC112907033, LOC140599527, LOC140599791, KCNH2, CFAP96, RESF1, USP18, LOC112907724, LOC112926784, GJA3, LOC140593877, LAMB2, TLE2, FGF22, LOC140594118, LOC112935181, LOC140593698, RNLS, GRIK3, TPPP2, SALL2, LOC140594238, HTATIP2, MRPL48, LOC140594558, LOC112927584, SPIDR, LOC140594677, LOC140594918, ZYG11A, MEAK7, PKD1L3, LOC140595033, LOC140595128, GNAL, LOC112925275, LOC140595077, MIXL1, ITPKB, LOC112923472, MMP24OS, ADGRV1, LOC112915617, RGMA, LOC140595466, BST1, YBEY, RIPK4, LOC140595748, PIGP, HDC, SEC31B, ECHS1, TMEM169, RPS19BP1, LOC140596253, ADGRG2, LYPD5, THAP8, FNDC1, VAT1, LGALS9, LOC140597768, PPM1N, LOC140595438, MMUT, LOC140598330, NOTUM, LOC140593665, SOX2, SEMA3C, IDO2, CPVL, PADI2, KLK8, CCN2, LOC112918520, LOC112934237, TENM4, GCNT2.

Приложение 3

Статистика секвенирования ChIP-seq библиотек, приготовленных из ДНК фронтальной коры головного мозга лис. Необогащенная ДНК - образец неиммунопреципитированной ДНК, использованный в качестве отрицательного контроля

Образец	Гистоновая модификация	Число прочтений	% картирования
FB17	H3K27ac	33 960 343	94,09
	H3K4me3	74 724 841	94,69
FBP75	H3K27ac	18 478 613	99,41
	H3K4me3	17 442 613	98,91
FBP76	H3K27ac	15 826 652	99,44
	H3K4me3	16 671 141	97,44
FBP77	H3K27ac	34 435 997	96,62
	H3K4me3	122 555 384	94,93
FBP78	H3K27ac	15 784 029	97,61
	H3K4me3	31 729 120	97,52
FBP79	H3K27ac	74 223 119	94,17
	H3K4me3	25 950 068	95,36
FBP80	H3K27ac	10 207 732	96,54
	H3K4me3	24 659 340	94,7
FBP81	H3K27ac	40 080 355	94,99
	H3K4me3	34 621 799	91,61
FBP82	H3K27ac	54 216 472	94,01
	H3K4me3	30 145 543	95,61
FBP83	H3K27ac	37 462 857	93,13
	H3K4me3	16 270 849	98,92
FBP84	H3K27ac	45 861 925	95,57
	H3K4me3	39 705 980	94,31
FBP85	H3K27ac	62 216 525	95,12
	H3K4me3	38 924 698	94,75

	необогащенная ДНК	50 005 474	96,5
FBP86	H3K27ac	16 692 330	99,25
	H3K4me3	16 088 581	96
FBP88	H3K27ac	20 868 503	99,16
	H3K4me3	18 295 362	98,37
FBP89	H3K27ac	20 908 970	99,47
	H3K4me3	18 878 557	98,78
FBP90	H3K27ac	14 807 306	98,14
	H3K4me3	16 312 310	96,59
FBP91	H3K27ac	8 650 252	98,56
	H3K4me3	16 389 464	97,36
FBP97	H3K27ac	11 198 347	98,99
	H3K4me3	28 471 812	98,13

Приложение И

Аннотация локусов с одновременным дифференциальным уровнем метилирования и ацетилирования гистона H3, находящихся на расстоянии не более 3000 пн от TSS аннотированных генов лисы. Сравнение агрессивных особей (A) с ручными (T) и неселекционированными (N), logFCcp - усредненное значение logFC для 2 сравнений по 2 гистоновым меткам

