

На правах рукописи



Кононкова Анна Дмитриевна

**Структура хроматина дрозофилы в контексте влияния белков
ядерной периферии и процессов, ассоциированных с ранними
стадиями сперматогенеза**

Специальность 1.5.8 —

«математическая биология, биоинформатика»

АВТОРЕФЕРАТ

диссертации на соискание ученой
степени кандидата биологических наук

Москва — 2024

Работа выполнена в Автономной некоммерческой образовательной организации высшего образования "Сколковский институт науки и технологий"

Научный руководитель:

Храмеева Екатерина Евгеньевна, доктор биологических наук, доцент (Автономная некоммерческая образовательная организация высшего образования "Сколковский институт науки и технологий")

Официальные оппоненты:

Шайтан Алексей Константинович, доктор физико-математических наук, профессор РАН, член-корреспондент РАН (Биологический факультет Московского государственного университета имени М. В. Ломоносова)

Ерохин Максим Максимович, доктор биологических наук, ведущий научный сотрудник (Федеральное государственное бюджетное учреждение науки «Институт биологии гена Российской академии наук»)

Ведущая организация:

Федеральное государственное бюджетное учреждение науки «Институт общей генетики им. Н. И. Вавилова Российской академии наук»

Защита состоится 27 января 2025 года в 15:00 часов на заседании диссертационного совета 24.1.101.01 при Институте проблем передачи информации имени А.А. Харкевича Российской академии наук (ИППИ РАН) по адресу: 127051, г. Москва, Большой Каретный переулок, д. 19, стр. 1.

С диссертацией можно ознакомиться в библиотеке ИППИ РАН, а также на сайте ИППИ РАН по адресу: http://iitp.ru/upload/content/1716/KAD_disser.pdf

Автореферат разослан « » 2025 года

Ученый секретарь диссертационного совета 24.1.101.01 д.б.н. Казенников Олег Васильевич



Общая характеристика работы

Актуальность темы. Исследования последних десятилетий говорят о том, что существенную роль в функциональной регуляции клетки играет пространственная упаковка ДНК, которая в свою очередь также является объектом регуляции. Подчиняясь базовым принципам трехмерной организации, структура хроматина варьируется в разных организмах, типах клеток, стадиях развития и иногда внешних условиях. Закономерности этих изменений, являясь предметом интенсивных исследований, остаются изученными не в полной мере, что объясняется не только масштабностью данной тематики, но и наличием определенных ограничивающих факторов. Выбор дрозофилы¹ в качестве экспериментального объекта отчасти снимает эти ограничения. Например, согласно имеющимся данным, процессы, связанные с клеточным старением, затрагивают также и структуру хроматина, а в число потенциальных белков-кандидатов, вовлеченных во взаимосвязь между старением и трехмерной структурой генома, входят белки периферии. Одним из таких белков является Elys. Elys способен связываться с хроматином и входит в состав ядерной поры, что делает его интересным объектом в вышеупомянутом контексте. Однако исследование влияния этого белка на компактизацию ДНК предполагает необходимость подавления его экспрессии. В этом смысле клеточная линия S2 дрозофилы дает уникальную возможность подобного рода экспериментов, т. к., в отличие от большинства других типов клеток, ее жизнеспособность сохраняется. Анализ данных до и после деплеции Elys, полученных на основе протокола Hi-C, позволяет подробно изучить влияние данного белка на структуру хроматина с разными свойствами и оценить изменения на разных масштабах взаимодействий.

Изучение закономерностей трехмерной организации хроматина включает поиск ассоциированных с ее изменениями факторов, и одним из них является экспрессия. Учитывая место, которое занимает согласованная и точная работа генов в функционировании организма, исследование взаимного влияния пространственной организации хроматина и транскрипции вызывает особенный интерес. Подходы к изучению этой взаимосвязи часто основываются на искусственном подавлении транскрипции или на сравнении разных типов клеток, однако, ввиду сложности получения такого материала из образцов тканей и последующего проведения протокола Hi-C, набор доступных на сегодня данных существенно ограничен: значительная доля исследований по структуре хроматина проводится на клеточных линиях, которые заметно отличаются по эпигенетическим и функциональным характеристикам от клеток, выделенных из биологических образцов. В этом отношении процесс сперматогенеза у дрозофилы

¹ Здесь и далее под дрозофилой следует понимать *Drosophila melanogaster*

оказывается особенно ценным для исследователя, т. к. на ранних его этапах, при переходе от стадии сперматогониев к стадии сперматоцитов, имеет место масштабная активация сперматоцитспецифичной транскрипции (активируется не менее 1000 генов), а возможность выведения мутантных особей у данного вида позволяет получить данные Hi-C, соответствующие двум этим стадиям. Поскольку большая часть подобных исследований основывается на экспериментах с подавлением транскрипции, такая постановка задачи потенциально может привести к дополняющим сложившуюся картину результатам.

Процесс сперматогенеза в дрозофиле интересен не только с точки зрения активации большого числа генов: для него характерно увеличение объема ядра клетки на стадии сперматоцитов. Каким образом это сказывается на структуре хроматина до конца не известно. Исследования такого рода ограничиваются лишь несколькими публикациями, в том числе с искусственным раздуванием ядра в клетках млекопитающих. Вместе с тем взаимосвязь между структурой хроматина и объемом ядра является высоко приоритетной темой для подробного изучения в силу потенциальной ассоциации с процессом старения.

Обобщая все вышеизложенное, можно сказать, что анализ данных Hi-C двух начальных стадий сперматогенеза дрозофилы дает возможность изучения сразу нескольких факторов в контексте их взаимосвязи с трехмерной организацией хроматина. Нуклеопорин Elys, будучи важной составляющей комплекса ядерной поры, также является потенциальным кандидатом на роль белка, напрямую участвующего в поддержании структуры хроматина. Однако работы в этой области, во-первых, немногочисленны, а во-вторых, не содержат информации о парных взаимодействиях и потому не реализуют возможности исследования хроматина на разных уровнях организации в масштабах всего генома.

Целью данной работы является исследование влияния нуклеопорина Elys на структуру хроматина дрозофилы и ее взаимосвязь с различными факторами, ассоциированными с ранними стадиями сперматогенеза, на основе данных Hi-C.

Анализ проведенных исследований по данной тематике показывает, что при наличии карт хроматиновых взаимодействий для ранних стадий сперматогенеза дрозофилы, а также данных Hi-C при деплеции Elys актуальным является решение перечисленных ниже **задач**:

- 1) выявление взаимосвязи между активацией сперматоцитспецифичных генов и реорганизацией структуры хроматина;
- 2) анализ структуры хроматина в контексте характерных для сперматоцитов цитологических изменений;

- 3) анализ изменений в организации активного и неактивного хроматина при деплеции Elys.

Положения, выносимые на защиту:

- 1) Активация сперматоцитспецифичных генов сопровождается повышением частоты их контактов, однако такой вид взаимодействий не является избирательным: так же часто гены данной группы формируют контакты с генами, активными в различных тканях дрозофилы. Функциональное сходство этой группы генов ассоциировано с линейной, но не пространственной колокализацией.
- 2) Значительная доля сперматоцитспецифичных генов приходится на неактивный хроматин. При активации их транскрипции имеет место локальная декомпактизация плотности окружения промоторов, но не формирование новых, ярко выраженных границ.
- 3) На ранних стадиях сперматогенеза в хроматине дрозофилы обнаруживаются изменения, не связанные с экспрессией и, вероятно, обусловленные изменением объема ядра клетки. Это наблюдение позволяет рассматривать данный процесс в качестве естественного эксперимента по увеличению объема ядра и учитывать его в будущих исследованиях по данной тематике.
- 4) В целом, при переходе от стадии сперматогониев к стадии сперматоцитов происходит связанное с активацией сперматоцитспецифичных генов изменение конформации хроматина на нескольких уровнях организации. Масштаб этих изменений зависит от исходного положения участка генома в более или менее активной области хроматина.
- 5) Деплеция нуклеопорина Elys также приводит к заметным изменениям на разных уровнях пространственной организации и затрагивает как активную, так и неактивную фракцию хроматина. Нуклеопорин Elys, наряду с белками ламины, является важным компонентом системы регуляции плотности упаковки ДНК и ее распределения в пространстве ядра.

Научная новизна данной работы во многом обусловлена выбором объекта исследования, поскольку анализ структуры хроматина на каких-либо стадиях сперматогенеза в дрозофиле ранее не проводился. Например, значительный интерес представляет экспрессия сперматоцитспецифичных генов, т. к. информация о влиянии запуска транскрипции большого числа генов на уже сформировавшуюся структуру хроматина дрозофилы в хоть и специфичных, но все же относительно естественных условиях, не представлена в литературе по данной тематике. Впервые были описаны изменения в организации хроматина при многократном увеличении объема ядра клетки у данного вида. Оригинальность

представленных результатов анализа данных Hi-C при деплеции нуклеопорина Elys объясняется несколькими факторами, касающимися выбора как объекта, так и метода исследования, и не ограничивается областью видовой специфики. Это первая работа, в которой, благодаря удачному выбору модельного организма и конкретной клеточной линии, удалось получить данные по хроматину при деплеции Elys, а применение Hi-C позволило проанализировать его влияния на разных уровнях организации.

