

**Жегалова Ирина Владимировна**

**Анализ трехмерной структуры хроматина эукариот**

Специальность 1.5.8. —  
«Математическая биология, биоинформатика»

Автореферат  
диссертации на соискание учёной степени  
кандидата биологических наук

Москва — 2024

Работа выполнена на факультете биоинженерии и биоинформатики МГУ имени М.В.Ломоносова и УНЦ «Биоинформатика» ИППИ РАН.

Научные руководители: доктор биологических наук, профессор  
**Гельфанд Михаил Сергеевич**

кандидат биологических наук  
**Храмеева Екатерина Евгеньевна**

Официальные оппоненты: **Васецкий Егор Сергеевич**,  
доктор биологических наук,  
Институт Гюстава Русси, Вильжюиф, Франция,  
старший научный сотрудник, директор научных исследований

**Баттулин Нариман Рашитович**,  
кандидат биологических наук,  
Институт цитологии и генетики Сибирского  
отделения Российской академии наук,  
ведущий научный сотрудник, заведующий  
лабораторией генетики развития

Ведущая организация: Федеральное государственное бюджетное  
образовательное учреждение высшего образова-  
ния «Высшая школа экономики»

Защита состоится 16 сентября 2024 г в 15-00 на заседании диссертационного  
совета Д24.1.101.01 при Федеральном государственном бюджетном  
учреждении науки Институт проблем передачи информации им. 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.

Отзывы на автореферат в двух экземплярах, заверенные печатью учреждения,  
просьба направлять по адресу: 127051, г. Москва, Большой Каретный переулок,  
д.19 стр. 1, ученому секретарю диссертационного совета Д 24.1.101.01.

Телефон для справок: +7 (495) 694-14-77.

Ученый секретарь  
диссертационного совета  
Д 24.1.101.01,  
доктор физико-математических наук Казенников Олег Васильевич

## Общая характеристика работы

**Актуальность темы.** *Dictyostelium discoideum* — гаплоидная амеба, живущая в лесной почве и питающаяся бактериями. При голодании этот одноклеточный организм начинает синтезировать цАМФ, что приводит к агрегации отдельных одноклеточных организмов в единый многоклеточный. Эта программа развития через 24 часа завершается формированием многоклеточного плодового тела и спор. Сходство между *D. discoideum* и клетками животных распространяется и на гены и белки, которые регулируют разнообразные консервативные процессы, такие как деление и рост клеток, хемотаксис, эндоцитоз, факультативная многоклеточность и взаимодействия между хозяином и патогеном<sup>1</sup>.

Более того, *D. discoideum* является важным объектом для изучения различных заболеваний человека, и именно эксперименты с ним вследствие легкости введения мутаций в гаплоидный организм и простоты культивации помогли внести важный вклад в изучение различных нозологий и состояний, например, устойчивости к химиотерапии опухоли<sup>2</sup>, биполярного расстройства<sup>3</sup>, митохондриальных заболеваний<sup>4</sup>, старения<sup>5</sup>, болезни Альцгеймера<sup>6</sup>, болезни Паркинсона<sup>7</sup> и других нейродегенеративных заболеваний<sup>8</sup>.

<sup>1</sup>Barry, N. P. *Dictyostelium amoebae and neutrophils can swim* / N. P. Barry, M. S. Bretscher // Proceedings of the National Academy of Sciences. 2010. Vol. 107, no. 25. P. 11376—11380; Maniak, M. *Dictyostelium as a model for human lysosomal and trafficking diseases* / M. Maniak // Seminars in Cell & Developmental Biology. 2011. Vol. 22, no. 1. P. 114—119. (GTPases in Intracellular Trafficking); Loomis, W. F. *Cell signaling during development of Dictyostelium* / W. F. Loomis // Developmental Biology. 2014. Vol. 391, no. 1. P. 1—16; Nichols, J. M. *Chemotaxis of a model organism: progress with Dictyostelium* / J. M. Nichols, D. Veltman, R. R. Kay // Current Opinion in Cell Biology. 2015. Vol. 36. P. 7—12. (Cell adhesion and migration); Tosetti, N. *Amoebae as a tool to isolate new bacterial species, to discover new virulence factors and to study the host–pathogen interactions* / N. Tosetti, A. Croxatto, G. Greub // Microbial Pathogenesis. 2014. Vol. 77. P. 125—130.

<sup>2</sup>Curcumin affects gene expression and reactive oxygen species via a PKA dependent mechanism in *Dictyostelium discoideum* / W. S. Swatson [et al.] // PLOS ONE. 2017. Vol. 12, no. 11. e0187562.

<sup>3</sup>Identifying an uptake mechanism for the antiepileptic and bipolar disorder treatment valproic acid using the simple biomedical model *Dictyostelium* / N. Terbach [et al.] // Journal of Cell Science. 2011. Vol. 124, no. 13. P. 2267—2276.

<sup>4</sup>The *Dictyostelium* model for mitochondrial disease / L. M. Francione [et al.] // Seminars in Cell & Developmental Biology. 2011. Vol. 22, no. 1. P. 120—130. (GTPases in Intracellular Trafficking).

<sup>5</sup>Integrated actions of mTOR complexes 1 and 2 for growth and development of *Dictyostelium* / P. Jaiswal [et al.] // International Journal of Developmental Biology. 2019. Vol. 63, no. 8. P. 521—527. <sup>6</sup>Myre, M. A. *Functional Analysis of Proteins Involved in Neurodegeneration Using the Model Organism Dictyostelium: Alzheimer's, Huntington's, and Batten Disease* / M. A. Myre, R. J. Huber, D. H. O'Day // Molecular-Genetic and Statistical Techniques for Behavioral and Neural Research / ed. by R. T. Gerlai. San Diego : Academic Press, 2018. P. 491—518.

<sup>7</sup>The Parkinson's Disease-Associated Protein DJ-1 Protects *Dictyostelium* Cells from AMPK Dependent Outcomes of Oxidative Stress / S. Chen [et al.] // Cells. 2021. Aug. Vol. 10, no. 8. P. 1874.

<sup>8</sup>*Dictyostelium*, a microbial model for brain disease / S. J. Annesley [et al.] // P. 1413—1432.

Губки (тип Porifera) являются донными животными и представляют, по разным оценкам, один из или самый древний тип существующих Metazoa. Класс Demospongiae включает около 85 % всех существующих и вымерших видов губок, распространённых от тропических до северных морей<sup>9</sup>; губка *Halisarca dujardini*, принадлежащая к классу Demospongiae, обитает в холодных северных морях. Древность губок делает их важным объектом исследования эволюционных процессов, поскольку губки являются одной из базальных ветвей Metazoa. Несмотря на упрощённую организацию, геномы губок имеют многие общие черты с представителями Eumetazoa. Например, было показано, что геном губки *Amphimedon queenslandica* содержит некоторые гены эмбрионального развития<sup>10</sup>, изменения экспрессии по стадиям развития схожи с таковыми у представителей Eumetazoa<sup>11</sup>. Более того, клетки губок обладают способностью переходить из одного типа в другой, подобно трансдифференцированным и стволовым клеткам млекопитающих<sup>12</sup>.

Другой уникальной особенностью губок является их способность к реагрегации из диссоциированных клеток, что было впервые обнаружено Вильсоном в 1907 году<sup>13</sup>. Показано, что *H. dujardini* проявляет высокую способность к реагрегации, поддерживаемую защитными и регенеративными механизмами<sup>14</sup>. Исследование *H. dujardini* предоставляет уникальную возможность расширения знаний о клеточной биологии и адаптивных стратегиях у самых ранних представителей многоклеточных организмов, а также о клеточных основах регенерации тканей.

Трисомии, т.е. наличие дополнительной копии одной из хромосом, в большинстве случаев являются летальными для эмбрионов, однако трисомии по хромосомам 13, 18 и 21 присутствуют у новорожденных с частотой

<sup>9</sup>Deep Phylogeny and Evolution of Sponges (Phylum Porifera) / G. Worheide [et al.] // *Adv Mar Biol.* 2012. Vol. 61. P. 1—78. pmid: 22560777.

<sup>10</sup>Deep Phylogeny and Evolution of Sponges (Phylum Porifera) / G. Worheide [et al.] // *Adv Mar Biol.* 2012. Vol. 61. P. 1—78. pmid: 22560777; The Demosponge *Amphimedon Queenslandica*: Reconstructing the Ancestral Metazoan Genome and Deciphering the Origin of Animal Multicellularity / B. M. Degnan [et al.] // *Cold Spring Harb Protoc.* 2008. Vol. 2008, no. 12. pdb.emo108. pmid: 21356734.

<sup>11</sup>Dynamic and Widespread lncRNA Expression in a Sponge and the Origin of Animal Complexity / F. Gaiti [et al.] // *Molecular Biology and Evolution.* 2015. Vol. 32, no. 9. P. 2367—2382; The Mid-Developmental Transition and the Evolution of Animal Body Plans / M. Levin [et al.] // *Nature.* 2016. Mar. Vol. 531, issue 7596, no. 7596. P. 637—641.

<sup>12</sup>Transdifferentiation and Mesenchymal-to-Epithelial Transition during Regeneration in Demospongiae (Porifera) / A. V. Ereskovsky [et al.] // *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution.* 2020. Vol. 334, no. 1. P. 37—58; Pluripotency and the Origin of Animal Multicellularity / S. Sogabe [et al.] // *Nature.* 2019. June. Vol. 570, issue 7762, no. 7762. P. 519—522; Whole-Body Regeneration in Sponges: Diversity, Fine Mechanisms, and Future Prospects / A. Ereskovsky [et al.] // *Genes.* 2021. Apr. Vol. 12, issue 4, no. 4. P. 506.