информация о гене				H3K4me3					H3K27ac					расстояние до TSS, пн
				сравнение	AvsT		AvsN		сравнение	AvsT		AvsN		
название	тип	цепь	logFC ср	локус	p-value	logF C	p-value	logFC	локус	p-value	log FC	p-value	log FC	
CPOX	белок-кодирующий	-	0,30	NC_132780.1:102900880-102905281	3,36E-02	0,19	6,12E-03	0,27	NC_132780.1:102881410-102903729	4,25E-03	0,33	2,67E-04	0,43	0
KLK6	белок-кодирующий	-	0,41	NC_132780.1:16928353-16931054	1,80E-02	0,38	4,35E-03	0,52	NC_132780.1:16928221-16931892	5,01E-03	0,41	3,53E-02	0,31	0
BBS2	белок-кодирующий	+	-0,39	NC_132781.1:139363387-139367835	4,03E-03	-0,3	1,32E-02	-0,27	NC_132781.1:139366191-139367909	3,23E-02	-0,47	2,39E-02	-0,51	0
MT4	белок-кодирующий	-	-0,39	NC_132781.1:139363387-139367835	4,03E-03	-0,3	1,32E-02	-0,27	NC_132781.1:139366191-139367909	3,23E-02	-0,47	2,39E-02	-0,51	0
EGFL7	псевдоген	-	-0,65	NC_132781.1:148056630-148057291	4,27E-02	-0,69	2,29E-02	-0,88	NC_132781.1:148055261-148058547	1,75E-02	-0,47	6,23E-03	-0,55	2493
MFAP2	белок-кодирующий	+	-0,97	NC_132781.1:160912728-160914385	2,64E-02	-0,5	4,62E-02	-0,5	NC_132781.1:160913183-160913938	2,62E-04	-1,4	4,56E-04	-1,49	-387
BTBD8	белок-кодирующий	-	0,28	NC_132782.1:131721873-131728182	2,98E-02	0,17	3,11E-03	0,25	NC_132782.1:131714035-131728607	9,17E-03	0,31	1,45E-03	0,38	0
LOC140598375	днРНК	+	0,28	NC_132782.1:131721873-131728182	2,98E-02	0,17	3,11E-03	0,25	NC_132782.1:131714035-131728607	9,17E-03	0,31	1,45E-03	0,38	0
LOC140598188	днРНК	-	0,42	NC_132782.1:20826565-20833923	1,67E-03	0,37	3,82E-05	0,57	NC_132782.1:20828255-20836618	3,34E-02	0,26	2,66E-04	0,47	0
HOXD1	белок-кодирующий	+	0,42	NC_132782.1:20826565-20833923	1,67E-03	0,37	3,82E-05	0,57	NC_132782.1:20828255-20836618	3,34E-02	0,26	2,66E-04	0,47	0
CLDN11	белок-кодирующий	-	0,38	NC_132782.1:39282596-39294412	8,66E-03	0,43	5,21E-03	0,49	NC_132782.1:39254961-39290489	1,95E-02	0,26	1,90E-03	0,36	0

ST6GAL1	белок-кодирующий	-	-0,38	NC_132782.1:54271821-54275140	8,17E-03	-0,31	1,62E-02	-0,31	NC_132782.1:54271558-54275086	4,93E-03	-0,45	6,37E-03	-0,44	1266
TAF5L	белок-кодирующий	+	0,29	NC_132783.1:69004463-69009342	2,65E-02	0,18	9,55E-03	0,23	NC_132783.1:68975476-69010114	8,82E-03	0,3	8,37E-05	0,46	0
URB2	белок-кодирующий	-	0,29	NC_132783.1:69004463-69009342	2,65E-02	0,18	9,55E-03	0,23	NC_132783.1:68975476-69010114	8,82E-03	0,3	8,37E-05	0,46	0
COMMD9	белок-кодирующий	-	0,40	NC_132784.1:84733543-84737623	3,01E-02	0,34	5,55E-04	0,64	NC_132784.1:84732093-84740770	3,07E-02	0,28	6,48E-03	0,35	0
PRR5L	белок-кодирующий	+	0,40	NC_132784.1:84733543-84737623	3,01E-02	0,34	5,55E-04	0,64	NC_132784.1:84732093-84740770	3,07E-02	0,28	6,48E-03	0,35	0
LOC140599275	днРНК	+	-0,94	NC_132785.1:957294-958280	4,90E-02	-0,39	3,81E-02	-0,44	NC_132785.1:957563-957760	9,95E-03	-1,5	1,79E-02	-1,44	0
LOC140599653	днРНК	+	0,38	NC_132786.1:26174476-26176648	9,94E-03	0,38	2,44E-03	0,48	NC_132786.1:26171983-26177467	3,67E-02	0,29	7,55E-03	0,37	0
DDHD2	белок-кодирующий	-	0,29	NC_132786.1:31629454-31634352	4,85E-02	0,15	2,18E-03	0,25	NC_132786.1:31628626-31637494	7,75E-03	0,33	8,61E-04	0,42	0
ANGPT2	белок-кодирующий	+	-0,32	NC_132786.1:60792478-60797089	4,54E-02	-0,16	5,88E-04	-0,31	NC_132786.1:60792878-60794968	1,83E-02	-0,44	4,22E-02	-0,39	0
ZC3HAV1	белок-кодирующий	+	-0,55	NC_132786.1:9959993-9963195	3,49E-02	-0,24	6,09E-03	-0,34	NC_132786.1:9960034-9961056	1,21E-02	-0,84	3,03E-02	-0,79	0
LOC140599968	белок-кодирующий	-	-1,14	NC_132787.1:55002852-55003914	1,49E-04	-1,38	2,17E-02	-0,92	NC_132787.1:55001840-55006570	6,87E-09	-1,38	1,49E-04	-0,9	2575
LOC140593724	псевдоген	-	1,38	NC_132787.1:90283939-90286896	5,31E-04	1,13	7,32E-03	0,82	NC_132787.1:90284962-90286435	9,98E-09	2,21	2,94E-05	1,36	0
LOC140600075	белок-кодирующий	+	1,38	NC_132787.1:90283939-90286896	5,31E-04	1,13	7,32E-03	0,82	NC_132787.1:90284962-90286435	9,98E-09	2,21	2,94E-05	1,36	0
PEX11G	белок-кодирующий	-	0,99	NC_132788.1:103723868-103726541	1,78E-04	0,69	2,58E-04	0,72	NC_132788.1:103721105-103726553	3,25E-02	1,18	2,01E-02	1,38	0
ZNF358	белок-кодирующий	+	0,99	NC_132788.1:103723868-103726541	1,78E-04	0,69	2,58E-04	0,72	NC_132788.1:103721105-103726553	3,25E-02	1,18	2,01E-02	1,38	0
TUBB4A	белок-кодирующий	+	0,77	NC_132788.1:105100282-105105662	6,72E-10	1,4	4,54E-07	1,07	NC_132788.1:105099946-105113031	3,43E-03	0,34	1,03E-02	0,3	2120