Практическая и теоретическая значимость проведенного исследования включает ряд методических наблюдений, полезных для дальнейших исследований организации хроматина в дрозофиле. Например, в ходе анализа был оптимизирован подход к детекции ТАДов. Выявленные закономерности позволяют заполнить некоторые пробелы в общей картине многообразия факторов, влияющих на структуру хроматина дрозофилы в целом и на ранних стадиях сперматогенеза в частности. Также, в ходе теоретической и практической работы, были обнаружены предпосылки для дальнейших исследований как в области методов анализа конформации хроматина, так и в направлении более глубокой биологической интерпретации полученных результатов. Кроме того, знания, которые были получены при анализе изменений, вызванных деплецией Elys, могут служить полезным референсным ресурсом при изучении состояний, характеризующихся нарушением метаболизма компонентов комплекса ядерной поры.

Достоверность полученных результатов обеспечивается такими аспектами исследования, как наличие биологических повторностей, оценка качества данных, проведение анализа как на группе тестируемых генов, так и контрольных, нормализация данных Hi-C, применение корректных статистических тестов с указанием статистической значимости, а также публикациями в индексируемых научных журналах.

Личный вклад. Автор принимал непосредственное участие в выборе стратегий исследований, планировании анализа данных и его выполнении, интерпретации и обобщении полученных результатов. Все результаты анализа данных, описанные в диссертационной работе, получены автором лично.

Объем и структура работы. Диссертация состоит из введения, трех глав, заключения, выводов и приложения. Полный объем диссертации составляет 109 страниц с 33 рисунками и 5 таблицами. Список литературы содержит 125 наименований.

Содержание работы

Глава 1 посвящена обзору литературы, содержащей актуальную информацию по теме диссертации. В параграфах главы изложены методы детекции структуры хроматина, подробно рассмотрены основные этапы протокола Hi-C и специфика обработки этого типа данных. Кроме того, освещаются общие закономерности трехмерной организации генома эукариот; значительная часть раздела приходится на описание текущих представлений о структуре хроматина у дрозофилы и ряда опосредующих ее факторов, в том числе белков периферии. Рассмотрены ключевые особенности протекания начальных стадий сперматогенеза у данного вида.

В **главе 2** содержится информация о материалах и методах, используемых в данном исследовании. В частности, описаны применяемые подходы для построения карт Hi-C и поиска основных паттернов структуры хроматина дрозофилы, перечислены параметры запуска инструментов, критерии классификации исследуемых генов или характеристик структуры хроматина на группы с различными свойствами, методы анализа структуры хроматина на разных масштабах взаимодействий и визуализации. Помимо прочего, в главе разобран алгоритм поиска значимых контактов для сперматоцитспецифичных генов, особенности применяемых контролей и анализ функционального обогащения.

В **главе 3** изложены основные результаты диссертационной работы.

Первый раздел главы 3 посвящен анализу изменений в структуре хроматина при активации транскрипции сперматоцитспецифичных генов на ранних стадиях сперматогенеза дрозофилы.

Исследование динамики трехмерной организации хроматина в процессе сперматогенеза проводилось на данных Hi-C, полученных для семенников *bamΔ86* мутантов, содержащих в основном сперматогонии (ранняя стадия), и семенников личинок третьей стадии, по большей части содержащих сперматоциты (более поздняя, предмейотическая стадия). Анализ был преимущественно сосредоточен на двух группах генов, сперматоцитспецифичных и повсеместно экспрессирующихся (убиквитных), также привлекались публичные данные Hi-C и экспрессии для клеточной линии Kc167. По результатам первичной оценки качества карт Hi-C на основе информации о сходстве реплик, форме кривой зависимости средней частоты контактов от расстояния и, наконец, визуального анализа, был сделан вывод о том, что качество карт контактов и выбор разрешения от 2 до 20 тыс. п.о. (в зависимости от цели) является подходящим для дальнейшего исследования.

Для поиска изменений, происходящих при переходе от стадии сперматогонии к стадии сперматоцитов, структура хроматина была проанализирована на нескольких уровнях организации. Анализ компарментной

структуры показал, что переход от стадии сперматогониев к стадии сперматоцитов сопровождался повышением значения первой главной компоненты (PC1) для участков генома, содержащих старты сперматоцитспецифичных генов, тогда как для убиквитных генов эта величина не менялась значимо между стадиями. Т. к. значение PC1 традиционно используется в качестве меры выраженности принадлежности к А- или В-компарменту, его повышение указывает на полное или частичное перемещение сперматоцитспецифичных генов в область активного компармента при запуске их транскрипции. Для 11% генов из этой группы наблюдался непосредственно переход из В- в А-компармент (т. е. имело место изменение знака PC1). В целом на долю событий смены компармента с неактивного в сперматогониях на активный в сперматоцитах приходится около 7.5% генома, из которых 25% содержит сперматоцитспецифичные гены. Для 42% сперматоцитспецифичных генов, несмотря на активацию и повышение значения PC1 (P -значение $< 10^{-60}$), перехода в А-компармент не происходило. В связи с этим наблюдением было выдвинуто предположение о наличии некоторых дополнительных факторов, ограничивающих или разрешающих такие события, одним из которых, предположительно, могла быть близость региона к границе ТАДа. Действительно, расстояние до ближайшей границы, как и размер ТАДа, при сравнении сперматоцитспецифичных генов, перешедших из В- в А-компармент и оставшихся в В, оказались меньше для первой группы (P -значение $< 10^{-14}$, рис. 1а). При этом оказалось, что уровень экспрессии сперматоцитспецифичных генов не связан с наличием или отсутствием перехода в активный компармент (P -значение ~ 0.3 , рис. 1б).

Данный раздел анализа был дополнен сравнением выраженности В-компармента в сперматогониях, сперматоцитах и клеточной линии Kc167. Мы показали, что для обеих исследуемых стадий сперматогенеза взаимодействия между регионами с отрицательным значением PC1, превышающие ожидаемые частоты контактов, распространяются на большую долю генома по сравнению с Kc167 (рис. 2а-г). При этом, вопреки ожиданиям, при сопоставлении транскрипции в регионах, вовлеченных в формирование выраженного В-компармента (рис. 2а,б, красная пунктирная линия), ее значения оказались в Kc167 значимо меньше по сравнению со стадией сперматогониев (P -значение = 0.023), а для стадии сперматоцитов эти участки генома и вовсе обогащены активными, специфичными для данной стадии, генами (P -значение $< 10^{-05}$). Таким образом, было показано, что экспрессия генов хотя и коррелирует с их расположением в А- или В-компарменте, все же не является единственным определяющим фактором ни для принадлежности региона к тому или иному компарменту, ни для выраженности последнего.

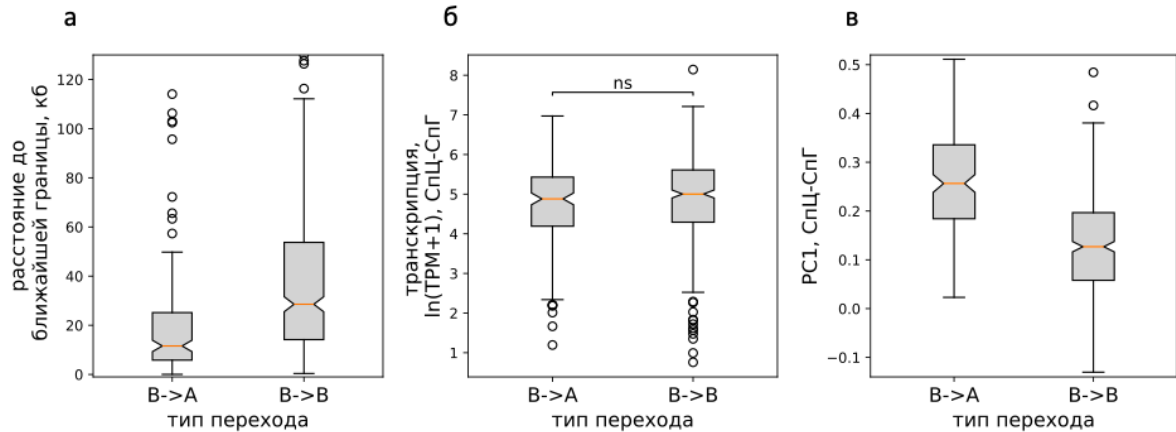


Рис.1 Сравнение характеристик регионов, перешедших в активный компартмент (B-A) и оставшихся в неактивном (B-B). (а) Расстояние до ближайшей границы ТАДа. (б) Изменение уровня транскрипции. (в) Изменение значения PC1.