<sup>13</sup>Wilson, H. V. On Some Phenomena of Coalescence and Regeneration in Sponges / H. V. Wilson // *Journal of Experimental Zoology.* 1907. Vol. 5, no. 2. P. 245—258.

<sup>14</sup>Transdifferentiation Is a Driving Force of Regeneration in *Halisarca Dujardini* (Demospongiae, Porifera) / I. E. Borisenko [et al.] // *PeerJ.* 2015. Vol. 3. e1211.

от 1:5000 до 1:1000<sup>15</sup>; реже встречаются мозаичные (представленные не во всех клетках тела) трисомии хромосом 8, 16, 19 и 22. Фенотипические проявления включают полиорганные нарушения: задержки развития, тяжёлые пороки сердца, рецидивирующие инфекции, а также предрасположенность к аутоиммунным заболеваниям. Существует несколько гипотез о роли дополнительной копии хромосомы в возникновении подобных фенотипов. Одна из них постулирует, что существуют дозозависимые гены, дополнительная копия которых ведёт к повышенной экспрессии и объясняет фенотип. Однако более поздние наблюдения поставили под сомнение эту гипотезу, показав, что сверхэкспрессия ограниченного числа генов недостаточна для проявления всех симптомов, представленных у людей с трисомиями<sup>16</sup>. В качестве альтернативы была предложена концепция нарушения клеточного гомеостаза<sup>17</sup>: фенотипические проявления трисомий обусловлены не приобретением конкретной дополнительной хромосомы, а наличием дополнительной хромосомы как таковой. Одной из характерных особенностей трисомий является пангеномная дисрегуляция транскрипции, которая может быть причиной нарушения гомеостаза.

Несмотря на то, что мы хорошо знаем, как устроены отдельные гены *D. discoideum* и как регулируется их экспрессия<sup>18</sup>, мы плохо понимаем, как устроена регуляция экспрессии у *D. discoideum* на уровне всего генома, а также не имеем сведений о его трёхмерной организации. Для губки *H. dujardini* отсутствует информация как о регуляции экспрессии генов, так и о трёхмерной организации генома. Выбор организмов для исследования обусловлен тем, что губки — одни из самых простых многоклеточных, способных к реагрегации, а *D. discoideum* при голодании запускает программу агрегации одноклеточных организмов с последующей дифференциацией на

<sup>15</sup>Current Estimate of Down Syndrome Population Prevalence in the United States / A. P. Presson [et al.] // J Pediatr. 2013. Oct. Vol. 163, no. 4. P. 1163—1168. pmid: 23885965; Survival and Surgical Interventions for Children With Trisomy 13 and 18 / K. E. Nelson [et al.] // JAMA. 2016. Vol. 316, no. 4. P. 420—428. pmid: 27458947.

<sup>16</sup>A Chromosome 21 Critical Region Does Not Cause Specific Down Syndrome Phenotypes / L. E. Olson [et al.] // Science. 2004. Vol. 306, no. 5696. P. 687—690. pmid: 15499018; Down Syndrome / S. E. Antonarakis [et al.] // Nat Rev Dis Primers. 2020. Vol. 6, no. 1. P. 9. pmid: 32029743.

<sup>17</sup>Krivega, M. Consequences of Chromosome Gain: A New View on Trisomy Syndromes / M. Krivega, C. M. Stiefel, Z. Storchova // Am J Hum Genet. 2022. Vol. 109, no. 12. P. 2126—2140. pmid: 36459979.

<sup>18</sup>Sequence-specific protein interaction with a transcriptional enhancer involved in the autoregulated expression of cAMP receptor 1 in Dictyostelium / X. Mu [et al.] // Development. 1998. Vol. 125, no. 18. P. 3689—3698; Hori, R. Identification and characterization of multiple A/T-rich cisacting elements that control expression from Dictyostelium actin promoters: the Dictyostelium actin upstream activating sequence confers growth phase expression and has enhancer-like properties / R. Hori, R. A. Firtel // Nucleic Acids Research. 1994. Vol. 22, no. 23. P. 5099—5111.

клеточные типы, а потому важно установить изменения структуры хроматина и координированы ли эти изменения с экспрессией генов при переходе от одноклеточности к клеточному агрегату. С другой стороны, изучение хроматина образцов с трисомиями человека поможет пролить свет на изменение укладки хроматина при наличии дополнительной копии хромосом в сравнении с образцами нормального кариотипа, что, с одной стороны, помогает понять патогенез заболевания, а с другой стороны, также является исследованием координированности между изменением организации хроматина и экспрессии генов по всему геному.

**Целью** данной работы является изучить устройство хроматина у *D. discoideum*, *H. dujardini* и образцов *Homo sapiens* с трисомиями по хромосомам 13, 16 и 19 на разных уровнях организации, а также взаимосвязь изменений трёхмерной организации с изменениями транскрипционной программы.

Для достижения поставленной цели необходимо было решить следующие **задачи**:

1. Проанализировать данные Hi-C, полученные для стадий развития *D. discoideum* и стадий реагрегации *H. dujardini*.
2. Изучить организацию хроматина на уровне межхромосомных взаимодействий, компартментов и экструзионных взаимодействий для *D. discoideum* и *H. dujardini*.
3. Исследовать взаимосвязи найденных хроматиновых структур с экспрессией генов, используя данные экспрессии для *D. discoideum*.
4. Провести сравнительный анализ хроматинового окружения гена на примере генов ферритинов у *H. dujardini* и других представителей Metazoa.
5. Изучить влияние дополнительной копии хромосом на организацию ДНК в ядре на примере образцов с трисомиями у человека в сравнении с образцами нормального кариотипа.
6. Изучить взаимосвязи хроматина и систем регуляции транскрипции, используя данные эпигенетики, полученные для *D. discoideum*, а также данные о лямбда-репликации для образцов с трисомиями у человека.

#### **Научная новизна:**

1. Впервые проанализировано устройство хроматина у гаплоидной амёбы *D. discoideum* и морской губки *H. dujardini*.
2. Впервые обнаружены вытянутые петли в хроматине *D. discoideum* и показана их связь с уровнями транскрипции.
3. Выполнено оригинальное исследование дальних контактов (преимущественно межхромосомных) хроматина *D. discoideum* и показана их взаимосвязь с длинными некодирующими РНК, значительно меняющими свои уровни экспрессии.
4. Впервые выдвинута гипотеза о конвергентной транскрипции как механизме формирования петель у *D. discoideum*.

5. Впервые аннотированы регуляторные элементы в геноме *D. discoideum* и показано их обогащение в основаниях петель.
6. Впервые аннотированы хроматиновые фонтаны в хроматине морской губки *H. dujardini*.
7. Впервые проанализирована пространственная структура образцов с трисомиями по хромосомам 13, 16 и 18 у человека.
8. Впервые показана связь изменения частот контактов хромосом с трисомиями по хромосомам 13, 16 и 18 с содержанием ламины ассоциированных доменов на таких хромосомах.

**Практическая значимость** *D. discoideum* — важный модельный организм и один из примеров независимого возникновения многоклеточности. При этом многие процессы в нем протекают похоже на высших эукариот, при чем отвечают за них гомологичные гены и белки. Совместно с простотой культивирования это делает *D. discoideum* важным объектом для изучения как фундаментальных процессов, таких как хемотаксис, эндоцитоз, рост и деление, так и различных нозологий и состояний. Однако при большой популярности данного модельного организма информация о трехмерной укладке его хроматина в литературе отсутствовала.

В настоящей работе нами была исследована трехмерная структура хроматина *D. discoideum* и *H. dujardini* на разных стадиях развития. Разработанная методология может быть применена для изучения устройства хроматина других беспозвоночных. Кроме того, настоящее исследование не только восполняет пробел об устройстве хроматина двух важных организмов, но и проливает свет на разнообразие устройства хроматина организмов в отсутствие белка CTCF. Результаты работы показали, что конвергентная пара генов может быть рассмотрена как барьер для экстрезии в других организмах.

Исследование, проведенное для образцов с трисомиями по хромосомам 13, 16 и 18 у человека, позволяет дополнить картину патогенеза анеуплоидий, добавляя наблюдения о глобальной реорганизации хромосом в ядре и изменении взаимодействия с белками ламины к имеющимся гипотезам о дозозависимых генах и нарушении гомеостаза.

**Основные положения, выносимые на защиту:**

1. Хроматин *D. discoideum* образует множество последовательных петель с высокой консервативностью между стадиями и низким уровнем иерархичности.
2. Основания хроматиновых петель *D. discoideum* обогащены энхансер-подобными элементами и границами синтенных блоков.
3. Хроматин губки *H. dujardini* образует фонтаны; различные хроматиновые петли и топологически ассоциированные домены отсутствуют.

4. Наличие дополнительной копии хромосом вызывает как изменения в межхромосомных взаимодействиях, так и реорганизацию локальной упаковки хроматина, причём степень этих изменений варьирует для различных хромосом и, по-видимому, связана с изменениями взаимодействий с белками ламины.

**Достоверность** полученных результатов основана на использовании современных и адекватных методов обработки полученных данных. Результаты находятся в соответствии с результатами, полученными другими авторами.

**Апробация работы.** Основные результаты работы докладывались на: Всероссийской конференции «Информационные технологии и системы» (Пермь, 2019; Москва, 2021, Москва, 2022), 10-ой Московской конференции по вычислительной молекулярной биологии МССМБ-2021 (Москва, 2021), ИТИС-2021, 13-ой Международной мультikonференции «Биоинформатика геномной регуляции и структурной/системной биологии» BGRS-2022 (Новосибирск, 2022).

