

На правах рукописи

Фишман Вениамин Семенович

**СРАВНЕНИЕ ПРОСТРАНСТВЕННОЙ
ОРГАНИЗАЦИИ ГЕНОМОВ ФИБРОБЛАСТОВ И
СПЕРМАТОЗОИДОВ МЫШИ МЕТОДОМ Hi-C**

03.02.07 – генетика

Автореферат

Диссертации на соискание ученой степени

кандидата биологических наук

Новосибирск 2015

Работа выполнена в Федеральном государственном бюджетном научном учреждении «Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук» в лаборатории генетики развития, г. Новосибирск.

Научный руководитель	Доктор биологических наук, профессор, заведующий лабораторией генетики развития Серов Олег Леонидович ИЦиГ СО РАН, г. Новосибирск
Официальные оппоненты	Белякин Степан Николаевич , кандидат биологических наук, заведующий лабораторией геномики, Институт молекулярной и клеточной биологии СО РАН, г. Новосибирск Лебедев Игорь Николаевич , доктор биологических наук, заведующий лабораторией цитогенетики, Научно-исследовательский Института медицинской генетики СО РАМН, г. Томск
Ведущее учреждение	Федеральное государственное бюджетное учреждение науки Институт биологии гена, г. Москва

Защита диссертации состоится «___» _____ 2015 года на утреннем заседании диссертационного совета по защите диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук Д 003.011.01, в ИЦиГ СО РАН, в конференц-зале Института по адресу: проспект Академика Лаврентьева, 10, г. Новосибирск, 630090 Тел. (383) 3634906, факс (383)3331278, email: dissov@bionet.nsc.ru
С диссертацией можно ознакомиться в библиотеке ИЦиГ СО РАН и на сайте Института: www.bionet.nsc.ru

Автореферат разослан «___» _____ 2015 г.

Ученый секретарь
Диссертационного совета,
Доктор биологических наук

Т.М. Хлебодарова

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность проблемы. В ходе эволюции, особенно при переходе от бактерий к многоклеточным эукариотам, произошло огромное – на 2-3 порядка – увеличение размера генома (от миллионов до миллиардов пар нуклеотидов). Увеличение числа нуклеотидов в геноме привело к увеличению физической длины молекул ДНК. Для упаковки и хранения отдельных молекул ДНК в эукариотических клетках появились специальные функциональные структуры – хромосомы, кардинально отличающиеся от бактериальных хромосом. Помимо решения основной функции – уменьшения длины молекулы ДНК за счет формирования многочисленных петель, белки, участвующие в формировании петель, а также сами петли, приобрели важные регуляторные функции.

Практически все клетки многоклеточного организма несут молекулы ДНК с одинаковой первичной последовательностью, однако каждый тип клеток обладает уникальным профилем экспрессии генов. Различия в генной экспрессии обеспечиваются эпигенетическими механизмами – в частности, особенностями пространственной организации ДНК. Таким образом, пространственная укладка ДНК играет роль в таких фундаментальных процессах, как клеточная специализация, дифференцировка и реализация программы развития.

Детали механизмов, которые связывают пространственную структуру ДНК и работу генома, изучены не до конца. В частности, малоизученным являются субхромосомные уровни организации ДНК – фрагментов длиной от десятков до сотен тысяч пар оснований (п.о.). В последние десятилетия, в связи с развитием новых молекулярно-биологических методик и технологий массового параллельного секвенирования, эта брешь в наших знаниях постепенно заполняется. Показано, что в геномах многих типов клеток млекопитающих можно выделить пространственные домены: крупные А/В-компарменты (с размером несколько миллионов п.о.) и более мелкие ТAD-домены (topologically associated domains, топологические домены, с размером в несколько сотен тысяч п.о.) (Dixon et al. 2012; Rao et al. 2014).

Сперматозоиды являются абсолютно особенными клетками с точки зрения пространственной организации генетического материала в ядре. Во-первых, ДНК сперматозоидов, в отличие от

других клеток млекопитающих, упакована протаминами, а не гистонами (Balhorn et al. 1999). Во-вторых, размер ядра сперматозоида на порядок меньше, чем ядер соматических клеток (Lee et al. 1997). По этому параметру, уровень конденсации ДНК в сперматозоидах сходен с конденсацией ДНК в митотической хромосоме. В-третьих, в зрелых сперматозоидах отсутствует процесс транскрипции (Mudrak et al. 2011). Существует прямая и обратная связь между транскрипцией и пространственной организацией генома. Поэтому, пространственная организация транскрипционно-неактивных клеток представляет особый интерес. Наконец, в-четвертых, основной функцией сперматозоидов является передача генетической информации от родителей к потомкам. Если механизм передачи о первичной последовательности генома сперматозоидами известен, то как происходит передача информации о трехмерной структуре, неизвестно.