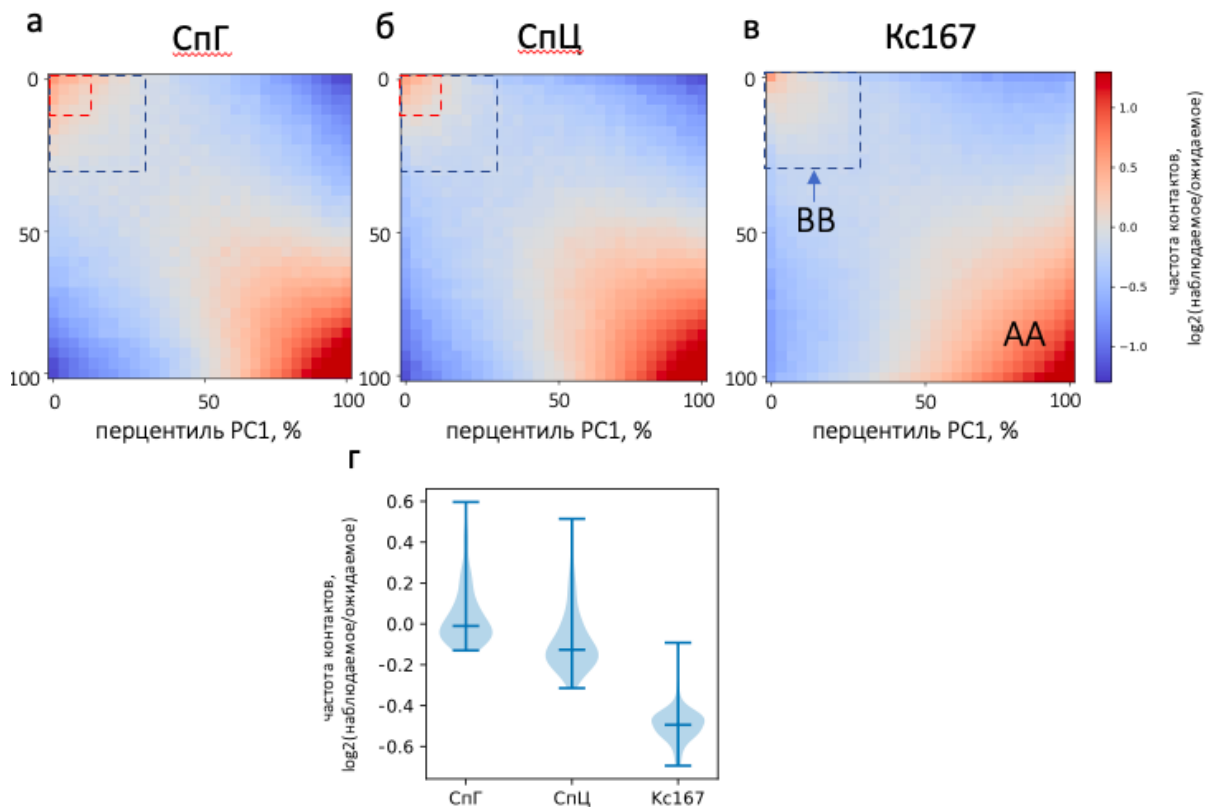


Рисунок 2. Структура компартментов. (а) Сперматогонии. (б) Сперматоциты. (в) Клеточная линия Kc167. (г) Частоты контактов в неактивном компартменте для области, очерченной синей пунктирной линией.

Т. к. изменение структуры компартментов в ходе сперматогенеза предполагает перестройку на уровне отдельных парных взаимодействий, для подробного исследования взаимосвязи между этой формой организации хроматина и активацией сперматоцитспецифичной транскрипции был проведен анализ частот контактов стартов генов из трех групп: сперматоцитспецифичных, неактивных и

убиквитных. Результаты анализа говорят о том, что активация сперматоцитспецифичных генов сопровождается повышением частот контактов их стартов транскрипции (P -значение $\sim 10^{-39}$) (рис. 3, 4). Данную зависимость удалось показать не только для генов, расположенных внутри одного и того же ТАДа, но и относящихся к разным ТАДам. При этом частота контактов неактивных генов не отличалась от фоновой и различий между стадиями обнаружено не было (P -значение ~ 0.8), тогда как для убиквитных генов наблюдалось повышение интенсивности взаимодействий на обеих стадиях по сравнению с прилегающими к ним участками генома. В соответствии с полученной ранее информацией о предпочтительном расположении убиквитных генов в границах ТАДов и с учетом описанной ранее в литературе тенденции границ ТАДов к пространственной кластеризации, для этой группы генов парные взаимодействия проявлялись в форме яркого контакта, напрямую связанного с формированием А-компартамента.

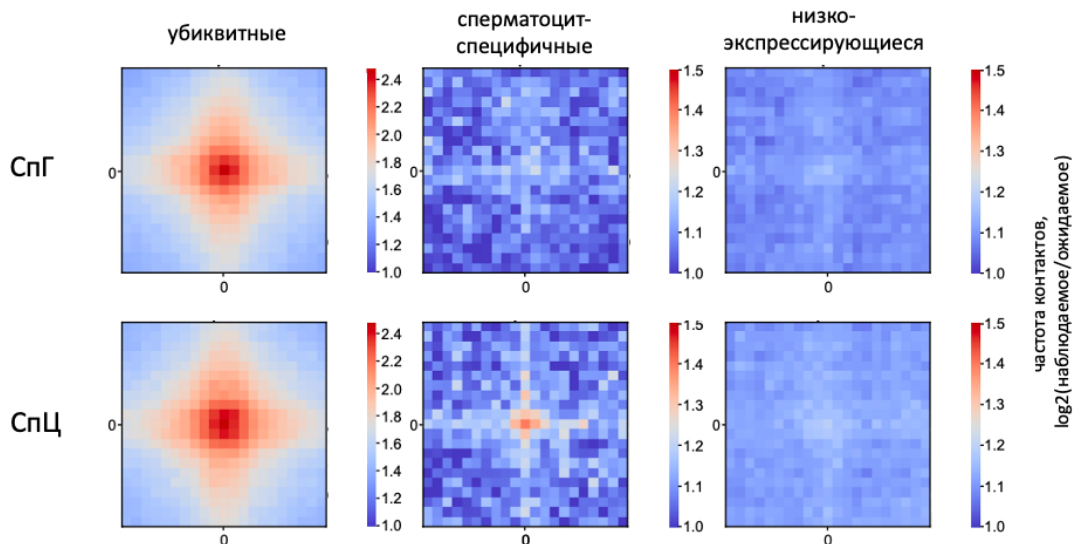


Рисунок 3. Парные взаимодействия генов из разных ТАДов.

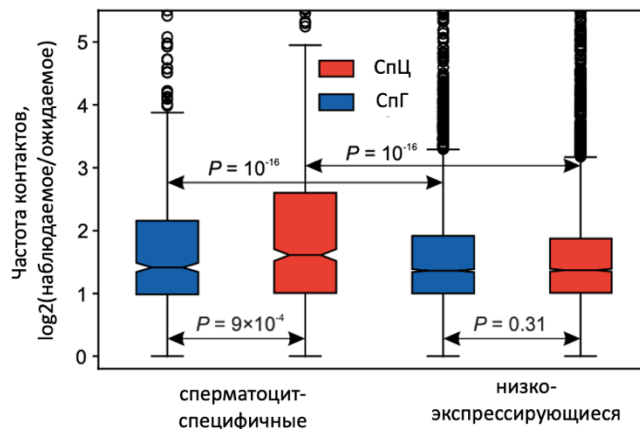


Рисунок 4. Парные взаимодействия генов внутри ТАДов.

Наличие этой закономерности для сперматоцитспецифичных генов было продемонстрировано и на уровне межхромосомных взаимодействий (P -значение $<$

10^{-20}). Однако в данном случае на стадии сперматогониев также было обнаружено повышение частот контактов по сравнению с окружающими фоновыми взаимодействиями, что может объясняться формированием структуры хроматина, которое предшествует и способствует дальнейшей активации транскрипции.