**Личный вклад.** Автор принимал ведущее участие в анализе данных Hi-C, ATAC-seq, RNA-seq и ChIP-seq, включая разработку и реализацию алгоритмов и скриптов для анализа данных, предварительную обработку и контроль качества, статистический анализ, визуализацию и интерпретацию результатов. Автором были аннотированы последовательные и вытянутые петли, конвергентные пары генов, энхансер-подобные элементы у *D. discoideum*, хроматиновые фонтаны *H. dujardini*, регионы с измененной упаковкой для образцов с трисомиями. Автором были исследованы особенности упаковки хроматина *D. discoideum* на вегетативной стадии. Автором было изучено хроматиновое окружение генов ферритинов у *H. dujardini* и представителей Metazoa. Автором был проведён анализ взаимосвязи изменения частот контактов между хромосомами и изменение локальной упаковки при наличии дополнительной копии хромосомы у образцов с трисомиями.

Подготовка материалов к публикации и написание статей проводилось при личном участии автора. Подготовка библиотек ДНК и РНК для экспериментов Hi-C и RNA-seq для *H. dujardini* была проведена А.В. Черкасовым и А.В. Рябовой. Подготовка библиотек ДНК и РНК для экспериментов Hi-C и RNA-seq *D. discoideum* была выполнена С.В. Ульяновым. Синхронизация клеточной культуры *D. discoideum* по клеточному циклу была проведена П.А. Васильевым. Подготовка библиотек для экспериментов Hi-C у образцов с трисомией была проведена С.В. Ульяновым, эксперимент FISH выполнен А.С. Штомпель.

**Публикации.** Основные результаты по теме диссертации изложены в 9 печатных изданиях, из которых 3 изданы в научных журналах, индексируемых Web of Science и Scopus, 5 — в тезисах докладов. 8

Диссертационная работа была выполнена при поддержке гранта РФ ФИ № 20-34-90058 по теме: «Изучение механизма формирования и функционального назначения хроматиновых петель у низших эукариот на примере диктиостелиума».

**Объем и структура работы.** Диссертация состоит из введения, трех глав, заключения и списка литературы. Полный объем диссертации 170 страниц текста с 62 рисунками и 9 таблицами. Список литературы содержит 301 наименованием.

## Содержание работы

Во **введении** обосновывается актуальность исследований, проводимых в рамках данной диссертационной работы, приводится краткий обзор научной литературы по изучаемой проблеме, формулируется цель, ставятся задачи работы, излагается научная новизна и практическая значимость представляемой работы.

**Первая глава** посвящена обзору литературы по тематике диссертации: особенность биологии *D. discoideum* и *H. dujardini*, их важность как организмов для изучения фундаментальных биологических процессов и заболевания человека. Уделено внимание особенностям генома и эпигенетическим особенностям; подробно разобраны работы, посвященные изучению принципов работы транскрипционных факторов у *D. discoideum* и *H. dujardini*. Отдельно рассмотрены белки ферритины, в том числе у беспозвоночных животных. Уделено внимание анеуплоидиям человека, как фенотипическим проявлениям, так и предполагаемым молекулярным нарушениям, лежащим в основе таких проявлений.

Далее в обзоре рассмотрено устройство хроматина, выделено особое место хроматиновым петлям и экструзии как механизму их образования. Дан подробный анализ барьерных функций белка CTCF и разобраны гипотезы о барьерных свойствах транскрипционных и репликационных комплексов, а также их роли в формировании границ петель у организмов в отсутствие белка CTCF.

Разобраны современные представления о данных Hi-C и методах работы с ними, уделено внимание методам нормализации, визуализации и оценке сходства данных Hi-C. Рассмотрены основные концепции анализа данных Hi-C, включая зависимость частоты контактов от расстояния и интуитивную. Описаны теоретические основы использования средних значений частот контактов на заданном расстоянии как «ожидаемых» частот контактов на том же расстоянии.

Обсуждена проблема выбора теоретического распределения для поиска значимых контактов, алгоритмы поиска хроматиновых петель и их математические основы.

Во **второй главе** приведено подробное описание использованных методов. Работа выполнена с использованием данных Hi-C и RNA-seq для временных стадий 0, 2, 5 и 8 часов после индукции голодания на культуре *D. discoideum* штамм AX4, полученных в лаборатории С.В. Разина Института биологии гена РАН, а также публично доступных данных ChIP-seq, ATAC seq<sup>19</sup> и аннотаций различных элементов генома *D. discoideum*, полученных из базы dictyBase. Для *H. dujardini* данные Hi-C были получены А. Черкасовым совместно с Институтом биологии развития РАН и Институтом биологии гена РАН. Данные Hi-C для анеуплоидий человека были получены в лаборатории С.В. Разина Института биологии гена РАН.

Для обработки результатов секвенирования экспериментов Hi-C был использован конвейер distiller-nf с параметрами по умолчанию. Полученные карты контактов *D. discoideum* и *H. dujardini* были разбиты на бины размером 100 п.н. и 1000 п.н. соответственно с помощью cooler. Разрешение для работы с петлями *D. discoideum* было установлено на 2 тыс. п.н. в связи с избыточным количеством белых полос и высокой разреженностью карт, однако, отдельные анализы, например, средняя карта контактов вокруг конвергентной пары, были проведены на разрешении 500 п.н. Фонтаны в картах *H. dujardini* были проанализированы на разрешении 5 тыс. п.н., а изучение изменений глобальной и локальной упаковки хроматина при анеуплоидиях было проведено на картах с разрешением 50 тыс. п.н.

Был проведен анализ средней карты межхромосомных взаимодействий с целью выявить, уложены ли хромосомы в ядрах *D. discoideum* и *H. dujardini* в структуры Рабля. Для этого карты межхромосомных взаимодействий были масштабированы до одного размера, нормализованы и усреднены. С помощью пакета cooltools был проведен анализ хроматиновых компартментов и подсчитана их сила.

При визуальной оценке на картах контактов хроматина *D. discoideum* были обнаружены точечные повышения частот контактов в сравнении с окружением, для их аннотации были использованы инструменты LASCA<sup>20</sup> и chromosight<sup>21</sup>. На картах контактов хроматина *H. dujardini* были обнаружены хроматиновые фонтаны, которые были аннотированы с помощью fontanka<sup>22</sup>.

<sup>19</sup>Role of epigenetics in unicellular to multicellular transition in Dictyostelium / S. Y. Wang [et al.] // Genome Biology. 2021. Vol. 22, no. 1. P. 134.

<sup>20</sup>LASCA: Loop and Significant Contact Annotation Pipeline / A. V. Luzhin [et al.] // Sci Rep. 2021. Vol. 11, issue 1, no. 1. P. 6361.

<sup>21</sup>Computer vision for pattern detection in chromosome contact maps / C. Matthey-Doret [et al.] // Nature Communications. 2020. Vol. 11, no. 1. P. 1—11.

<sup>22</sup>Extrusion Fountains Are Hallmarks of Chromosome Organization Emerging upon Zygotic Genome

Для изучения свойств найденных структур *D. discoideum* были использованы данные ATAC-seq и ChIP-seq из работы<sup>23</sup>. Данные обрабатывались с использованием конвейера nf-coqe для ATAC-seq и ChIP-seq. Далее с использованием этих данных были аннотированы энхансер-подобные элементы (Рисунок 1).

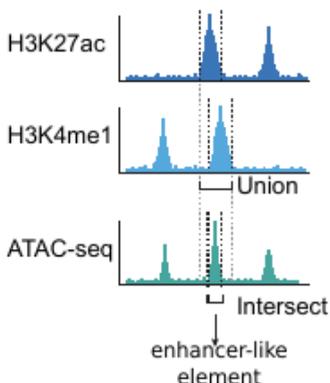


Рис. 1 — Схема аннотации энхансер-подобных элементов. Если полученные пики H3K3me1 и H3K27ac пересекались, использовалось их объединение. Затем полученные пики пересекались с пиками ATAC-seq.

Для оценки изменений локальной упаковки хроматина при анеуплоидиях по сравнению с нормальным кариотипом был введен показатель  $R$ . Этот показатель рассчитывается как отношение долей частот контактов на расстоянии не более 1 млн. п.н. к сумме всех внутрихромосомных контактов данного локуса для образцов с трисомиями и контроля.

**Третья глава** посвящена описанию полученных в ходе диссертационной работы результатов. Контроль качества обработанных данных секвенирования библиотек Hi-C был проведен с оценкой доли картирования, доли пар прочтений, где оба прочтения были закартированы, а также доли цис-контактов.

## Организация хроматина *Dictyostelium discoideum*

При помощи инструмента HiCRep установлена высокая корреляция как между биологическими репликами (около 98 %, Рисунок 2), так и между стадиями развития *D. discoideum*. На основании этой высокой корреляции и визуального сходства карт контактов, последующие анализы структуры



Рис. 2 — Матрица корреляций между биологическими репликами и часами развития. Значения коэффициента корреляции рассчитаны с помощью инструмента HiCRep.

хроматина *D. discoideum* были проведены на карте 0 часа развития, если не указано иное.