Целью данной работы является сравнение пространственной организации геномов сперматозоидов и фибробластов мышцы. Для выполнения данной цели были поставлены следующие задачи.

1. Построить карту пространственных контактов ДНК в геномах сперматозоидов и фибробластов мышцы.
2. Провести анализ и сравнение пространственных доменов геномов фибробластов и сперматозоидов.
3. Сравнить пространственную укладку и частоту контактов индивидуальных локусов в этих типах клеток.
4. Оценить зависимость частоты контактов участков генома от их удаленности в линейной молекуле ДНК для сперматозоидов и фибробластов.
5. Оценить влияние компактизации генома сперматозоидов на особенности укладки ДНК этих клеток.

Научная новизна работы. На основе новейшего метода Hi-C впервые получены пространственные карты геномов фибробластов и сперматозоидов мышцы. Важно отметить, что в данной работе впервые исследованы пространственные контакты транскрипционно-неактивных клеток. Впервые показано наличие пространственных доменов в геномах этих клеток. В работе проведено сравнение трехмерной организации геномов соматических и половых клеток с использованием как ранее описанных, так и оригинальных, разработанных автором методов.

Впервые получен список локусов генома сперматозоидов, пространственная укладка которых наиболее значительно отличается от укладки соматических клеток. Более того, в ходе работы разработана методика нормализации частот контактов, учитывающая эффект компактизации генома. Разработанный автором алгоритм статистического сравнения частот контактов, в сочетании с методикой нормализации, позволил адекватно оценить влияние компактизации генома сперматозоидов и других особенностей этих клеток на специфику трехмерной организации ДНК.

Теоретическая и практическая значимость исследования. С теоретической точки зрения, анализ пространственной архитектуры генома сперматозоидов способствует расширению наших знаний о таких фундаментальных вопросах, как роль процессов транскрипции в трехмерной организации ДНК, влияние компактизации генома на структуру пространственных доменов и влияние белков-упаковщиков ДНК (таких как гистоны и протамины) на организацию этой молекулы на макроуровне. Понимание теоретических механизмов, лежащих в основе формирования и поддержания структуры пространственных доменов, помогает объяснить причины заболеваний, связанных с нарушениями пространственной укладки генома и, в перспективе, может способствовать разработкам методов их прогнозирования и лечения.

На защиту выносятся **следующие положения и результаты.**

- 1.Общность принципов пространственной организации геномов соматических и половых клеток выражается в характерной для обоих типов клеток степенной зависимости распределения частот контактов локусов от расстояния между ними в линейной молекуле и наличии топологических доменов в геномах фибробластов и сперматозоидов.
- 2.Особенностью пространственной укладки генома сперматозоидов является увеличение частот как внутрихромосомных контактов между удаленными локусами ДНК, так и межхромосомных взаимодействий.
3. Различия в частотах пространственных контактов в геномах фибробластов и сперматозоидов в 25% случаев объясняются более высокой компактизацией генома последних, а в 75% – другими причинами.

Вклад автора. Автором самостоятельно получены все основные результаты. Hi-C библиотеки фибробластов и сперматозоидов мыши были получены Баттулиным Н.Р. и Хабаровой А.А., при участии автора. Алгоритм анализа зависимости частоты взаимодействий от расстояния между локусами в линейной молекуле был реализован Помазным М.Ю., при участии автора.

Апробация работы. Работа была доложена на двух научных конференциях. По материалам работы опубликованы три статьи в научных журналах. Результаты работы использовались при составлении курса лекций для учащихся СУНЦ НГУ.

Структура и объем работы. Диссертация состоит из оглавления, списка сокращений, введения, обзора литературы, описания используемых материалов и методов, результатов, обсуждения, выводов и списка литературы. Работа изложена на 113 страницах, содержит 14 рисунков, 2 таблицы и 3 приложения.

Материалы и методы

Hi-C библиотеки. Получение Hi-C библиотек было выполнено Баттулиным Н.Р. и Хабаровой А.А. с участием автора, по протоколу TSS, предложенному Калхором с соавторами (Kalhor et al. 2012). Библиотеки были получены из подвижных сперматозоидов мыши C57/Bl, выделенных из придатков семенников и эмбриональных фибробластов мыши, полученных из 12,5-дневных эмбрионов мыши линии C57/Bl. При приготовлении библиотек рестрикция ДНК проводилась ферментом HindIII. Секвенирование образцов проводилось А. М. Мазуром и Е.В. Прохорчуком на платформе Illumina GAL II. Длина ридов (прочтений) составляла 100 п.н. В работе использовались данные Hi-C о эмбриональных стволовых клетках (ЭСК) и кортексе мыши, опубликованные Диксоном с соавторами (Dixon et al. 2012).