Обнаруженное повышение частот контактов сперматоцитспецифичных генов привело нас к вопросу о специфичности таких взаимодействий: проявляют ли эти гены избирательность в формировании контактов по принципу «пространственной близости с себе подобными»? Существует ли взаимосвязь между повышением частоты контакта и функциональным сходством лежащих в его основаниях сперматоцитспецифичных генов? Для ответа на эти вопросы, на основе анализа данных Hi-C стадии сперматоцитов, были выявлены контакты, содержащие хотя бы в одном из оснований сперматоцитспецифичный ген и характеризующиеся статистически значимым повышением интенсивности по сравнению со всеми остальными контактами, которые были образованы геномными локусами, разделенными тем же линейным расстоянием. В результате было получено около 13 тыс. пар взаимодействующих генов, из которых затем было выделено ~700 пар контактов двух сперматоцитспецифичных генов и ~4800 пар контактов между сперматоцитспецифичным и убиквитным геном. Сравнение с числом потенциально возможных пар генов двух вышеуказанных типов показало отсутствие предпочтения в формировании контактов сперматоцитспецифичных генов с себе подобными: так же часто они формируют контакты и с убиквитными генами (P-значение = 0.38). Более того, в ходе анализа семантического сходства взаимодействующих друг с другом сперматоцитспецифичных генов, при использовании пермутационного контроля, взаимосвязи между пространственной колокализацией и функциональным сходством обнаружено не было (пермутационное P-значение > 0.05 для всех трех онтологий: биологические процессы, молекулярные функции, клеточные компоненты). Интересно, что иная картина наблюдалась для сперматоцитспецифичных генов, организованных в кластеры на основе линейной близости в геноме. Семантическое сходство между парами генов, относящихся к одному кластеру, оказалось значимо выше, чем между генами из разных кластеров (пермутационное P-значение = 0.01 для каждой из онтологий). Примечательно, что только 18% кластеров сперматоцитспецифичных генов было разделено границей, тогда как в основном каждый кластер располагался в пределах одного ТАДа. Таким образом, было показано, что общая регуляция сперматоцитспецифичных генов, объединенных сходными функциями, по крайней мере отчасти может обеспечиваться за счет их последовательного расположения в геноме.

Поскольку одним из актуальных направлений в области исследования организации хроматина является выявление факторов, вовлеченных в

формирование границ ТАДов, и одним из них может быть транскрипция, мы также проанализировали структуру хроматина в данном контексте. Для выровненных по суммарному числу контактов карт было детектировано ~3800 ТАДов со средней длиной 35 тыс.п.о. на каждой из стадий, что соотносится в описанными в литературе параметрами для данных Hi-C высокого разрешения. При сравнении разметки ТАДов для обеих стадий было обнаружено, что около 90% границ остаются неизменными с точностью до одного бина размером 5 тыс.п.о. Для границ, специфичных для стадии сперматоцитов и приходящихся на ТАДы в сперматогониях, была выявлена значимая ассоциация с содержанием в них сперматоцитспецифичных генов при сравнении с остальными границами (P-значение = 0.005), убиквитные гены, напротив, значимо реже встречались в таких границах по сравнению со всеми прочими (P-значение < 10^{-16}). Также для границ этого типа характерно небольшое повышение транскрипции в сперматоцитах по сравнению со сперматогониями, однако оно не являлось статистически значимым (P-значение ~ 0.08). При этом границы, детектированные на стадии сперматоцитов, и приходящиеся на ТАДы в сперматогониях, содержали только 1.2% сперматоцитспецифичных генов и еще 9% располагались в общих для обеих стадий границах. При визуализации участков карт Hi-C в окрестности новых границ, содержащих сперматоцитспецифичные гены, оказалось, что часть из них вызывает сомнения относительно статуса границы как такового, а для других локальная структура визуально не отличается от наблюдаемой в сперматогониях или рядом с границей обнаруживаются участки карт с пропущенными значениями, т. е. такие события в целом могли быть артефактами работы алгоритма Armatus. В целом, большинство генов из этой группы приходится на внутреннюю область ТАДов: для нее не характерна перепредставленность в границах. Обнаруженные изменения были интерпретированы нами как преимущественно свидетельствующие об отсутствии ассоциации между активацией сперматоцитспецифичных генов и формированием новых границ. Вместе с тем, исследование локальной плотности хроматина, выраженной через индекс инсуляции, показало наличие заметных и значимых изменений в области промоторов при активации транскрипции. В результате данного анализа было обнаружено, что при активации сперматоцитспецифичных генов происходит локальное разрыхление хроматина в радиусе 5-6 тыс.п.о. от промотора гена (рис. 5а). Кроме того, мы показали, что разница в компактизации напрямую связана с изменением транскрипции: активация генов, приводящая к более выраженному увеличению транскрипции, соотносится с более сильным разрыхлением, и наоборот (рис. 5б, коэффициенты корреляции Спирмена для стадии сперматогониев и сперматоцитов равны 0,89 и 0,99, P-значение = 0,037 и P-значение < 10^{-24} , соответственно). Для убиквитных генов наблюдался ожидаемо более сильный провал в кривой инсуляции, поскольку

эти гены преимущественно располагаются в границах ТАДов (P -значение $< 10^{-4}$, пермутационный тест), где по определению хроматин характеризуется наименьшей компактизацией, при этом уменьшения локальной плотности на стадии сперматоцитов обнаружено не было.

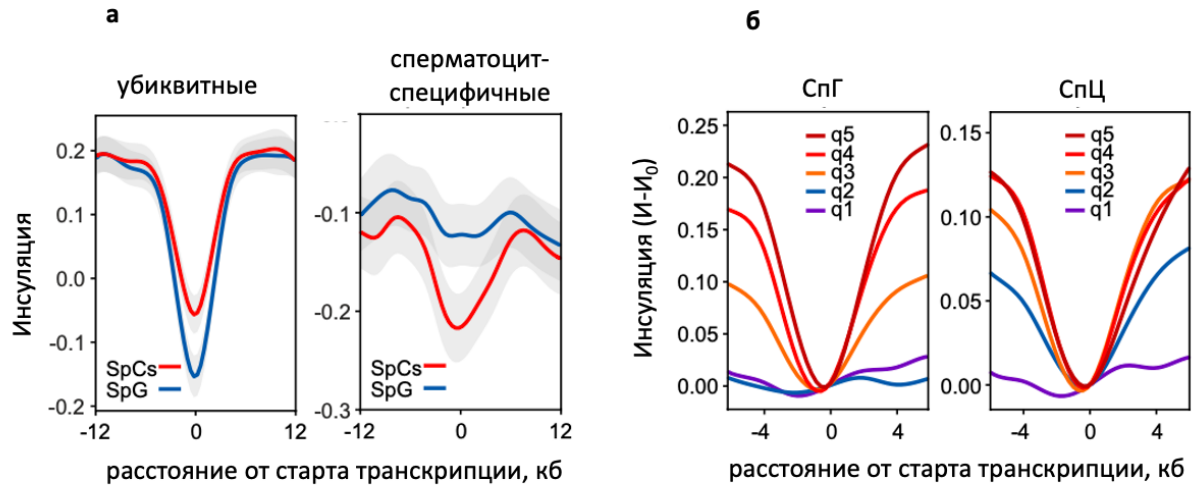


Рисунок 5. Кривые инсуляции вокруг стартов транскрипции. (а) Группы убиквитных и сперматоцитспецифичных генов. Погрешность значений - 95%-ный доверительный интервал. (б) Все гены, поделенные на квантили экспрессии.

В завершении данного раздела исследования мы также проанализировали взаимное расположение *aly*-зависимых и независимых сперматоцитспецифичных генов и границ ТАДов в предположении о возможной их колокализации, которое было обусловлено следующими соображениями: активация сперматоцитспецифичных генов происходит не одновременно и регулируется целым набором белков, среди которых особенно выделяется белок - продукт гена *aly* (один из основных компонентов комплекса tMAC, необходимого для второй волны активации сперматоцитспецифичной экспрессии); согласно имеющимся данным, *aly*-независимые гены экспрессируются раньше по сравнению с остальными сперматоцитспецифичными генами. При сравнении распределения *aly*-зависимых и независимых сперматоцитспецифичных генов относительно границ ТАДов была выявлена слабая, но значимая ассоциация между локализацией в границах и принадлежностью к группе *aly*-независимых генов (P -значение = 0.041 для общих границ, P -значение ~ 0.0068 для всех границ в сперматоцитах). Значение PC1 в сперматогониях для участков генома, содержащих старты *aly*-независимых генов, также оказалась несколько выше, чем для *aly*-зависимых (P -значение = 0.042). В качестве одного из возможных объяснений для обнаруженных закономерностей мы рассматривали предрасположенность генов, активирующихся раньше в ходе развития, находиться в заведомо более транскрипционно выгодном с точки зрения структуры хроматина состоянии.

Второй раздел главы 3 посвящен анализу структуры хроматина в контексте изменения размера клеток на ранних стадиях сперматогенеза. На стадии сперматоцитов имеет место ~ 25 -кратное увеличение размера клетки, что приводит также к увеличению размера ядра и может сказываться на структуре хроматина, при этом отсутствие гомологичной рекомбинации у самцов дрозофилы позволяет не учитывать данный фактор (у млекопитающих этот процесс сопровождается конденсацией хроматина, проявляющейся отчасти уже на ранних стадиях сперматогенеза) и дает возможность более однозначно интерпретировать полученные результаты.