На уровне межхромосомных взаимодействий с использованием визуализации в HiGlass<sup>24</sup> обнаружено, что концы разных хромосом *D. discoideum* характеризуются повышенными относительно фона частотой контактов. Для анализа силы транс-взаимодействий (предполагаемо центромерных) и подтверждения структуры Рабля был проведен анализ средней карты межхромосомных взаимодействий на разных стадиях развития (Рисунок 3). Наблюдаемое повышение частот контактов между концами хромосом по сравнению с фоном на всех точках развития свидетельствует о сильном межцентромерном и межтеломерном взаимодействии разных хромосом. Результаты анализа также демонстрируют асимметрию частот межхромосомных контактов для 5'- и 3'-концов разных хромосом, на всех точках развития, что, вероятно, связано с наличием на 5'-конце центромера. Стоит отметить, что для акроцентрических хромосом *D. discoideum* 5'-конец хромосомы содержит и центромерный, и теломерный регионы, поэтому окончательно разделить теломерные и центромерные взаимодействия 5'-концов хромосом не представляется возможным. Мы заметили также, что контакты между 5'- и 3'-концами хромосом у *D. discoideum* редки в сравнении со средней частотой межхромосомных взаимодействий, что

свидетельствует о том, что хромосомы *D. discoideum* избегают контактов друг с другом разными концами.

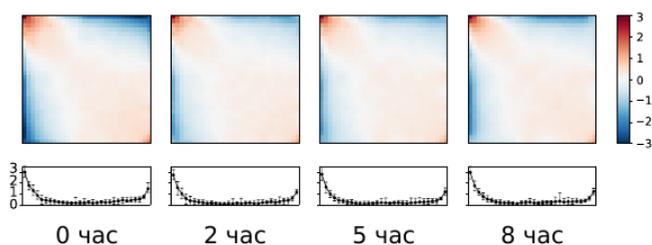


Рис. 3 — Структура межхромосомных контактов на разных стадиях развития *D. discoideum*. Верхняя панель: средняя карта нормализованных частот межхромосомных взаимодействий на четырех стадиях развития. Нижняя панель: частоты контактов на диагонали, взятые из полученных матриц межхромосомных взаимодействий. Интервалы, представленные на нижней панели отражают разброс значений нормализованных частот контактов от минимальных до максимальных.

Был проведен сравнительный анализ средних карт межхромосомных взаимодействий для *S. cerevisiae* и *D. discoideum* (Рисунок 4). Использование *S. cerevisiae* для сравнения обусловлено тем, что это один из немногих представителей низших эукариотов, для которых имеются карты хорошего разрешения и описана структура Рабля. Для сравнительного анализа межхромосомные взаимодействия двух организмов были приведены к одному масштабу, и была построена карта средних межхромосомных взаимодействий. Карты межхромосомных контактов для *S. cerevisiae* были «разрезаны» по центромерам (4 матрицы для межхромосомных взаимодействий каждой пары хромосом) и ориентированы так, чтобы в левом верхнем углу были межцентромерные взаимодействия, а в нижнем правом углу — межтеломерные. Результаты анализа показали, что межцентромерные и межтеломерные взаимодействия *D. discoideum* и *S. cerevisiae* одинаково обогащены по сравнению со средней частотой межхромосомных контактов.

При визуальном анализе карт Hi-C были выявлены обогащения над фоном

точечных дальних взаимодействий (как межхромосомных, так и внутрихромосомных) на 0 и 8 часах развития (соответственно 12 и 11 точечных взаимодействий). Было проведено изучение наличия некодирующих РНК в основании дальних контактов. Для этого была взята разметка

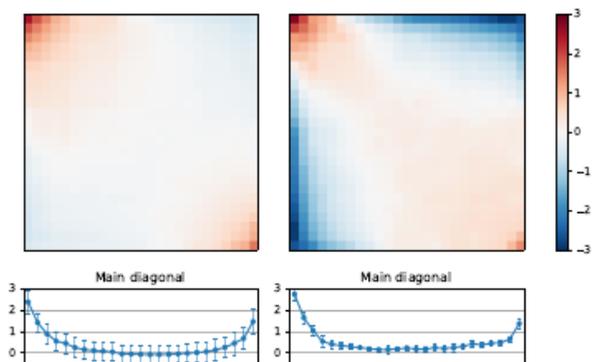


Рис. 4 — Сравнительный анализ структур Рабля в хроматине *D. discoideum* и *S. cerevisiae*. Слева: *D. discoideum*, справа: *S. cerevisiae*. Верхняя панель: средняя карта нормализованных частот межхромосомных взаимодействий. Нижняя панель: графики распределения частот контактов вдоль верхней строки карты межхромосомных взаимодействий с верхней панели.

длинных некодирующих РНК, аннотированных в работе<sup>25</sup>, и подсчитано сколько из них пересекаются с основаниями дальних контактов, значимость рассчитана пермутационным тестом. В результате проведенного анализа оказалось, что в основаниях дальних контактов находятся длинные некодирующие РНК общим количеством 22 и 17 соответственно, что статистически ( $p$ -значение  $< 0.001$ ) больше, чем в случайных областях. Выявлено так же, что изменения экспрессии между 0 и 8 часами развития, подсчитанные как разница значений Transcripts Per Million (TPM) генов, кодирующих длинные некодирующие РНК, которые лежат в основаниях дальних петель в сравнении случайно выбранными длинными некодирующими РНК, то статистически значимые отличия возникают на обеих стадиях, причем на стадии наличия дальних контактов экспрессия ожидаемо выше ( $p$ -значения  $< 0.0001$  и  $< 0.01$  соответственно) (Рисунок 5).

Анализ карт Hi-C на разрешении 2 тыс. п.н. показал наличие ярких точек обогащенных над фоном контактов (называемых в литературе петлями<sup>26</sup>. Визуальный анализ показал, что большинство точек стабильны

<sup>25</sup>The Long Noncoding RNA Transcriptome of Dictyostelium discoideum Development / R. D. Rosengarten [et al.] // G3: Genes, Genomes, Genetics. 2017. Vol. 7, no. 2. P. 387—398. <sup>26</sup>Comparison of computational methods for Hi-C data analysis / M. Forcato [et al.] // Nature Methods. 2017. July. Vol. 14, no. 7. P. 679—685; A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping / S. S. P. Rao [et al.] // Cell. 2014. Vol. 159, no. 7. P. 1665—1680.

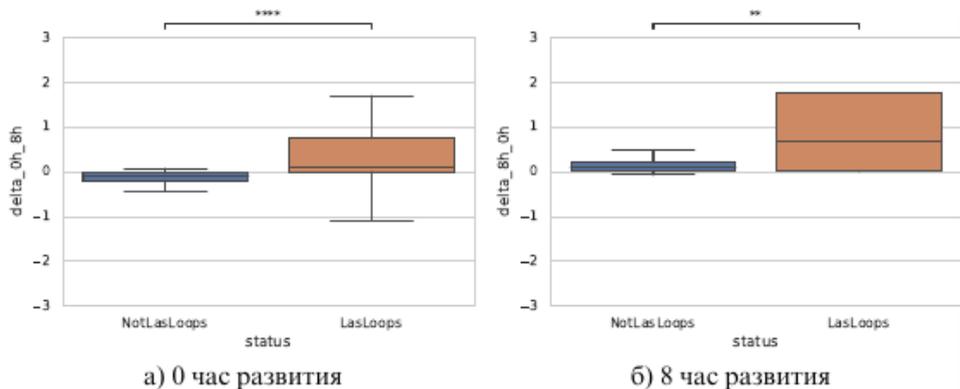


Рис. 5 — Изменения уровня экспрессии генов длинных некодирующих РНК, лежащих в основании дальних контактов, в сравнении со средним значением изменения экспрессии, рассчитанного пермутационным тестом.

между стадиями развития. Для изучения особенностей этих взаимодействий (в дальнейшем называемыми петлевыми) были использованы инструменты для аннотации петель LASCA<sup>27</sup> и chromosight<sup>28</sup>. Финальной разметкой были выбраны петли, размеченные с помощью обоих инструментов одновременно; такой подход позволил добиться оптимальной интенсивности средней карты хроматиновых контактов вокруг петель и снижения количества ложных находок.

Из аннотации петель следует, что преобладающее большинство из них (75,5%) действительно является последовательными, а также, что петли не являются иерархичными. Для подтверждения визуальных наблюдений был проведен анализ средней хроматиновой петли разных уровней иерархии. Для этого петлями первого порядка названы петли, полученные из разметки; петлями второго порядка — те, для которых 5'-основание взято из одной петли, а 3'-основание из соседней. Для петель третьего и четвертого порядков 3'-основание взято из петли, находящейся, соответственно, через одну или две петли от той петли, чье 5'-основание взято в анализ. Средняя карта контактов вокруг петель разных порядков показывает (Рисунок 6), что сила петли, полученная как усреднение по позициям первого порядка в окне 3×3 вокруг центра петли, в 1,63 раза выше, чем ожидаемая (0,705 в логарифмической шкале), в то время как для петель второго порядка сила петли составляет 1,17, что может свидетельствовать об отдельных петлях второго порядка. В то же

время интенсивность петли третьего порядка неотличима

<sup>27</sup>LASCA: loop and significant contact annotation pipeline / A. V. Luzhin [et al.] // Scientific Reports. 2021. Vol. 11, no. 1. P. 6361.

<sup>28</sup>Computer vision for pattern detection in chromosome contact maps / C. Matthey-Doret [et al.] // Nature Communications. 2020. Vol. 11, no. 1. P. 1—11.

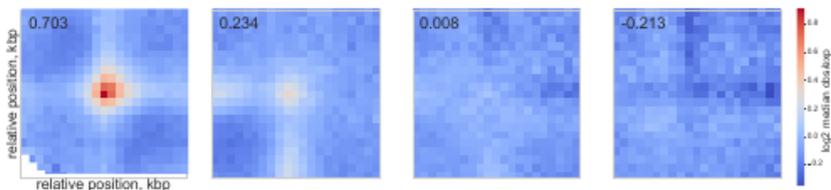


Рис. 6 — Изучение иерархии в петлях хроматина *D. discoideum*. Слева на право представлены график средней петли первого, второго, третьего и четвертого порядков.