Построение матрицы пространственных контактов. Картирование ридов на геном, фильтрация ридов, построение матрицы контактов и её коррекция проводились в соответствии с алгоритмами, описанными в (Imakaev et al. 2012) и (Lieberman-Aiden et al. 2009). Анализ зависимости частоты взаимодействий (P) от расстояния между локусами в линейной молекуле (s) выполнялся как предложено в (Lieberman-Aiden et al. 2009).

Поиск различий в профилях контактов локусов фибробластов и сперматозоидов. Вычисление Евклидова расстояния (E) проводилось по формуле

$$E = \sqrt{\sum_{j=c_{st}..c_{end}} (Sp_{ij} - Fib_{ij})^2}$$
, где E – Евклидово расстояние между бинами (локусами) i и j на хромосоме C, c_{st} – первый бин хромосомы C, c_{end} – последний бин хромосомы C, Sp_{ij} и Fib_{ij} – число контактов между бинами i и j для сперматозоидов и фибробластов соответственно. Чем большее значение E, тем больше считалось различие между бинами.

Для вычисления коэффициентов корреляции Пирсона и Спирмена, использовали внутривромосомные контакты каждого бина в виде одномерного массива. Для локусов, в контактах которых, по каким-либо причинам, соотношение сигнал\шум значительно ниже, чем в других, корреляция контактов будет ниже, чем для локусов с меньшим уровнем шума (будем называть в дальнейшем такие локусы «неинформативные»). Мы предложили решение этой проблемы на основе анализа «референсных» выборок (фактически, метод статистического бутстрэппинга, bootstrapping (Efron 1979)). Под референсными выборками имеются в виду две выборки, сгенерированные из данных Hi-C ЭСК случайным образом, количество данных в которых был сходным с количеством данных для фибробластов и сперматозоидов. Мы определили неинформативные локусы как такие, для которых выполняется $C_i < M - SD$, где C_i – коэффициент корреляции Спирмена/Пирсона для бина i в референсных выборках, M и SD – медиана и стандартное отклонение распределения коэффициентов корреляции Спирмена/Пирсона всех локусов референсных выборок, соответственно. Мы исключили из дальнейшего рассмотрения все неинформативные и смежные с ними бины, и рассчитали коэффициент корреляции для оставшихся локусов фибробластов и сперматозоидов.

Для сравнения полученных значений первого собственного вектора (E^1) фибробластов и сперматозоидов, каждый бин был представлен в виде точки в двухмерном пространстве, X- и Y-координатами которой являлись значение E^1 фибробластов и сперматозоидов. Затем, для полученного множества точек была

подсчитана линия линейной регрессии. Наконец, было подсчитано геометрическое расстояние от этой линии до каждой точки. Это расстояние считалось критерием схожести значений E^1 .

После вычисления Евклидова расстояния, коэффициента корреляции Пирсона и различий E^1 , мы присвоили каждому бину 3 характеристики, соответствующие полученным в этих трех методах сравнения значениям. Таким образом, нами было получено три списка бинов. Мы считали наиболее сильно различающимися локусы, оказавшиеся в пересечении всех трех полученных выборок бинов.

Выявление статистических различий в частотах индивидуальных контактов между сперматозоидами и фибробластами. Предполагая нормальную аппроксимацию биномиального распределения, мы рассчитали критерий (p-value) для нулевой гипотезы. Умножив каждое из значений pval на число проанализированных локусов, мы получили значения q (q-value), учитывающие поправку на множественное тестирование гипотез.

Моделирование «компрессии» генома. Вычисляли значения K_j следующим образом: $K_j = Sp_j / Fib_j$, где Sp_j и Fib_j представляют собой суммы элементов диагонали номер j в матрице контактов сперматозоидов и фибробластов. Затем, генерировали «сжатую» матрицу, умножая каждый элемент диагонали j матрицы контактов фибробластов на соответствующий коэффициент K_j . Мы проводили коррекцию только в том случае, если погрешность определения K_j составляла менее 5%. «Сжатую» матрицу умножали на коэффициент нормировки $K_{norm} = \frac{\sum_{i,j} Fib}{\sum_{i,j} Fib_{compressed}}$, где Fib – исходная матрица контактов, $Fib_{compressed}$ – матрица контактов, полученная в ходе «компрессии».