В данном разделе исследования мы проанализировали выраженность компарментов, петель и границ ТАДов, а также оценили изменение в средней частоте контактов в зависимости от расстояния (кривая шкалирования, $P(s)$). Было обнаружено, что на расстоянии от ~ 100 до ~ 300 тыс.п.о. наблюдается изменение скорости убывания средней частоты контактов (P -значение=0.014) и далее практически до конца кривой интенсивность взаимодействий остается в сперматоцитах ниже по сравнению со сперматогониями (рис. 6а). Компарменты обоих типов становятся менее плотными и обособленными (рис. 6б). При этом межхромосомные и особенно межплечевые контакты становятся более выраженными (табл. 1). Также проведенный анализ показал, что на фоне отсутствия изменений в плотности ТАДов в сперматоцитах наблюдается значимое ослабление силы границ (P -значение ~ 0.0022), а также петель (P -значение $< 10^{-57}$, рис. 7). Интересно, что в сперматоцитах в целом детектируется примерно в три раза меньше петель по сравнению со сперматогониями, причем различие в их числе и интенсивности не связано с уровнем экспрессии генов, лежащих в основаниях, поскольку этот фактор был учтен дополнительно.

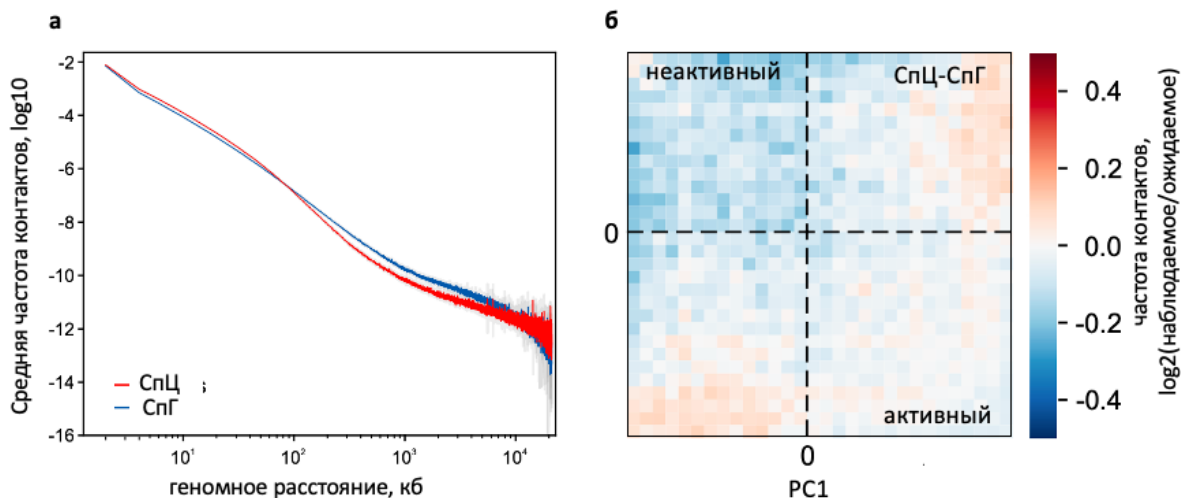


Рисунок 6. Крупномасштабные изменения в структуре хроматина сперматогониев и сперматоцитов. (а) Средняя частота контактов в зависимости от геномного расстояния. (б) Различия на уровне компарментов.

Таблица 1. Средняя частота межхромосомных взаимодействий, нормированная на общее число контактов. Выше диагонали – сперматоциты, ниже диагонали – сперматогонии.

	2L	2R	3L	3R	X
2L		1.40	0.44	0.46	0.45
2R	0.67		0.48	0.50	0.49
3L	0.35	0.38		1.25	0.46
3R	0.35	0.38	0.58		0.44
X	0.35	0.38	0.33	0.33	

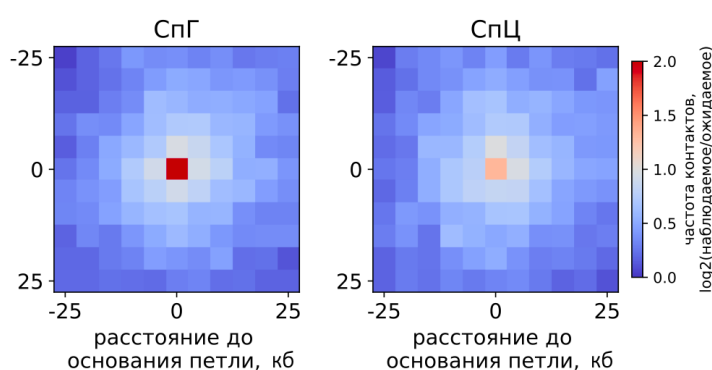


Рисунок 7. Средняя интенсивность контактов для объединения детектированных на каждой из стадий петель, в основаниях которых средняя экспрессия варьируется в пределах 10% между сперматогониями и сперматоцитами.

Изменение в структуре компартментов и кривой скейлинга объяснялось нами как более релаксированное и перемешанное состояние хроматина, связанное с расширением хромосомных территорий на фоне увеличения объема ядра, при сохранении плотности ТАДов и средней частоты контактов на коротких расстояниях за счет взаимодействия с ламиной. Полученные результаты во многом согласуются с результатами работы по искусственному расширению ядра в клетках млекопитающих, где было показано уменьшение выраженности компартментов, а также повышение межхромосомных взаимодействий у ассоциированных с ламиной хромосом [1]. При этом повышение дальних внутрихромосомных взаимодействий, наблюдаемое при искусственном увеличении объема ядра на расстоянии более 1 Мб, проявляется в сперматоцитах только с отметки 10 Мб. Кроме того, продемонстрированное в указанной работе ослабление петель было характерно только для CTCF-ассоциированных. Как известно, у дрозофилы CTCF не является основным архитектурным белком, ассоциированным с петлями, однако в сперматоцитах их интенсивность заметно падает. Отсутствие на сегодняшний день данных по интенсивности связывания архитектурных белков на ранних

стадиях сперматогенеза препятствует более подробному анализу в этом направлении. Кроме того, для более уверенного сравнительного анализа между естественным экспериментом по увеличению ядра в сперматоцитах и его искусственным раздуванием, необходимо учитывать множество параметров, например, продолжительность периода увеличения ядра, его размеры и размеры генома.

В третьем разделе главы 3 подробно рассмотрено влияние одного из белков ядерной поры, нуклеопорина Elys, на плотность упаковки хроматина в клетках S2 дрозофилы. Эта часть анализа проводилась на основе сравнения данных Hi-C для образцов после деплеции этого белка методом РНК-интерференции и контрольных образцов (клеточная линия с добавлением двуцепочечной РНК, комплементарной последовательности гена бактериального белка LacZ). Также использовалась полученная сотрудниками лаборатории Ю.Я.Шевелева информация о сайтах связывания Elys в нуклеоплазме и в области ядерной поры ($Elys_{\text{нукл}}$ и $Elys_{\text{янк}}$ соответственно), данные транскрипции в контроле и публичные данные по активности и взаимодействию с ламиной. На предварительном этапе была проведена оценка качества полученных карт контактов. Кластеризация реплик показала хорошее разделение образцов на две группы, что, в совокупности с визуальной оценкой карт Hi-C, позволило сделать вывод о достаточном для последующего анализа качестве исходных данных и выбрать разрешение от 2 до 20 тыс. п.о.

Исследование влияния Elys на структуру хроматина проводилось как для крупномасштабных форм его организации, так и на уровне локальной упаковки в окрестности сайтов связывания Elys. Для изучения взаимодействий внутри А- и В-компартамента были построены тепловые карты («седловые графики»), отражающие интенсивность контактов внутри них и между ними. При деплеции Elys наблюдалось усиление неактивного и ослабление активного компартамента (рис. 8а, ВВ и АА область соответственно). При этом частоты контактов между участками генома, принадлежащими к разным компартаментам, в целом не изменялись, т. е. в отсутствие белка Elys перемешивания между компартаментами не происходило (рис. 8б). Анализ их протяженности не выявил преобладающих переходов из компартамента типа А в тип В или наоборот (Р-значение = 0.36).