от ожидаемой, а интенсивность петли четвертого порядка даже ниже ожидаемой. Совместно, полученные результаты свидетельствуют об отсутствии иерархии петель *D. discoideum*, что, возможно, связано с низкой процессивностью когезина или с очень низкой проницаемостью барьеров экстррузии.

Для изучения консервативности петель *D. discoideum* были использованы данные из работы<sup>29</sup>, в которых определены синтенные блоки между двумя видами рода *Dictyostelium*: *D. discoideum* и *D. purpureum*. Визуальный анализ показал, что границы синтенных блоков часто соответствуют границам петель (Рисунок 7а), а отдельные синтенные блоки покрывают сразу несколько петель. Пермутационный тест свидетельствует о более частой колокации границ синтенных блоков и границ петель, чем ожидается для структур такого количества и размеров (Рисунок 7б). Таким образом, проведенный анализ свидетельствует о том, что петли часто являются частью синтенных блоков, а границы петель и границы синтенных блоков статистически значимо колокализуются.

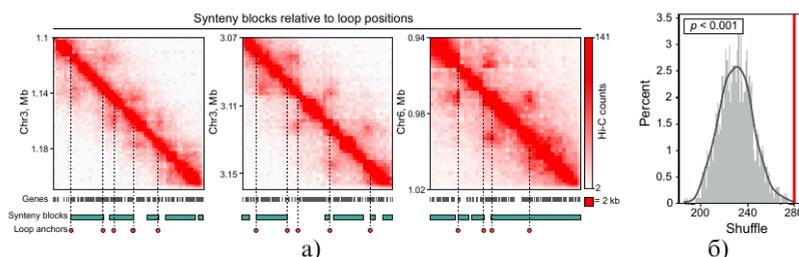


Рис. 7 — Колокализация синтенных блоков и петель *D. discoideum*. (а) Примеры колокации синтенных блоков и границ петель *D. discoideum*. (б) Численная оценка колокации границ синтенных блоков и границ петель *D.*

*discoideum*, выполненная с помощью пермутационного теста.

<sup>29</sup>Comparative Genomics of the Social Amoebae *Dictyostelium Discoideum* and *Dictyostelium Purpureum* / R. Sucgang [et al.] // *Genome Biology*. 2011. Vol. 12, no. 2. R20.

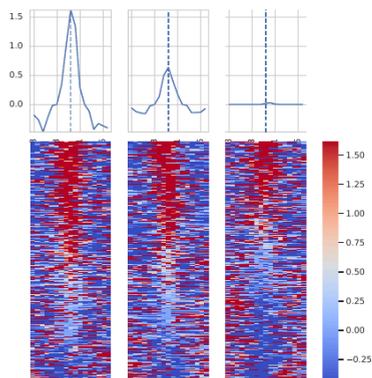


Рис. 8 — Профиль показателей конвергентности вокруг оснований последовательных и одиночных петель в сравнении с контролем. Верхняя панель: медианные значения показателей вокруг оснований петель. Нижняя панель: тепловые карты профиля сигнала вокруг каждого отдельного основания петель. Пунктиром обозначены основания петель.

Для того чтобы выяснить, обладают ли последовательные и одиночные петли разными свойствами, был проведен сравнительный анализ оснований последовательных и одиночных петель. Для оценки ориентации генов вокруг оснований петель использовался показатель конвергентности, который позволяет оценивать направления генов на цепях ДНК в заданном окне (20 тыс. п.н.) вокруг каждого бина генома, причем чем выше показатель, тем больше генов на плюс-цепи расположено левее от бина, для которого рассчитывается показатель, и больше генов на минус-цепи ДНК правее от бина, для которого рассчитывается показатель. Результаты анализа показали (Рисунок 8), что основания последовательных петель имеют более выраженный пик показателя конвергентности (медианные значения достигают 1,5), что приблизительно в 3 раза выше, чем соответствующий показатель для одиночных петель. Таким образом, проведенный анализ свидетельствует о роли конвергентно ориентированных генов в образовании границ петель, причем пик показателя конвергентности значительно более выражен для оснований последовательных петель, чем для оснований одиночных петель.

При визуальном анализе хроматина *D. discoideum* мы заметили, что,

помимо симметричных петель, LASCA<sup>30</sup> также аннотирует асимметричные петли (el-петли), вытянутые как вдоль 5'-основания, так и 3'-основания. Для

<sup>30</sup>LASCA: loop and significant contact annotation pipeline / A. V. Luzhin [et al.] // Scientific Reports. 2021. Vol. 11, no. 1. P. 6361.

их выделения из аннотации петель, был разработан простой алгоритм на основе отношения сигнала вдоль 5'- и 3'-оснований петель. С целью выяснения свойств вытянутых петель, мы провели анализ распределения значений экспрессии генов и показателя инсуляции вокруг петель. Для этого el-петли были ориентированы так, чтобы петля была вытянута вдоль 5'-основания. Полученный график для показателя инсуляции (Рисунок 9а) свидетельствует о сильной инсуляции 5'-основания и слабой инсуляции 3'-основания. Эти результаты соотносятся также с изменением профиля экспрессии (Рисунок 9б) вокруг петель, который демонстрирует снижение уровня экспрессии в 3'-основаниях el-петлей по сравнению с 5'-основанием el-петель, а также по сравнению с сигналом в обоих основаниях симметричных петель. Это свидетельствует о том, что сигнал петли вытягивается вдоль одного основания, если уровень транскрипции в другом ниже. Таким образом, предполагаемым барьером экстрюзии в хроматине *D. discoideum* является активная транскрипция, как это было показано для хроматина млекопитающих<sup>31</sup>.

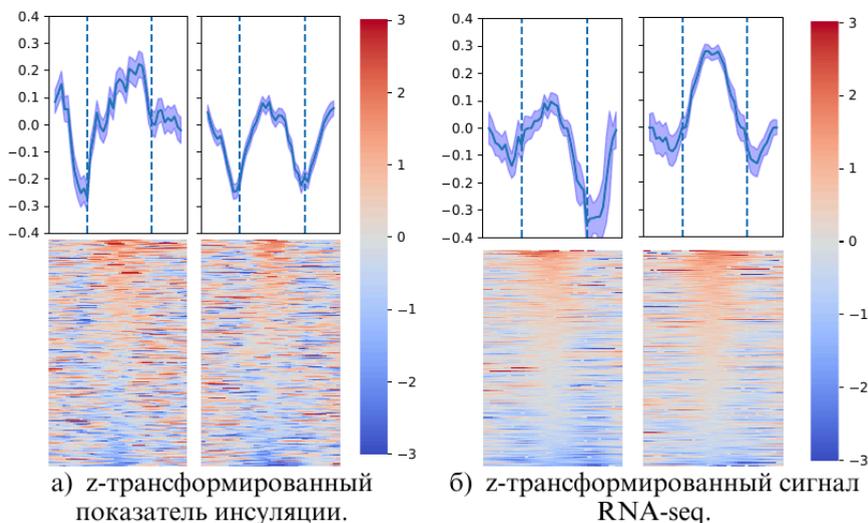


Рис. 9 — Графики значений сигнала RNA-seq (а) и показателя инсуляции (б) вокруг вытянутых (левая панель) и симметричных (правая панель) петель. Тепловая карта показывает z-трансформированные значения вокруг отдельных петель, верхние панели — медианные значения этих параметров. Пунктирами обозначены основания петель.

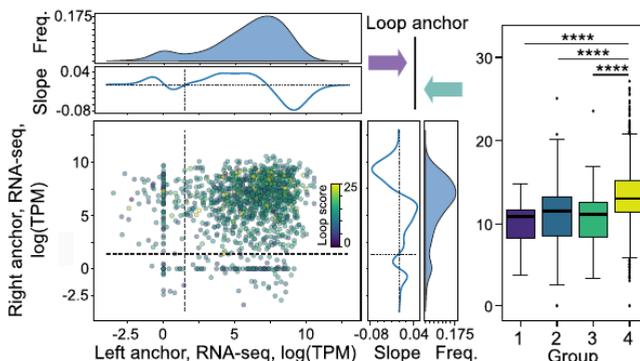


Рис. 10 — Уровень конвергентной транскрипции вокруг оснований петель и его связь с силой петель. Ось  $x$  — сумма конвергентной транскрипции вокруг 3'-оснований петель, ось  $y$  — сумма конвергентной транскрипции во круг 5'-оснований петель.

Полученные для вытянутых петель результаты свидетельствуют о влиянии уровня транскрипции на структуру петель. В то же время, мы показали повышение показателей конвергентности в основаниях петель. Для более подробного понимания общих закономерностей связи формы и силы петли и транскрипции, мы решили провести анализ уровней транскрипции конвергентно ориентированных генов вокруг оснований петель. Для этого был рассчитан показатель конвергентной транскрипции вокруг оснований как сумма уровня транскрипции генов, лежащих на плюс цепи, расположенных на расстоянии 2-4 тыс. п.н. со стороны 5'-конца основания петли, и уровня транскрипции генов, лежащих на минус цепи, расположенных на 2-4 тыс. п.н. со стороны 3'-конца основания петли. В результате каждой петле соответствуют два значения показателей конвергентной транскрипции вокруг 5'- и 3'-оснований петель, которые были отложены по  $y$ - и  $x$ -осям соответственно (Рисунок 10). Сила петли рассчитана как сумма нормализованной на ожидаемое частоты контактов в окне  $3 \times 3$  бина размером 2 тыс. п.н. каждый и показана на графике цветом соответствующей точки.

Значения экспрессии вокруг основания, соответствующие изменению знака первой производной, были использованы как пороги для разделения на группы. Для каждой группы был построен график распределения силы петли (правая панель Рисунка 10). Рассчитанные с помощью одностороннего теста Манна-Уитни  $p$ -значения для силы петли в разных группах демонстрируют высокую значимость отличий между группами (Рисунок 10).