Идентификация TAD доменов проводилась как предложено в (Dixon et al. 2012). Моделирование пространственного расположения локусов в TAD и расчет HD-параметра доменов проводились как указано в (Hu et al. 2013).

Анализ межхромосомных контактов проводили по аналогии с алгоритмами, описанными ранее (Kalhor et al. 2012; Lieberman-Aiden et al. 2009).

Результаты

Оценка количества и качества данных массового параллельного секвенирования. Мы провели анализ данных массового параллельного секвенирования Hi-C библиотек фибробластов и сперматозоидов мыши. Библиотеки включали около 150 и 400 миллионов ридов для фибробластов и сперматозоидов, соответственно. Мы провели картирование ридов финальной библиотеки на геном и осуществили фильтрацию, позволившую исключить, по крайней мере частично, следующие неспецифические продукты: ПЦР дубликаты, продукты случайного (не опосредованного хроматином) взаимодействия фрагментов ДНК, продукты внутримолекулярного лигирования с образованием колец и т.д. После применения фильтров число ридов сократилось до 36% (54,5 млн) для фибробластов и 5% (19,97 млн) для сперматозоидов. Основываясь на ридах, прошедших фильтры, мы построили матрицы пространственных контактов для сперматозоидов, фибробластов и ЭСК (рис. 1).

Рисунок контактов внутри индивидуальных хромосом имеет явно выраженный «клетчатый» паттерн, что подразумевает наличие внутри хромосом протяженных участков, разделенных крупными областями в линейной молекуле, но взаимодействующих друг с другом в пространстве ядра.

Мы проанализировали размер погрешности определения частоты пространственных контактов, и обнаружили явно выраженный рост этого параметра при увеличении разрешения матрицы (т.е. при уменьшении размера бина). Так, при разрешении 1 Mb (1 миллион п.о.), средняя ошибка для внутривнутрихромосомных контактов составляла ~24 %, а при увеличении разрешения на порядок (до 100 kb, 100 тысяч п.о.) средняя ошибка выросла до ~88 %. Поэтому, в большинстве дальнейших вычислений использовалась матрица, составленная с разрешением 1 Mb.

При визуальном анализе мы обнаружили высокое сходство карт пространственных контактов фибробластов и сперматозоидов (рис. 1). Чтобы подтвердить это сходство, мы провели анализ значений E^1 для этих матриц.

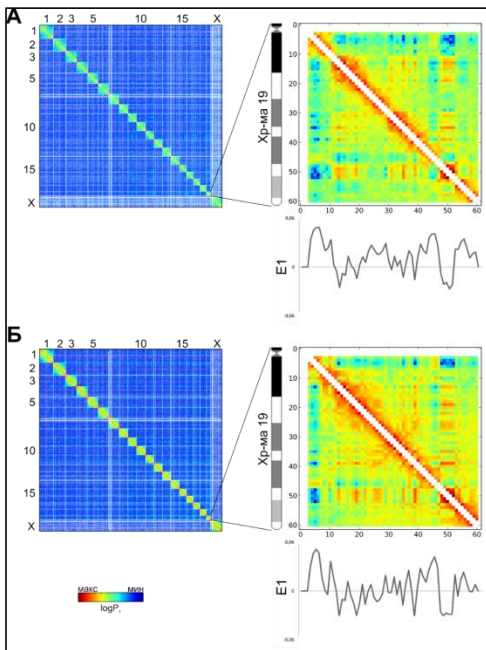


Рисунок 1. Матрицы пространственных контактов фибробластов и сперматозоидов. Показаны полногеномные (слева) контакты фибробластов (А) и сперматозоидов (Б) и внутривнутрихромосомные контакты хромосомы 19 (справа). Сбоку и сверху каждой матрицы контактов указаны номера бинов (последовательно расположенных фрагментов генома), а в ячейках матрицы отображен логарифм относительных частот контактов между парами

бинов ($\log P_c$), в цветовой шкале: красный – максимальный (наибольшая вероятность контактов), синий – минимальный. Сверху матриц указаны номера соответствующих хромосом. Под матрицей контактов хромосомы 19 показаны значения E^1 соответствующих локусов. Матрицы приведены с разрешением 1 Mb, после коррекции.

A/B-домены в геномах сперматозоидов и фибробластов.