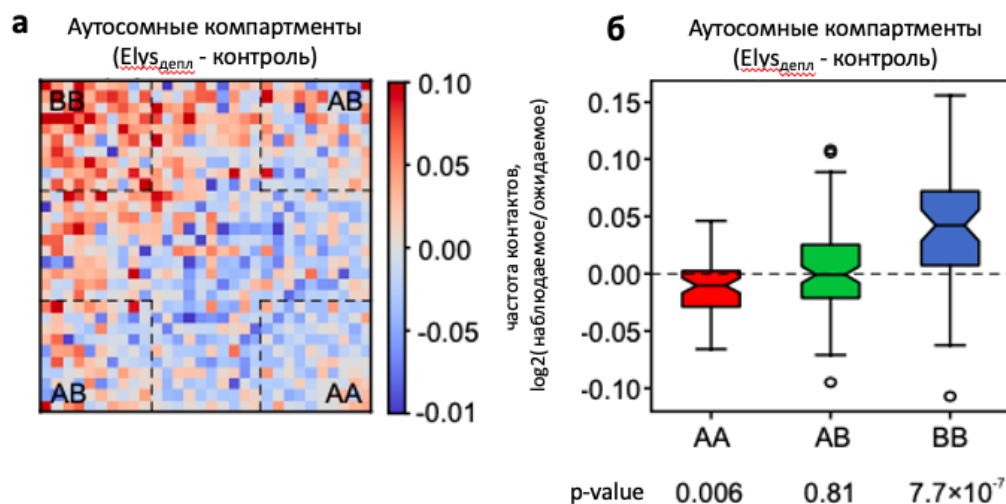


Рисунок 8. Взаимодействия хроматина на уровне компартментов. (а) Агрегированные частоты контактов внутри и между компартментами. (б) Статистическая оценка значимости различий в частотах контактов (частоты контактов в областях, выделенных пунктирной линией на панели а).

Для получения более подробной картины компартментализации до и после деплеции Elys эта форма организации хроматина была проанализирована в трех интервалах расстояний между участками генома, контактирующими за пределами ТАДов: ближние (0.2-4 Мб), средние (4-10 Мб) и дальние (10-20 Мб) (рис. 9). Взаимодействия внутри неактивного компартмента оказались после деплеции Elys более интенсивными для всех указанных диапазонов расстояний, тогда как для активного компартмента значимых отличий обнаружено не было, хотя в группе ближних контактов наблюдалось небольшое снижение их частот (рис. 9б). По-видимому, наблюдаемый на общей тепловой карте (рис.8) эффект ослабления А-компартмента на самом деле является слабо выраженным, основной вклад в эти изменения могут вносить ближние контакты, а его статистическая значимость существенно зависит от баланса между числом измерений и амплитудой наблюдаемых изменений.

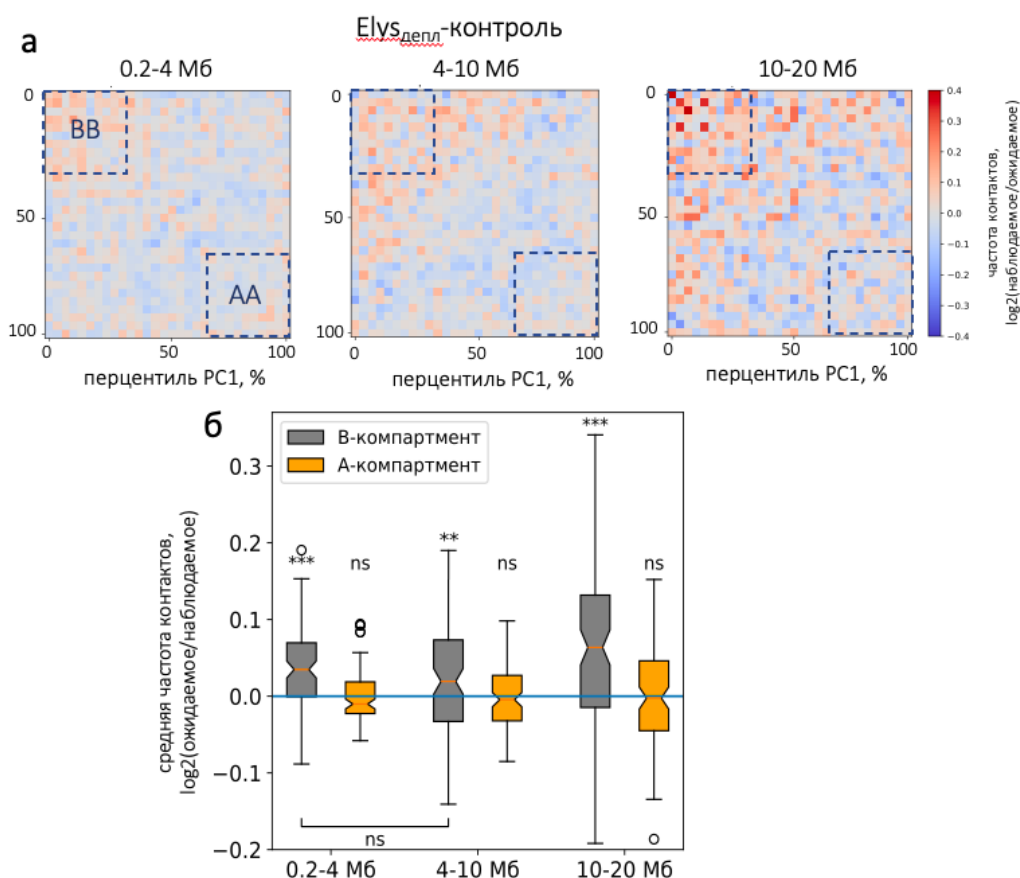


Рисунок 9. Изменение структуры компарментов, образованных взаимодействием участков генома, расположенных на разных расстояниях за пределами ТАДов. (а) «Седловые графики». (б) Разница частот контактов в компартаментах между картами для деплеции Elus и контроля (очерчены пунктирной линией на панели а).

Мы также изучили влияние деплеции Elus на плотность ТАДов с различными свойствами. Все ТАДы были разделены на три группы: активные – с высоким содержанием активных цветов хроматина и низким содержанием ЛАДов, неактивные – с низким содержанием активных цветов хроматина и высоким содержанием ЛАДов и промежуточные. Для всех групп ТАДов с границами, совпадающими в контрольных картах и после деплеции, была рассчитана их плотность, выраженная через среднюю частоту контактов. Для активных ТАДов деплеция Elus привела к повышению плотности, тогда как для неактивных наблюдалась противоположная тенденция (рис. 10а). Аналогичный результат был получен при делении ТАДов на группы по квартилям экспрессии (рис. 10б).

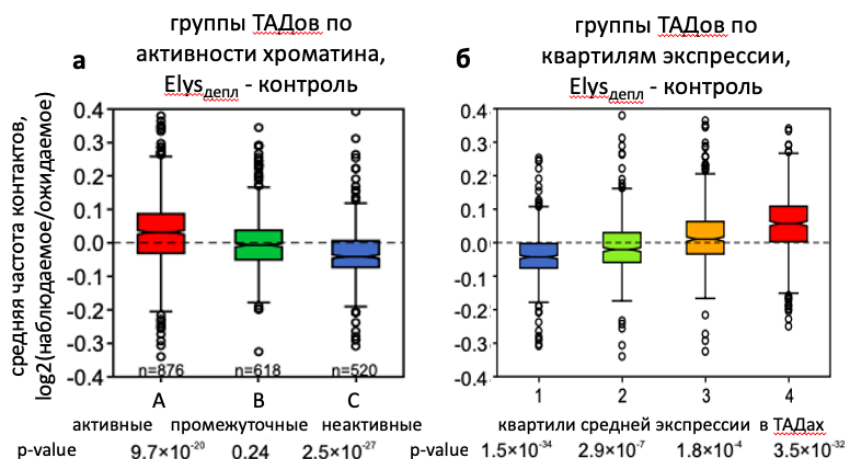


Рисунок 10. Плотность ТАДов. (а) ТАДы, разделенные на группы по активности, Elys_{депл} - контроль. (б) ТАДы, разделенные на квантили экспрессии, Elys_{депл} - контроль.

В связи с предполагаемым наличием у Elys_{нукл} ремоделирующего потенциала, опосредованного взаимодействием с комплексом РВАР, наряду с повышением плотности малых активных ТАДов при его деплеции ожидалось также повышение компактизации в области границ, которое могло проявляться в ослаблении сегрегации соседних ТАДов и их частичном перемешивании. Действительно, было показано, что при нокдауне Elys частота взаимодействий через границы заметно повышается, особенно ярко это проявляется в сильных границах (рис. 11а), для которых характерно выраженное обогащение Elys_{нукл} по сравнению с его содержанием в прилегающих к границе ТАДах (рис. 11б). Наблюдаемые изменения послужили подтверждением способности Elys к ремоделированию хроматина (возможно, через взаимодействие с другими белками) и в целом поддержали концепцию о формировании ТАДов и их границ у дрозофилы в результате чередования участков активного и неактивного хроматина: деплеция Elys может приводить к компактизации активного хроматина, формирующего границы между неактивными ТАДами, и приводить к увеличению частоты контактов между ними. С другой стороны, мы обнаружили, что 86% ТАДов остается в прежних границах (± 5 тыс.п.о.). Образование новых границ происходит при деплеции не чаще, чем исчезновение (Р-значение ~ 0.74), и для профиля инсуляции вокруг дифференциальных границ значимых отличий выявлено не было. Эти наблюдения соотносятся с полученными ранее представлениями о сложной природе границ ТАДов дрозофилы, в формировании которых участвует множество факторов.