Полученные нами результаты свидетельствуют, что именно конвергентная ориентация генов важна для формирования основания петель. Поэтому далее мы сосредоточили внимание на генах, образующих конвергентные пары. Конвергентными парами генов (КПГ) были названы пары генов, которые расположены на цепях в конвергентной ориентации, расстояние между концами этих генов не более 8 тысяч п.н., и такие гены являются ближайшими соседями друг друга. Далее в тексте КПГ, расположенные в геномном бине размером 2 тыс. п.н., идентифицированном как основание петли, и/или в его окрестности размером 4 тыс. п.н., называются КПГ, ассоциированными с основанием петли (оп-КПГ).

Далее, мы сосредоточили исследование на влиянии транскрипции генов, входящих в оп-КПГ, на контакты хроматина. Для этого, мы разделили гены, входящие в оп-КПГ, на высокоэкспрессируемые и низкоэкспрессируемые, используя пороговые значения 8 и 2 Transcripts Per Million (TPM) соответственно. Мы получили три группы пар: оба гена с низким уровнем транскрипции (нн-КПГ), оба гена с высоким уровнем транскрипции (вв-КПГ) и один ген с высоким уровнем транскрипции, а другой с низким (вн-КПГ). Для полученных трех категорий были построены средние карты хроматиновых взаимодействий, причем последний тип пар был ориентирован так, чтобы высокоэкспрессируемый ген находился с 5'-стороны пары (Рисунок 11а). Результаты анализа показали, что вв-КПГ образуют контактные полосы («stripe» на Рисунке 11а)<sup>32</sup> как с 5'-конца (upstream), так и с 3'-конца (downstream). Кроме того, слабоэкспрессируемый ген из вн-КПГ теряет контакты с участком выше пары (со стороны высокоэкспрессируемого гена; третья панель Рисунка 11а, черная стрелка). Мы также наблюдали асимметричный прирост контактов между участками, расположенными с 5'-конца и 3'-конца вн-КПГ, поскольку участки, расположенные с 5'-конца вн-КПГ, увеличивают частоту взаимодействия с регионами, непосредственно примыкающими к 3'-концу вн-КПГ (со стороны слабоэкспрессируемого гена; третья панель Рисунка 11а, пунктирная область). Наконец, контактная полоса с 3'-конца пары значительно ослаблена в вн-КПГ по сравнению с вв-КПГ, тогда как контактная полоса с 5'-конца пары, как и профиль инсуляции, в вн-КПГ сохраняются. Для оценки наблюдаемых изменений было произведено вычитание логарифмированных карт вн-КПГ и вв-КПГ (правая панель Рисунка 11а).

Проведенный анализ свидетельствует, что пара генов с низкой экспрессией слабо тормозит когезин (отсутствует контактная полоса), пара с высокой экспрессией является высокоэффективным барьером для экстразии (симметричная контактная полоса), а пара генов с высокой и с низкой экспрессией является, по-видимому, сильным барьером со стороны гена с высокой экспрессией и слабым барьером со стороны гена с низкой экспрессией, о чем свидетельствует асимметричность контактной полосы.

<sup>32</sup>Transcription Shapes 3D Chromatin Organization by Interacting with Loop Extrusion / E. J. Banigan [et al.] // Proceedings of the National Academy of Sciences. 2023. Vol. 120, no. 11. e2210480120; *Mirny, L. A.* Two Major Mechanisms of Chromosome Organization / L. A. Mirny, M. Imakaev, N. Abdennur // Current Opinion in Cell Biology. 2019. T. 58. C. 142—152. (Cell Nucleus); Formation of Chromosomal Domains by Loop Extrusion / G. Fudenberg [et al.] // Cell Rep. 2016. Vol. 15, no. 9. P. 2038—2049. pmid: 27210764.

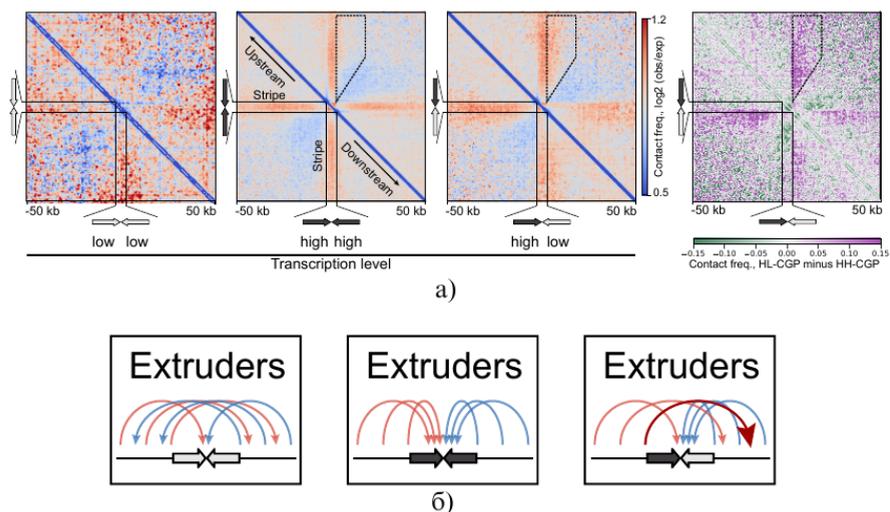


Рис. 11 — Влияние уровней транскрипции пары конвергентных генов на частоты контактов в окне 100 тыс. п.н. (а) Средняя карта хроматинового окружения конвергентных пар генов (КПГ). Панели слева направо: средняя карта хроматинового окружения нн-КПГ, вв-КПГ и вн-КПГ; результат вычитания средних карт контактов вокруг вн-КПГ и вв-КПГ. Три левые панели содержат нормализованные частоты контактов между регионами, правая — результат вычитания средних карт контактов вокруг вн-КПГ и вв-КПГ. (б) Схема экструзии для трех групп уровней транскрипции. Красными и синими стрелками обозначены контакты между областью со стороны со стороны 5'- и 3'-концов КПГ, соответственно, с областями, обозначенным концом стрелки. Бордовая стрелка крупнее остальных обозначает контакты регионов в выделенном четырехугольнике. Ген с высокой экспрессией обозначен стрелкой с черной заливкой, ген с низкой экспрессией обозначен стрелкой с белой заливкой.

Описанная модель изображена на Рисунке 11б. Полученные результаты свидетельствуют о важности не только конвергентной ориентации генов, но и их уровня транскрипции для образования основания петли.

Была проведена аннотация энхансер-подобных элементов (далее — эн

хансеры) путем пересечения пиков ATAC-seq, ChIP-seq H3K4me1 и H3K27ac (Рисунок 1). Каждый полученный энхансер был присвоен ближайшему гену с помощью функции `bedtools closest`, и для таких генов был построен график распределения уровней транскрипции; статистическая значимость разницы экспрессии между генами с энхансером и без него была подсчитана с помощью двустороннего теста Манн-Уитни (Рисунок 12а). Группа генов, не имеющая энхансеров, ожидаемо имеет меньшую экспрессию, чем гены, содержащие энхансер. Более того, подсчитана доля оснований петель, содержащих энхансеры, в сравнении с пермутационным тестом — около 30% оснований содержат энхансеры ( $p$ -значение  $< 0,001$ , пермутационный тест) (Рисунок 12б). Интересно также отметить, что гены, пересекающие энхансеры, имеют значительно более высокий в среднем уровень экспрессии на всех часах развития, а не только на тех часах развития, на которых они бы ли аннотированы.

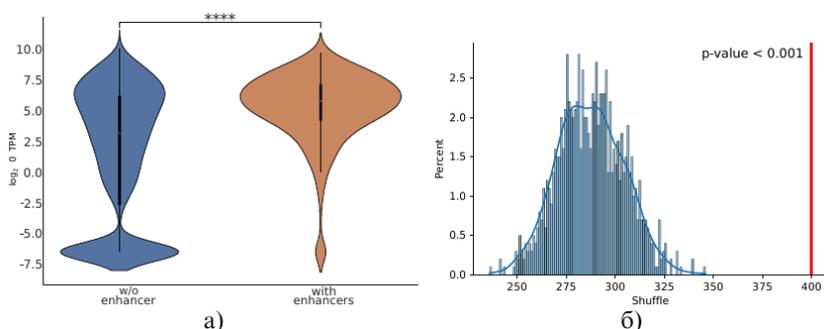


Рис. 12 — Анализ энхансеров генома: (а) Уровни транскрипции генов, пересекающих энхансеры, в сравнении с уровнем транскрипции генов, энхансеров не пересекающих. Ось  $y$  — TPM-нормализованные значения транскрипции генов, преобразованные в логарифмические значения. (б) Количество энхансеров, пересекающих основания петель, в сравнении со случайными последовательностями. Значимость рассчитана пермутационным тестом.

По совокупности полученных результатов можно предположить, что энхансеры в геноме *D. discoideum* представляют собой не специфичные для определенных стадий развития регуляторные элементы, как это известно для млекопитающих<sup>33</sup> и дрозофилы<sup>34</sup>, а скорее элементы, поддерживающие стабильную экспрессию на всех стадиях развития.

## Организация хроматина *Halisarca dujardini*

Прочтения Hi-C, используемые в данной работе, были получены А. Черкасовым в коллаборации с Институтом биологии развития им.

<sup>33</sup>Enhancer Evolution across 20 Mammalian Species / D. Villar [et al.] // Cell. 2015. Vol. 160, no. 3. P. 554—566. pmid: 25635462.