Ранее было показано, что значения E^1 характеризуют пространственную организацию локусов в клетках млекопитающих (Imakaev et al. 2012; Lieberman-Aiden et al. 2009). Мы провели вычисление значений собственных векторов для полученных матриц пространственных контактов фибробластов и сперматозоидов, а также матриц ЭСК и клеток кортекса, описанных в (Dixon et al. 2012). Полученные данные хорошо согласуются с визуальным сходством матриц пространственных контактов: коэффициент корреляции Спирмена для значений E^1 сперматозоидов и фибробластов составил 0.878. Следует отметить, что значения E^1 сперматозоидов и других типов клеток для других типов клеток также показывали высокий уровень сходства: коэффициент корреляции для значений E^1 сперматозоидов и ЭСК

составил 0.878, сперматозоидов и кортекса - 0.901. Эти результаты были подтверждены при использовании альтернативной меры сходства двух наборов данных: коэффициента максимальной информации (Reshef et al. 2011).

В работе (Lieberman-Aiden et al. 2009) было показано, что геном может быть поделен на различные по своим характеристикам А- и В- домены, основываясь на значениях E^1 . Чередование областей, характеризующихся положительными и отрицательными значениями E^1 , указывает на присутствие этих доменов в геномах сперматозоидов и фибробластов. А сходство значений E^1 сперматозоидов и фибробластов свидетельствует о сходстве этих доменов в геномах исследованных клеток.

Анализ TAD-доменов в геномах сперматозоидов и фибробластов. Нами был проведен поиск TAD-доменов в геномах фибробластов и сперматозоидов. Мы идентифицировали 2 590 TAD-доменов в геноме фибробластов, средний размер которых составлял 928 kb (медиана распределения размеров доменов составила 680 kb). Число TAD-доменов фибробластов и TAD-доменов, описанных ранее для ЭСК (2 200 доменов с медианой 880 kb (Dixon et al. 2012)) было сходным. Интересно, что число доменов, идентифицированных в сперматозоидах, оказалось несколько меньшим: 1 856 доменов со средним размером 1 226 kb и медианой 1 000 kb.

Проанализировав значение переменной DI (domain index, переменная, значение которой вычисляется для каждого участка генома и косвенно характеризует его принадлежность к домену; описана более подробно в (Dixon et al. 2012)) в данных фибробластов и сперматозоидов, мы обнаружили высокий уровень сходства (коэффициент корреляции Спирмена 0,759). Это свидетельствует о высоком сходстве TAD-доменов фибробластов и сперматозоидов. Сходство также подтверждается при визуальном анализе TAD-доменов этих клеток (рис. 2). Мы обнаружили, что, в ряде случаев, расположенные последовательно в геноме фибробластов домены «сливаются» друг с другом в геноме сперматозоидов (рис 2). Этот эффект может частично объяснить увеличение средних размеров TAD генома сперматозоидов по сравнению с фибробластами.

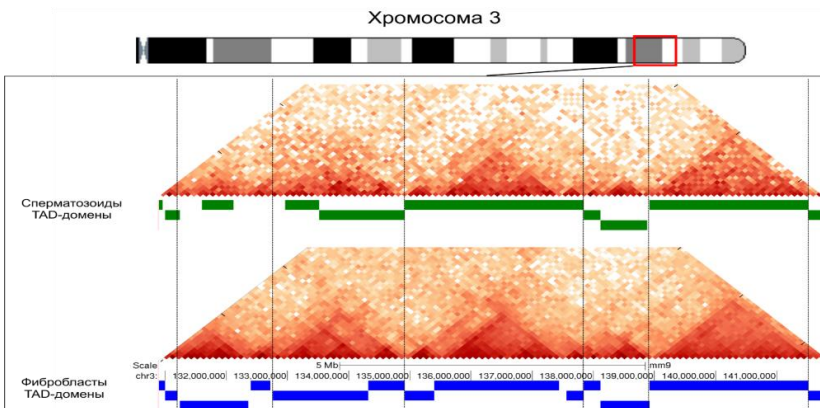


Рисунок 2. TAD-домены фибробластов и сперматозоидов. Репрезентативный участок генома (район хромосомы 3), представлен матрицей пространственных контактов с разрешением 40 000 п.о. Показан участок матрицы с одной стороны от главной диагонали, развернутый на 90 градусов. Вероятность контактов показана цветом в красно-белой шкале (красный – наибольшая вероятность). Границы доменов показаны под матрицей в виде прямоугольников: синего цвета для фибробластов, зеленого для сперматозоидов. Вертикальными линиями отмечены три различных региона, на одном из них TAD-домены фибробластов и сперматозоидов различаются (слева), на другом – домены фибробластов сливаются в сперматозоидах (в центре) и на третьем домены совпадают (справа).