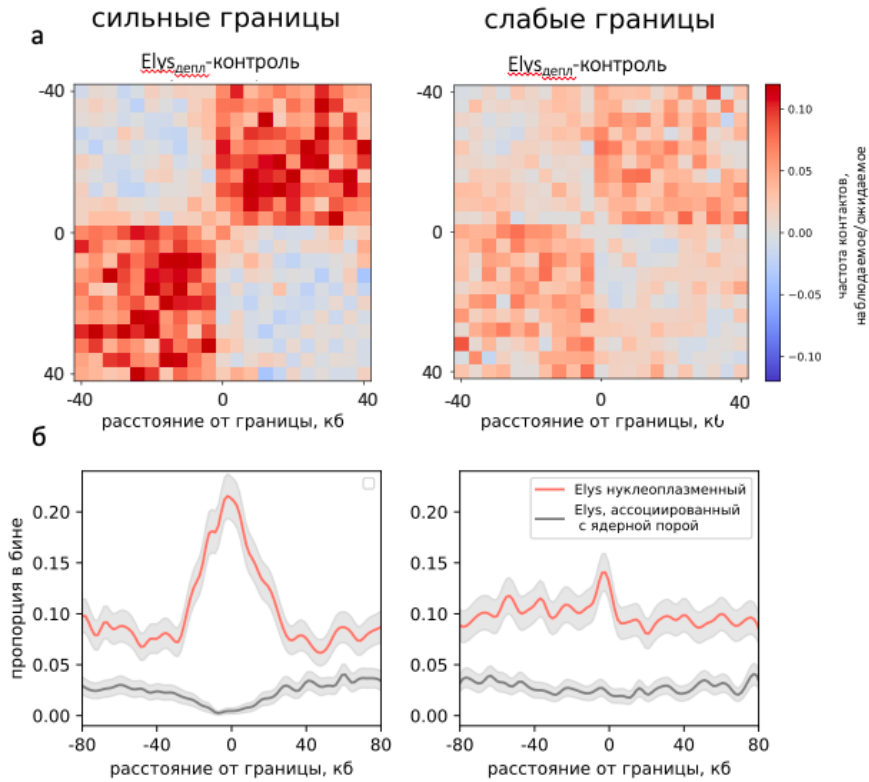


Рисунок 11. Влияние деплеции Elys на структуру хроматина в границах. (а) Средняя карта Hi-C вокруг сильных (слева) и слабых (справа) границ. (б) Профили пропорции Elys_{нукл} и Elys_{япк} вокруг сильных (слева) и слабых (справа) границ.

Elys обладает способностью напрямую связываться с ДНК в определенных участках. Благодаря наличию информации об их геномных координатах, мы имели возможность изучить влияние деплеции этого белка на конформацию хроматина непосредственно в области сайтов связывания. Так как Elys представлен в ядре двумя фракциями, нуклеоплазменной и ассоциированной с комплексом ядерной поры, исследование в данном направлении проводилось с учетом пространственной локализации. Для оценки изменения локальной плотности хроматина в окрестности сайтов связывания были построены средние тепловые карты контактов, которые содержат информацию о взаимодействиях внутри самого этого сайта и в пределах его ближайшего окружения. При деплеции Elys для ЯПК-ассоциированной фракции было показано снижение частоты контактов в местах его посадки по сравнению с контролем, свидетельствующее о локальной декомпактизации хроматина (рис. 12а). Для Elys_{нукл} картина оказалась не столь однозначна. С одной стороны, характер взаимодействий здесь сильно отличался от наблюдаемого в области связывания с ядерной порой, а с другой, обнаруженное повышение частот взаимодействий не ограничивалось сайтом связывания Elys_{нукл}, а распространялось и на близлежащие участки генома (рис. 12б). Влияние Elys на коммуникацию участков генома, расположенных по разные стороны от его сайта связывания, было проанализировано в терминах кривых инсуляции,

представляющих собой профиль локальных взаимодействий. Деpletion приводила к повышению частот контактов такого типа для сайтов связывания $Elys_{\text{нукл}}$, что, на наш взгляд, указывало на компактизацию активного хроматина в результате деплеции этого белка (рис. 13). Для $Elys_{\text{япк}}$ масштаб наблюдаемого эффекта не позволял говорить о его существенном влиянии на частоты взаимодействий участков генома, расположенных по разные стороны от сайта связывания, хотя направление изменений соответствовало общей картине, наблюдаемой при деплеции этого белка.

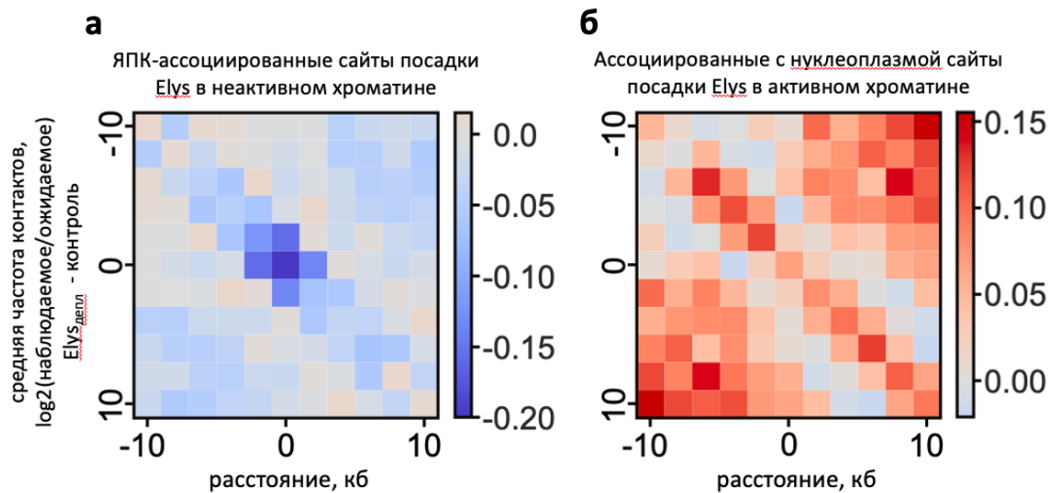


Рисунок 12. Средняя частота контактов вокруг сайтов связывания $Elys$. (а) В составе ядерной поры. (б) В нуклеоплазме.

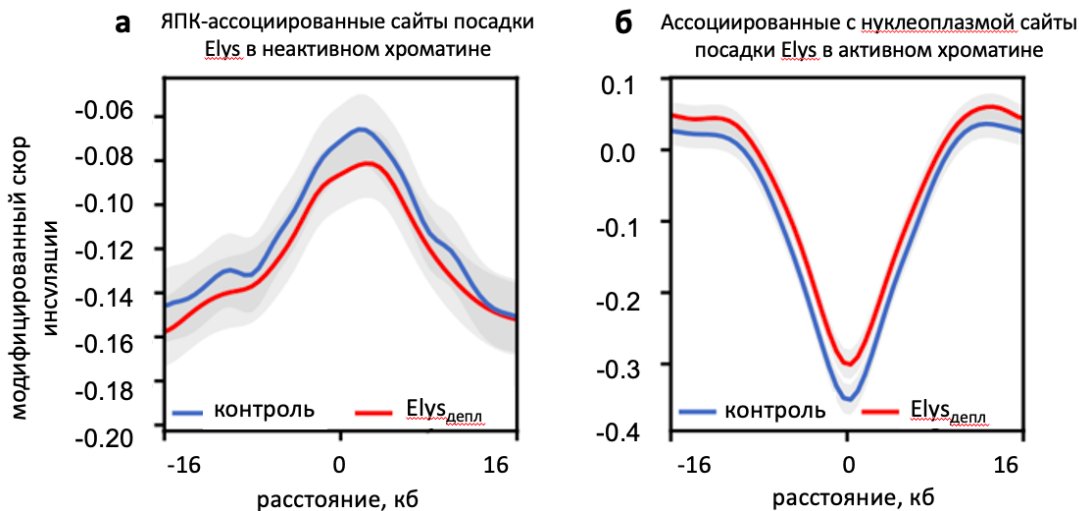


Рисунок 13. Модифицированные кривые инсуляции вокруг сайтов связывания. (а) $Elys$ в составе ядерной поры. (б) $Elys$ в нуклеоплазме.