<sup>34</sup>Small, S. Transcriptional Enhancers in Drosophila / S. Small, D. N. Arnosti // Genetics. 2020. T. 216, № 1. С. 1—26. pmid: 32878914.

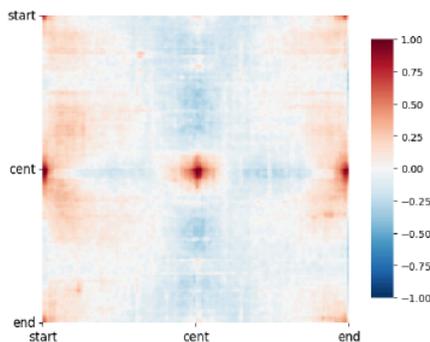


Рис. 13 — Средняя карта межхромосомных взаимодействий *H. dujardini*.

Для анализа взаимодействий между центромерами и теломерами хромосом *H. dujardini* была построена средняя карта межхромосомных взаимодействий. В отличие от акроцентрических хромосом *D. discoideum*, у губок центромера разделяет два примерно одинаковых плеча хромосомы, поэтому для губки возможно разделить теломерные и центромерные взаимодействия. Полученный график свидетельствует о наличии центромер-центромерных и теломер-теломерных взаимодействий между хромосомами (центромер-центромерные взаимодействия представлены центральной точкой на Рисунке 13, теломер-теломерные — взаимодействия в углах графика). Однако важно ещё отметить, что для хромосом *H. dujardini* также наблюдаются центромер-теломерные взаимодействия между разными хромосомами (повышенные частоты контактов, расположенные в середине около границ графика, Рисунок 13). Описанные характеристики свидетельствуют о том, что хромосомы *H. dujardini* образуют структуру Рабля с повышенными частотами контактов не только между теломерами и центромерами разных хромосом, но и между теломерами одних хромосом и центромерами других хромосом.

Для изучения компартментной организации хроматина *H. dujardini* были посчитаны собственные вектора карт Hi-C. Визуальный анализ (Рисунок 14а) показывает, что в хроматине *H. dujardini* присутствуют выраженная «шахматная доска», а первый собственный вектор хорошо отражает изменения в паттернах взаимодействия локусов друг с другом внутри хромосом. Была также построена седловая диаграмма для изучения силы предполагаемых компартментов (Рисунок 14б). Полученный график свидетельствует о наличии выраженных взаимодействий как между локусами активного компартмента, так и внутри неактивного компартмента. При этом частоты взаимодействий локусов, принадлежащих разным

компартаментам, оказываются ниже ожидаемых (правая верхняя и левая нижняя части Рисунка 14б). Проведенный анализ свидетельствует о выраженной компарментализации хроматина *H. dujardini*.

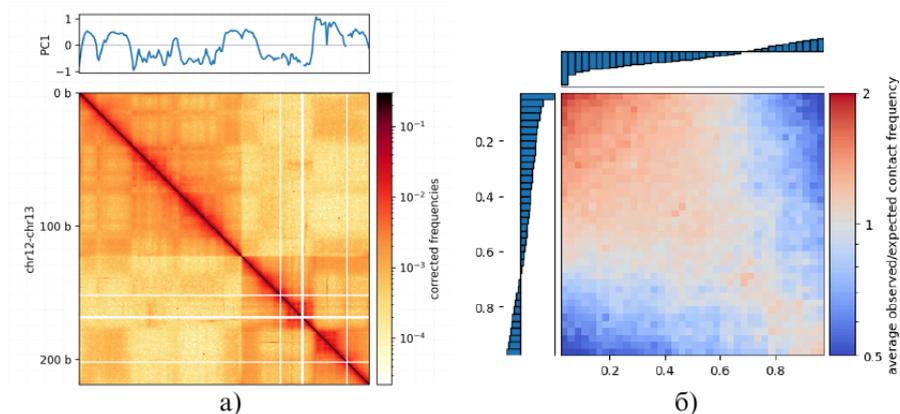


Рис. 14 — Компарменты в хроматине *H. dujardini*. (а) Пример смены знака значений первого собственного вектора для нескольких хромосом *H. dujardini*. (б) Седловая диаграмма для компарментов хроматина *H. dujardini*.

При визуальной оценке карты контактов губки *H. dujardini* с использованием хроматинового браузера HiGlass<sup>35</sup> на разрешении 3-5 тысяч п.н. мы обнаружили яркие вытянутые структуры, похожие на хроматиновые шпильки, или фонтаны. Для аннотации данных объектов был использован инструмент fontanka<sup>36</sup>, специально разработанный для поиска подобных структур.

С помощью этого инструмента мы смогли получить 986 фонтанов со средней силой 1,12 (Рисунок 15). Для губки *Petrosia ficiformis* данные Hi C были взяты из проекта Tree of Life (SRA: ERR10297828). Аналогичная процедура, проведенная с картой контактов губки *P. ficiformis*, позволила получить 1309 фонтанов со средней силой 1,1. Проведенный анализ позволяет говорить о большой распространенности фонтанов в хроматине представителей Porifera.

Для морских губок наличие ионов железа является обязательным условием для роста, поэтому отдельно интересно изучить, какие именно гены участвуют в данном процессе и какую роль в этом играет организация хроматина. В геноме губки *H. dujardini* обнаружено три последовательности, продукты которых имеют статистически значимое сходство с аминокислотной последовательностью ферритина человека HuHF, причем две последовательности, *HdF1a* и *HdF1b*, отличаются друг от друга на несколько

<sup>35</sup>HiGlass: Web-Based Visual Exploration and Analysis of Genome Interaction Maps / P. Kerpe jiev [et al.] // Genome Biology. 2018. Vol. 19, no. 1. P. 125.

<sup>36</sup>Extrusion Fountains Are Hallmarks of Chromosome Organization Emerging upon Zygotic Genome Activation / A. Galitsyna [et al.]. 2023. preprint.

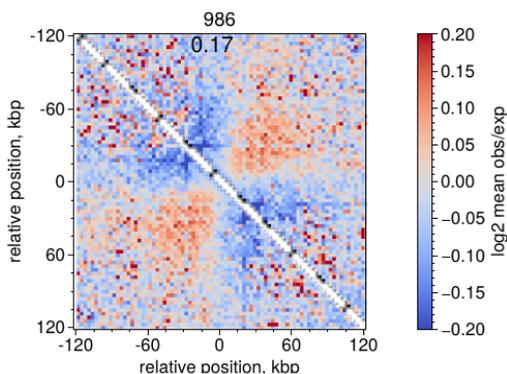


Рис. 15 — Средняя карта контактов вокруг хроматиновых фонтанов, аннотированных с помощью инструмента fontanka.

аминокислот, а третья, *HdF2*, значительно отличается как от NuHF, так и от *HdF1a* и *HdF1b*.

Далее мы провели сравнительный анализ аминокислотных последовательностей гомологов ферритинов из различных представителей типа Porifera и ферритина человека NuHF, оценив, в частности, количества аминокислотных замен в консервативных позициях активных центров. Результаты анализа демонстрируют, что для большинства последовательностей число аминокислотных замен в сравнении с NuHF не превышает трех, за исключением упомянутого выше белка *H. dujardini* HdF2 (шесть аминокислотных замен) и ферритина губки *Ephydatia muelleri* EmF2 (шесть аминокислотных замен). По совокупности признаков было решено называть ферритины HdF2 и EmF2 атипичными ферритинами.

Кроме того, был проведен анализ хроматинового окружения ортологов ферритина для *H. dujardini* и других представителей Metazoa. Результаты анализа показывают, что основание одного из хроматиновых фонтанов у *H. dujardini* расположено в непосредственной близости от локуса ферритина. На примере локуса ферритина в геноме *Danio rerio* мы проследили изменения хроматина эмбрионов до 20 часов после оплодотворения, где фонтаны ранних стадий превращаются в ТАДы и петли. Анализ карты контактов хроматина *Mus musculus* на разрешении 1 тыс. п.н. вокруг локуса ферритина показал наличие множества иерархических петель по геному, в том числе и вблизи локуса гена ферритина.

## Организация хроматина при анеуплоидиях

Для изучения возможной взаимосвязи между экспрессией генов и организацией хроматина была выбрана модель анеуплоидий, в частности, трисомий по хромосомам 13, 16 и 18. С. Ульяновым из Института биологии

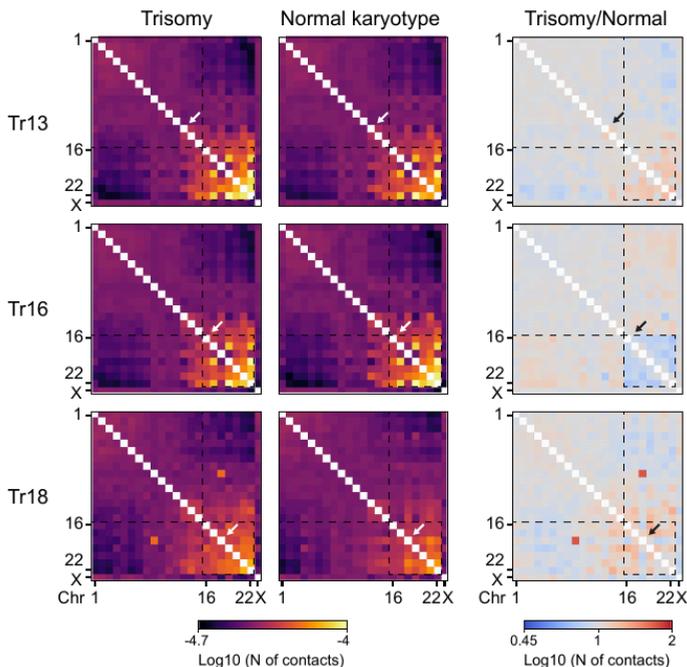


Рис. 16 — Карт межхромосомных контактов для образцов с трисомиями по 13, 16, 18 хромосомам человека *H. sapiens*.