LGI3	белок-кодирующий	-	0,29	NC_132788.1:16586667-16589934	2,88E-02	0,25	7,29E-03	0,34	NC_132788.1:16576434-16598360	1,25E-02	0,29	2,30E-02	0,26	0
DOCK5	белок-кодирующий	+	0,31	NC_132788.1:19419678-19424837	4,20E-02	0,19	1,34E-03	0,34	NC_132788.1:19419913-19444655	2,18E-02	0,27	2,10E-04	0,44	0
LOC140594221	днРНК	+	-0,80	NC_132789.1:66171330-66173062	3,75E-02	-0,6	3,80E-02	-0,66	NC_132789.1:66171187-66172109	4,22E-03	-0,85	5,09E-04	-1,09	-4
LOC112911331	днРНК	+	0,73	NC_132790.1:99351725-99354366	3,25E-03	0,51	8,40E-06	1,06	NC_132790.1:99351025-99354176	4,93E-02	0,31	2,95E-08	1,02	0
LOC112911947	днРНК	-	0,40	NC_132791.1:101908166-101913057	4,72E-02	0,3	1,05E-03	0,56	NC_132791.1:101908149-101913286	1,43E-02	0,35	4,40E-03	0,41	348
ACTL7B	белок-кодирующий	+	-0,70	NC_132791.1:10458132-10461555	6,99E-03	-0,4	2,70E-02	-0,35	NC_132791.1:10459270-10460385	1,65E-03	-0,92	1,59E-04	-1,13	-502
ANKRD65	белок-кодирующий	-	0,40	NC_132791.1:157974771-157976284	3,00E-02	0,52	3,19E-02	0,59	NC_132791.1:157970314-157997143	2,49E-02	0,25	4,37E-02	0,23	176
LOC140594598	днРНК	+	-2,47	NC_132791.1:67773158-67773641	3,63E-05	-2,87	3,14E-04	-2,73	NC_132791.1:67773270-67773601	2,22E-04	-2,47	1,27E-02	-1,82	0
FEM1C	белок-кодирующий	+	0,34	NC_132791.1:68537425-68540792	1,11E-02	0,29	2,80E-03	0,38	NC_132791.1:68526677-68543450	1,37E-02	0,3	1,17E-03	0,4	0
LOC140594612	белок-кодирующий	-	-0,58	NC_132791.1:76902747-76908044	6,12E-05	-0,59	7,14E-03	-0,37	NC_132791.1:76902964-76908159	3,62E-06	-0,73	8,93E-05	-0,62	0
LOC112920442	белок-кодирующий	-	-0,58	NC_132791.1:76902747-76908044	6,12E-05	-0,59	7,14E-03	-0,37	NC_132791.1:76902964-76908159	3,62E-06	-0,73	8,93E-05	-0,62	0
LOC140594962	днРНК	-	0,70	NC_132792.1:116800874-116803218	1,34E-03	0,73	4,89E-03	0,68	NC_132792.1:116800592-116802575	4,40E-04	0,83	1,34E-02	0,59	0
FMN2	белок-кодирующий	-	-0,35	NC_132792.1:128767319-128776337	1,26E-02	-0,18	2,49E-04	-0,28	NC_132792.1:128768114-128771239	2,70E-02	-0,37	1,35E-03	-0,55	0
FMN2	белок-кодирующий	-	-0,32	NC_132792.1:128767319-128776337	1,26E-02	-0,18	2,49E-04	-0,28	NC_132792.1:128772302-128777017	3,37E-02	-0,34	3,70E-03	-0,48	0
RUNX1T1	белок-кодирующий	+	-1,05	NC_132792.1:42095488-42101797	2,11E-03	-0,33	3,76E-04	-0,41	NC_132792.1:42097434-42097851	3,19E-02	-1,46	5,74E-03	-1,99	2741
LOC140595371	днРНК	-	0,50	NC_132793.1:26510339-26512067	2,59E-02	0,53	3,36E-03	0,81	NC_132793.1:26501667-26514015	4,77E-02	0,23	8,79E-04	0,41	0