В рамках данного раздела исследования был также проведен анализ петель и энхансерно-промоторных взаимодействий. Как при деплеции $Elys$, так и в контрольных образцах наблюдалось обогащение в интенсивности взаимодействия оснований петель, т. е. используемый алгоритм в целом позволял их детектировать,

однако изменения частоты контактов при сравнении контрольных данных и после деплеции Elys обнаружено не было (P-значение ~ 0.051). Участие Elys_{нукл} в процессе активации транскрипции на уровне регуляции структуры хроматина и пересечение 42% энхансеров клеток S2 с сайтами связывания Elys_{нукл} послужили основанием к проведению анализа взаимодействий этих энхансеров с соответствующими им промоторами. Информация о парах энхансер-промотор была получена из опубликованных ранее данных (метод STARR-seq), в анализ были взяты только такие случаи, когда расстояние между энхансером и промотором составляло не менее 10 кб. Такой подход также не выявил заметных изменений в частоте контактов соответствующих участков генома при деплеции Elys по сравнению с контролем.

Учитывая продемонстрированную при деплеции Elys тенденцию к перемещению хроматина от периферии ближе к центру ядра (часть исследования, выполненная группой Ю.Я.Шевелева, метод FISH), изменения, наблюдаемые в плотности ТАДов, были интерпретированы нами как ослабление связи неактивного хроматина с ядерной порой (и как следствие с ядерной оболочкой), приводящее к его деконденсации. Интересно, что сходный характер изменения компактизации ТАДов наблюдался ранее при деплеции ламины [2]. Это наблюдение говорит о том, что в распределение активного и неактивного хроматина в пространстве ядра и в поддержание уровня его компактизации вовлечен не один определенный фактор, а, по-видимому, целый комплекс разнообразных белков периферии.

Заключение

В настоящей работе представлены результаты исследования роли некоторых факторов в организации пространственной архитектуры генома у дрозофилы, в частности белков ядерной периферии и активации транскрипции на ранних стадиях сперматогенеза. Проведенный анализ литературы выявил неравномерность в степени изученности вклада того или иного фактора: значительная доля работ приходится на исследование влияния архитектурных белков, модификации гистонов, а также результатов индукции или подавления транскрипции, тогда как другие, не менее важные, компоненты структуры хроматина или протекающие в ядре процессы изучены в данном контексте в меньшей степени.

Наличие данных Hi-C для двух ранних стадий сперматогенеза, в силу многогранности этого процесса, позволило изучить несколько факторов, потенциально связанных со структурой хроматина. Исследование особенностей упаковки ДНК с точки зрения активации сперматоцитспецифичных генов выявило заметные отличия между сперматогониями и сперматоцитами, обусловленные этим процессом. Обнаруженные изменения затрагивают разные уровни

организации хроматина: от локального разрыхления до перехода между компартментами. С другой стороны, в работе продемонстрированы примеры отсутствия прямой взаимосвязи между трехмерной структурой хроматина, транскрипцией и функциональным сходством сперматоцитспецифичных генов.

Поскольку переход к стадии сперматоцитов сопровождается не только активацией транскрипции множества генов, но и увеличением объема ядра, было выдвинуто предположение о том, что ряд изменений в структуре хроматина может объясняться именно этим процессом. Сопоставление с результатами эксперимента по искусственному раздуванию ядра показало, что, действительно, некоторые отличия, наблюдаемые между двумя стадиями, согласуются с картиной, характерной для индуцированного увеличения объема ядра, и не связаны с активацией транскрипции.

Понимание роли различных факторов, определяющих структуру хроматина в дрозофиле, было дополнено результатами анализа данных Hi-C при деплеции белка Elys. Выявленные изменения указывают на существенный вклад этого белка в поддержание плотности упаковки и локализации неактивного хроматина на периферии ядра, а также в декомпактизацию активных участков генома. Учитывая согласованность обнаруженных изменений с полученной ранее картиной при деплеции белков ламины, интерпретация наблюдаемой динамики хроматина выходит за пределы описания функции отдельно взятого белка и на основе полученных результатов позволяет выдвинуть предположение о том, что регуляция распределения активного и неактивного хроматина может осуществляться за счет целого комплекса белков периферии.

В целом, результаты данной работы обогащают совокупную картину факторов, влияющих на структуру хроматина плодовой мушки, подчеркивают важность изучения других составляющих, потенциально ассоциированных с изменением конформации хроматина, и открывают новые перспективы исследований как в биологии сперматогенеза дрозофилы, так и в области технических аспектов анализа данных Hi-C.

На основании сделанных в работе наблюдений, можно сформулировать следующие **выводы**:

1. При переходе от стадии сперматогониев к стадии сперматоцитов происходят связанные с активацией сперматоцитспецифичных генов изменения конформации хроматина дрозофилы, такие как:
 - а) частичный переход в активный компартмент участков генома, содержащих промоторы сперматоцитспецифичных генов;
 - б) повышение частот контактов сперматоцитспецифичных генов;

- в) локальное снижение плотности хроматина вокруг стартов сперматоцитспецифичных генов.
2. С помощью усовершенствованного и реализованного алгоритма поиска значимых контактов для заданного списка геномных локусов было установлено, что повышение интенсивности взаимодействий сперматоцитспецифичных генов не является специфичным: частота таких контактов не отличается от наблюдаемой в парах сперматоцитспецифичных и активных в различных тканях генов.
 3. Функциональное сходство сперматоцитспецифичных генов ассоциировано с их линейной, но не пространственной колокализацией.
 4. Переход из неактивного компартмента в активный характерен для промоторов сперматоцитспецифичных генов, расположенных близко к границе ТАДов, т. е. масштаб изменений в структуре хроматина может зависеть от исходного расположения геномного локуса в более или менее активной области.
 5. Неактивный компартмент сперматогониев и сперматоцитов является более выраженным при сравнении с эмбриональной клеточной линией. Наблюдаемые различия не связаны напрямую с уровнем транскрипции, что подчеркивает значимость других факторов, предположительно вовлеченных в формирование неактивного компартмента.
 6. Обнаруженные изменения в структуре хроматина на уровне выраженности дальних взаимодействий, компартментов, границ и петель, вероятно, объясняются увеличением размера ядра сперматоцитов, поскольку соотносятся с закономерностями, продемонстрированными ранее в экспериментах по индуцированному увеличению объема ядра.
 7. Деpletion нуклеопорина Elys приводит к изменению плотности упаковки хроматина, сопоставимому с наблюдаемым при деpletion белков ламины, а именно:
 - а) в области связывания Elys в неактивном хроматине происходит заметная декомпактизация;
 - б) плотность неактивных ТАДов снижается, а активных повышается;
 - в) неактивный компартмент становится более выраженным на всех масштабах внутрихромосомных взаимодействий.

Процитированная литература

- 1 Sanders JT, Gollosi R, Das P *et al.* Loops, topologically associating domains, compartments, and territories are elastic and robust to dramatic nuclear volume swelling. *Sci Rep* 12(1), 4721 (2022).
- 2 Ulianov SV, Doronin SA, Khrameeva EE *et al.* Nuclear lamina integrity is required for proper spatial organization of chromatin in *Drosophila*. *Nat Commun* 10(1), 1176 (2019).

Список публикаций по теме диссертации:

1. Doronin, S. A.*, Ilyin, A. A.*, Kononkova, A. D.*, Solovyev, M. A.*, Olenkina, O. M., Nenasheva, V. V., Mikhaleva, E. A., Lavrov, S. A., Ivannikova, A. Y., Simonov, R. A., Fedotova, A. A., Khrameeva, E. E., Ulianov, S. V., Razin, S. V., & Shevelyov, Y. Y.* (2024). Nucleoporin Elys attaches peripheral chromatin to the nuclear pores in interphase nuclei. *Communications biology*, 7(1), 783.
2. Ilyin, A. A.*, Kononkova, A. D.*, Golova, A. V.*, Shloma, V. V.*, Olenkina, O. M.*, Nenasheva, V. V., Abramov, Y. A., Kotov, A. A., Maksimov, D. A., Laktionov, P. P., Pindyurin, A. V., Galitsyna, A. A., Ulianov, S. V., Khrameeva, E. E., Gelfand, M. S., Belyakin, S. N., Razin, S. V., & Shevelyov, Y. Y. (2022). Comparison of genome architecture at two stages of male germline cell differentiation in *Drosophila*. *Nucleic acids research*, 50(6), 3203–3225.

* - равный вклад авторов

Результаты работы, опубликованные в сборниках тезисов международных конференций:

1. ContactHunter: поиск значимых взаимодействий для заданных геномных локусов на основе данных Hi-C. Proceedings of 11th Moscow Conference on Computational Molecular Biology MCCMB'23.
2. Влияние изменения уровня ацетилирования на структуру хроматина в клеточной линии дрозофилы. Роль нуклеопорина Элис в поддержании структуры хроматина. Сборник трудов 47-й междисциплинарной школы-конференции ИППИ РАН. 2023
3. Chromatin structure changes during drosophila spermatogenesis. Сборник трудов 43-й междисциплинарной школы-конференции ИППИ РАН. 2019