Поиск различий в пространственной укладке геномов сперматозоидов и фибробластов. Мы использовали ряд методов для поиска отдельных локусов, контакты которых различаются в фибробластах и сперматозоидах: коэффициенты корреляции Пирсона и Спирмена, Евклидово расстояние и близость значений E^1 . Общее сходство пространственной укладки сперматозоидов и фибробластов было выявлено всеми тремя методами: оно выражается в высоких коэффициентах корреляции, небольшом Евклидовом расстоянии и корреляции значений E^1 .

Мы идентифицировали бины, различия между которыми были выявлены всеми тремя методами. Было найдено 7 таких локусов, на хромосомах 5, 12, 13 и 19, причем три из семи участков оказались расположенными на хромосоме 19.

Пространственная организация генома может различаться не только на уровне укладки целых локусов, но и за счет индивидуальных различий в попарных контактах отдельных

локусов. Мы протестировали в сперматозоидах и фибробластах все контакты, представленные более чем 1 рядом («информативные» контакты). Из 153 363 информативных для сперматозоидов и фибробластов контактов, 8 947 (5,85%) имели статистически значимые различия в частотах взаимодействия (q -value <0.05). Более того, из этих 8 947 взаимодействий, для 6 586 контактов частоты различались более чем в 2 раза. Следует отметить, что вышеупомянутые участки хромосомы 19, различающиеся по укладке в сперматозоидах и фибробластах, имели в этих клетках большое количество статистически значимо различающихся контактов с различными бинами генома

Анализ зависимости частоты контактов локусов от расстояния в линейной молекуле. Анализ зависимости частоты контактов локусов (P) от расстояния в линейной молекуле (s), $P(s)$, позволяет сделать определить наиболее вероятный тип укладки ДНК в ядре (Mirny 2011; Lieberman-Aiden et al. 2009; Naumova et al. 2013). Мы рассчитали зависимость $P(s)$ для сперматозоидов и фибробластов (рис. 3). И для тех, и для других, наблюдалось сильно выраженное уменьшение числа контактов с ростом расстояния между локусами в линейной молекуле. Для сперматозоидов зависимость выражалась как $P(s) \sim s^{-1,07}$, для фибробластов - $P(s) \sim s^{-1,27}$ (рис. 12, А). Мы оценили стандартную ошибку для показателей степени (-1,07 и -1,27) как не более, чем 0,01 и показали, что эти значения статистически значимо различаются между собой. Полученные значения также значимо отличались от значения -1 ($P(s) \sim s^{-1}$), характерного для гипотетической идеальной фрактальной глобулы ДНК. Однако тип упаковки сперматозоидов оказался ближе к фрактальному, чем упаковка фибробластов.

Интересно, в сперматозоидах наблюдалось увеличение частоты контактов локусов, удаленных на расстояние более 40 Мб. Для регионов, разделенных 50-150 Мб, сперматозоиды показывали более чем двукратное увеличение частот контактов, по сравнению с фибробластами. Это означает, что сперматозоиды имеют больше контактов между удаленными локусами, чем фибробласты.

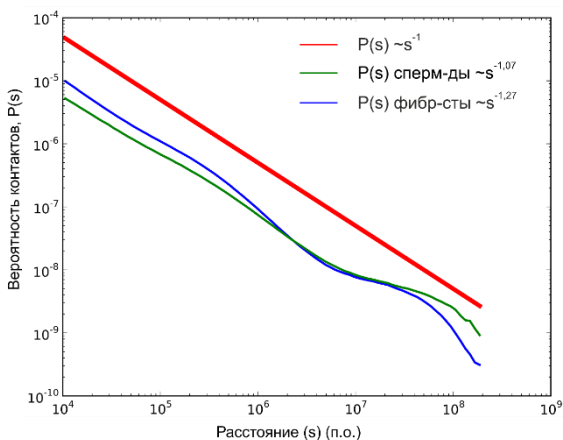


Рисунок 3. Геном сперматозоидов обогащен взаимодействиями удаленных участков. Приведен график зависимости $P(s)$ для фибробластов, сперматозоидов и теоретический график, ожидаемый для идеальной фрактальной глобулы.

Чтобы выявить влияние различий $P(s)$ на характеристики пространственной организации, мы, используя алгоритм WACH (Hu et al. 2013), смоделировали пространственную структуру выявленных в сперматозоидах и фибробластах TAD-доменов. После моделирования, мы представили каждый TAD-домен как цилиндр, и измерили соотношение длины такого цилиндра и его диаметра (HD-ratio). У более вытянутых цилиндров такое соотношение будет больше, чем у более компактных. Мы обнаружили, что TAD-домены сперматозоидов являются более компактными (имеют меньшее HD-ratio), чем TAD-домены фибробластов (последние являются более вытянутыми).