LARP6	белок-кодирующий	-	0,31	NC_132794.1:74312102-74316229	3,24E-02	0,24	5,74E-03	0,35	NC_132794.1:74314912-74333599	1,72E-02	0,27	1,91E-03	0,36	0
KCNIP2	белок-кодирующий	-	-0,72	NC_132794.1:91812725-91817604	4,07E-02	-0,22	1,38E-02	-0,28	NC_132794.1:91817488-91817738	8,23E-03	-1,31	4,24E-02	-1,06	0
LOC112929931	днРНК	+	0,48	NC_132796.1:36613943-36619225	2,01E-03	0,41	9,01E-06	0,68	NC_132796.1:36613403-36620558	4,74E-02	0,26	4,32E-05	0,56	0
VLDLR	белок-кодирующий	-	0,24	NW_027325769.1:154597-159506	1,39E-02	0,17	1,49E-02	0,18	NW_027325769.1:140674-158904	4,92E-02	0,23	6,32E-04	0,41	0
LOC112922134	белок-кодирующий	+	0,24	NW_027325769.1:154597-159506	1,39E-02	0,17	1,49E-02	0,18	NW_027325769.1:140674-158904	4,92E-02	0,23	6,32E-04	0,41	0

Приложение К

Пересечение локусов с дифференциальным уровнем метилирования или ацетилирования гистона H3, находящихся на расстоянии не более 3000 пн от TSS аннотированных генов лисы, с координатами ОНВ, специфичных для агрессивной линии. Сравнение агрессивных особей (A) с ручными (T) и неселектированными (N)

	сравнение	AvsT		AvsN				
гистоновая метка	координаты локуса	p-value	logFC	p-value	logFC	координаты ОНВ	цепь	ген
H3K27ac	NC_132780.1:13333879-13335142	9,41E-03	-0,72	1,22E-02	-0,71	NC_132780.1:13334823	+	*HIF3A*
H3K27ac	NC_132780.1:16928221-16931892	5,01E-03	0,41	3,53E-02	0,31	NC_132780.1:16928678	-	*KLK6*
H3K4me3	NC_132780.1:16928353-16931054	1,80E-02	0,38	4,35E-03	0,52		-	
H3K27ac	NC_132780.1:19874111-19875771	2,09E-02	-0,66	8,9E-04	-1,02	NC_132780.1:19875697	-	*RPS9*
							+	*LOC112909393*
						NC_132780.1:19875703	-	*RPS9*
							+	*LOC112909393*
						NC_132780.1:19875713	-	*RPS9*
							+	*LOC112909393*
H3K27ac	NC_132783.1:31388069-31486384	3,40E-02	0,23	7,27E-04	0,81	NC_132783.1:31424848	-	*LOC140598570*
							+	*LOC140598569*
							-	*UGT8*
						NC_132783.1:31446672	-	*LOC140598570*
							+	*LOC140598569*
							-	*UGT8*
H3K27ac	NC_132788.1:13495201-13515425	3,81E-03	0,34	3,77E-05	0,5	NC_132788.1:13512230	+	*DOCK10*

H3K27ac	NC_132793.1:100472827-100477196	2,74E-02	-0,36	3,36E-05	-0,7	NC_132793.1:100475602	+	*LOC112921736*
H3K27ac	NC_132794.1:105825376-105827090	1,05E-02	0,62	1,23E-03	0,81	NC_132794.1:105826213	-	*RAB11FIP2*
						NC_132794.1:105826850	-	*RAB11FIP2*
						NC_132794.1:105826856	-	*RAB11FIP2*