Анализ межхромосомных контактов геномов фибробластов и сперматозоидов. В ряде работ было показано, что количество межхромосомных контактов много меньше, чем количество внутривхромосомных (Lieberman-Aiden et al. 2009; Kalhor et al. 2012; Rao et al. 2014). Мы обнаружили, что для всех хромосом фибробластов и сперматозоидов наблюдаются две одинаковые тенденции. Во-первых, число межхромосомных контактов в 10-40 раз меньше, чем внутривхромосомных. Во-вторых, в сперматозоидах это соотношение (внутри- к межхромосомным) ниже, чем в фибробластах: для сперматозоидов составляет 10-20 раз, тогда как для фибробластов – 20-40 раз. Это означает, что в сперматозоидах наблюдается много больше межхромосомных контактов, чем в фибробластах.

Мы также оценили частоты контактов индивидуальных хромосом друг с другом (рис. 4). Оказалось, что длинные хромосомы (1-5 и хромосома X) имеют тенденцию взаимодействовать друг с другом больше, чем с короткими (хромосомы 10-19). Эта тенденция может быть визуализирована как обогаченный красным сигнал квадрат в левом верхнем углу матриц, представленных на рисунке 4. Наблюдалась ещё одна аналогичная тенденция – увеличение частоты контактов длинных хромосом друг с другом, по сравнению с частотами контактов длинных хромосом с короткими. Однако, это тенденция была менее выраженной. Мы подтвердили полученные результаты о предпочтениях межхромосомных контактов, зависящих от длин хромосом, проведя анализ распределения частот межхромосомных контактов от соотношения их длин. Нами была обнаружена обратная корреляция этих двух параметров, характеризующаяся коэффициентом корреляции Пирсона $-0,44$.

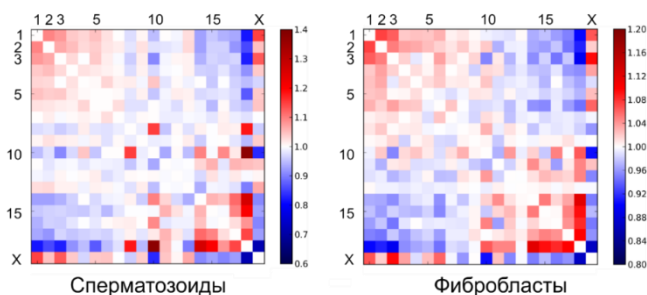


Рисунок 4. Анализ межхромосомных контактов в геномах фибробластов и сперматозоидов. Каждый элемент матрицы относится

к контактам хромосом, указанных в соответствующей строке и столбце. Элементы матрицы отражают соотношение полученной в эксперименте частоты контактов к ожидаемому для случайного (равномерного) распределения, в соответствии с цветовой шкалой приведенной сбоку. Красный цвет соответствует превышению числа контактов над ожидаемым, синий – уменьшению по сравнению с ожидаемым.

Влияние компактизации генома сперматозоидов на особенности пространственной организации этих клеток. Различия в пространственной организации сперматозоидов и фибробластов, описанные выше, могут иметь как минимум две потенциальные причины. Во-первых, характерная для сперматозоидов более плотная упаковка генома, связанная с уменьшением размера ядра и более плотной упаковкой ДНК при

помощи протаминов, приводит к локус-неспецифическим изменениям контактов в масштабе всего генома и может, в конечном счете, быть причиной различий в частотах индивидуальных контактов. Во-вторых, различия в частотах контактов могут объясняться локальными, локус-специфическими изменениями в структуре пространственных взаимодействий, имеющие значение для функционирования данных локусов.

Чтобы оценить роль этих причин, мы разработали метод «компрессии» генома соматических клеток (фибробластов) в соответствии с параметрами сперматозоидов. Проводя «компрессию» генома соматических клеток, мы хотим получить теоретическую клетку, ДНК которой компактизована также, как и в сперматозоидах, но содержит локус-специфические контакты, характерные для исходных соматических клеток. Идея метода состоит в том, чтобы, не меняя распределение частот контактов для локусов, расположенных в линейной молекуле на одинаковом расстоянии, сблизить все локусы друг с другом. Таким образом «компрессия» фибробластов в соответствии с параметрами сперматозоидов приводит к тому, что функция $P(s)$ для этих клеток становится одинаковой, однако при этом в фибробластах сохраняется соотношение контактов для локусов, расположенных на одинаковом расстоянии друг от друга.

Мы обнаружили, что число значимо различающихся контактов между C_{sp} -фибробластами («компрессированные» по типу сперматозоидов фибробласты) и сперматозоидами после компрессии уменьшилось приблизительно на 25%. В качестве контроля, мы выполнили компрессию фибробластов по типу ЭСК, получив C_{ESC} -фибробласты. Общее число статистически различающихся контактов в сперматозоидах и C_{ESC} -фибробластах оказалось на 20% больше, чем число различий между сперматозоидами и фибробластами до компрессии. Следует отметить, что при этом C_{ESC} -фибробласты были более сходны с ЭСК, чем исходные фибробласты. Это свидетельствует о том, что «компрессия» уменьшает число различий в пространственной организации между разными типами клеток, только если она проводится в соответствии со специфическими для данных типов клеток параметрами.

Выводы

1. Метод высокоэффективного конформационного захвата хромосом, в сочетании с биоинформационными алгоритмами фильтрации и нормализации данных, позволил впервые получить полногеномные карты пространственных контактов для фибробластов и сперматозоидов мыши с разрешением не менее одного миллиона п.о.

2. Анализ карт пространственных контактов фибробластов и сперматозоидов мыши позволил идентифицировать А- и В-компарменты и топологические домены, характерный размер которых составляет 680 и 1000 тысяч нуклеотидов, соответственно, в геномах исследованных клеток.

3. Сравнение пространственной укладки геномов фибробластов и сперматозоидов, проведенное тремя независимыми методами, показало высокий уровень сходства их организации. Статистический анализ различий частот индивидуальных пространственных контактов обнаружил 5% взаимодействий, различающихся в фибробластах и сперматозоидах. Выявлено семь локусов на хромосомах 5, 12, 13 и 19, трехмерная организация которых наиболее значительно различается в фибробластах и сперматозоидах.

4. Зависимость частоты контактов локусов от расстояния между ними в линейной молекуле для фибробластов и сперматозоидов характеризуется степенной функцией с коэффициентами степени -1,27 и -1,07. Укладка ДНК сперматозоидов лучше соответствует фрактальной модели, чем организация ДНК фибробластов. В ДНК сперматозоидов выявлено увеличение частоты контактов между локусами, располагающимися на расстоянии более 40 миллионов п.о., в сравнении с геномом фибробластов, что свидетельствует о большей степени компактизации генома половых клеток.

5. Разработана оригинальная модель «компрессии» генома фибробластов, позволившая установить, что около четверти выявленных различий в частотах пространственных контактов могут быть связаны с большим уровнем компактизации генома сперматозоидов, по сравнению с фибробластами.

6. Выявлены преференции в распределении межхромосомных контактов в геномах фибробластов и сперматозоидов, заключающиеся в преимущественных контактах либо длинных

хромосом 1-6 и хромосомы X, либо коротких хромосом 15-19 между собой. В целом геном сперматозоидов демонстрирует большее количество межхромосомных контактов, чем геном фибробластов, что может быть следствием его большей компактизации.

7. Несмотря на поразительные морфологические и функциональные различия, геномы фибробластов и сперматозоидов имеют сходные параметры пространственной архитектуры, что указывает на консерватизм принципов трехмерной организации ДНК в интерфазном ядре.

Список работ, опубликованных по теме диссертации

1. Н.Р. Баттулин, **В.С. Фишман**, Ю.Л. Орлов, А.Г. Мензоров, Д.А. Афонников, О.Л. Серов 3С-методы в исследованиях пространственной организации генома. // Вавиловский журнал генетики и селекции. – 2012. – Т 16. - № 4/2. - с 872-878.

2. N.R. Battulin, V.S. Fishman, A.A. Khabarova, M.Yu. Pomaznoy, T.A. Shnaider, D.A. Afonnikov, O.L. Serov, Investigation of the spatial genome organization of mouse sperm and fibroblasts by the Hi-C method // Russian Journal of Genetics: Applied Research. – 2014. - V 4. - № 6. - p 556-560.

3. N. Battulin*, **V.S. Fishman***, A.M. Mazur, M. Pomaznoy, A.A. Khabarova, D.A. Afonnikov, E.B. Prokhortchouk, O.L. Serov. Comparison of the three-dimensional organization of sperm and fibroblast genomes using the Hi-C approach // Genome Biology, 2015, V 16, I 77, doi:10.1186/s13059-015-0642-0

* - равный вклад авторов

4. N. Battulin , **V.S. Fishman**, A.M. Mazur, M. Pomaznoy, A.A. Khabarova, D.A. Afonnikov, E.B. Prokhortchouk, O.L. Serov. Comparison of the 3D organization of sperm and fibroblast genomes using the Hi-C approach // Сборник трудов IV международной научно-практической конференции “Постгеномные методы анализа в биологии, лабораторной и клинической медицине”, стр. 15.