гена РАН были выполнены эксперименты Hi-C на клетках хориона, несущих дополнительную хромосому 13 (синдром Патау, Tr13) или хромосому 16 (Tr16, наиболее частая трисомия при анализе материала самопроизвольных аборт), а также на культивируемых первичных фибробластах кожи с третьей копией хромосомы 18 (синдром Эдвардса, Tr18).

Анализ парных взаимодействий между хромосомами в клетках нормального кариотипа показал, что хромосомы небольшого размера образуют контактный кластер (т.е. имеют повышенные частоты контактов взаимодействия друг с другом) и имеют низкую частоту взаимодействия с крупными хромосомами, которые образуют менее выраженный, но заметный кластер (Рисунок 16). Поэтому мы определили хромосомы 16-22 как «малые», а хромосомы 1-15, X — как «большие».

Паттерн межхромосомных контактов в целом схож для образцов с различными дополнительными хромосомами: Tr13, Tr16 и Tr18 (сравни те левую и среднюю панели на Рисунке 16). Однако мы обнаружили, что наличие дополнительных хромосом вследствие трисомии вызывает умеренные, но статистически значимые и противоположные изменения в контактах между малыми и большими хромосомами (Рисунок 17). Для Tr13 и Tr18 мы

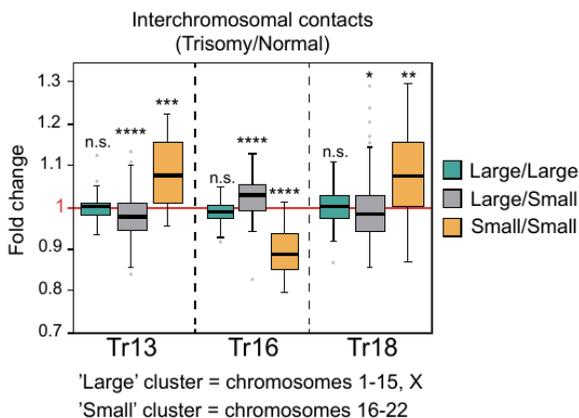


Рис. 17 — Различия частот межхромосомных контактов для больших и малых хромосом с трисомиями по хромосомам 13, 16, 18. ns  $p$ -значение  $> 0,05$ ; \*  $p$ -значение  $< 0,05$ ; \*\*  $p$ -значение  $< 0,01$ ; \*\*\*  $p$ -значение  $< 0,001$ ; \*\*\*\*  $p$  значение  $< 0,0001$  по тесту Манна-Уитни.

наблюдали значительное различие между образцами с трисомией и контролем в группах больших и малых хромосом, сопровождающееся увеличением частоты контактов между малыми хромосомами, тогда как для образцов Tr16 характерно частичное перемешивание кластеров больших и малых хромосом, а также выраженное снижение взаимодействий между малыми хромосомами.

Визуальный сравнительный анализ карт хроматиновых контактов показал значительные изменения локальной упаковки хроматина в образцах с трисомиями по сравнению с контролем. Для численной оценки геномных масштабов изменения локальной упаковки в образцах с трисомией мы подсчитали долю суммарной длины регионов, изменивших упаковку к общей длине хромосом, отдельно для больших и малых хромосом. Результаты численной оценки показывают, что для образцов Tr18 и Tr13, но не образца Tr16, доли длины хромосомы с измененной упаковкой значительно больше у больших хромосом, чем для малых (Рисунок 18а).

В образцах Tr13 и Tr18 мы обнаружили статистически значимую корреляцию между покрытием хромосом ламина-ассоциированными доменами и долей компактизованных участков на больших, но не на малых хромосомах

(Рисунок 18б), что также свидетельствует о существенных различиях между этими двумя группами.

В **четвертой главе** приведено подробное обсуждение полученных в ходе диссертационной работы результатов и сравнение их с современными литературными данными.

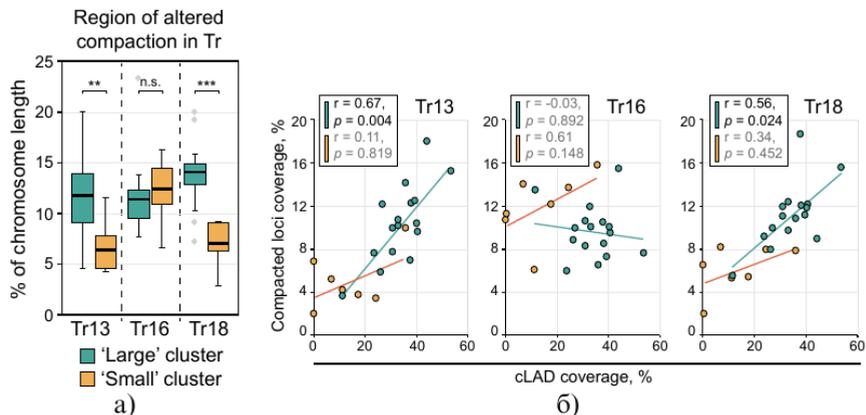


Рис. 18 — Сравнительная характеристика регионов хроматина с измененной упаковкой. (а) Сравнение доли длины хромосом с измененной упаковкой для образцов с анеуплоидиями для больших и малых хромосом. ns  $p$  значение  $> 0,05$ ; \*  $p$ -значение  $< 0,05$ ; \*\*  $p$ -значение  $< 0,01$ ; \*\*\*  $p$ -значение  $< 0,001$  по тесту Манна-Уитни. (б) Взаимоотношения доли длины хромосом, для которых изменилась компактизация хроматина, и доли покрытия ЛА Дами.  $r$  — коэффициент корреляции Спирмена.

Структура Рабля в литературе описаны, например, в делящихся и почкующихся дрожжах, дрозофиле и растениях. В основе подобных взаимодействий, по-видимому, лежит дефектный или отсутствующий конденсин II<sup>37</sup>. В настоящей работе было обнаружено, что хромосомы *D. discoideum* не склонны к образованию явных хромосомных территорий, а средние карты межхромосомных контактов как для *D. discoideum*, так и для *H. dujardini* показывают обогащение контактов в сравнении со средней частотой межхромосомных взаимодействий между центромерами и теломерами разными хромосомами.

Из литературы известен феномен дальних взаимодействий локусов *piwi*РНК у *D. melanogaster*, также в геноме *A. thaliana* были описаны структуры KNOT, которые представляют собой дальние взаимодействия локусов, обогащенные генами малых РНК<sup>38</sup>. В нашей работе были обнаружены дальние цис- и транс-контакты, специфичные для 0 и 8 часов развития

<sup>37</sup>3D Genomics across the Tree of Life Reveals Condensin II as a Determinant of Architecture Type / C. Hoencamp [et al.] // Science. 2021. Vol. 372, no. 6545. P. 984—989. <sup>38</sup>Grob, S. Hi-C Analysis in Arabidopsis Identifies the KNOT, a Structure with Similarities to the flamenco Locus of Drosophila / S. Grob, M. W. Schmid, U. Grossniklaus // Molecular Cell. 2014. T. 55, № 5. С. 678—693; Grob, S. Invasive DNA elements modify the nuclear architecture of their insertion site by KNOT-linked silencing in Arabidopsis thaliana / S. Grob, U. Grossniklaus // Genome Biology. 2019. T. 20. С. 120.

## 28

*D. discoideum*. Нами было показано, что в основаниях таких контактов лежат гены длинных некодирующих РНК. Было также показано, что гены длинных некодирующих РНК в основаниях дальних контактов меняют свой уровень экспрессии между часами развития, однако, вследствие недостаточно точной аннотации генов длинных некодирующих РНК, выявить их функции в рамках данной работы не удалось.

В литературе описаны компартменты у растений<sup>39</sup>, *Drosophila melanogaster*<sup>40</sup>, *Danio rerio*<sup>41</sup>, а также различных млекопитающих<sup>42</sup>. На стоящая работа позволяет добавить губку *Halisarca dujardini* к этому ряду организмов, поскольку в её хроматине присутствует выраженный паттерн «шахматная доска», а численная оценка подтверждает визуальные наблюдения. Также, по-видимому, по результатам настоящей работы, к таким организмам стоит отнести *D. discoideum*, для которого результаты численного анализа свидетельствуют о наличии слабой компартментализации генома.

Кроме того, описаны эффекты нарушения компартментализации как часть патогенеза различных заболеваний<sup>43</sup>. В нашей работе на примере трисомии человека мы показали, что наличие дополнительной копии хромосомы нарушает глобальную укладку хроматина, что проявляется в виде компактизации одних регионов и декомпактизации других.

Хромосомы в ядрах эукариот прикрепляются к ламине и ядрышку, а участки их крепления называются соответственно ламина-ассоциированные домены (ЛАДы) и ядрышко-ассоциированные домены<sup>44</sup>. В работе<sup>45</sup> было получено, что лишь 30 % ЛАДов ассоциированы с ламинной в каждой от дельно клетке. Было также показано, что именно расположение на ядерной

<sup>39</sup>Genome-Wide Hi-C Analyses in Wild-Type and Mutants Reveal High-Resolution Chromatin Interactions in Arabidopsis / S. Feng [et al.] // Molecular Cell. 2014. Vol. 55, no. 5. P. 694—707; Grob, S. Hi-C Analysis in Arabidopsis Identifies the KNOT, a Structure with Similarities to the Flamenco Locus of Drosophila / S. Grob, M. W. Schmid, U. Grossniklaus // Molecular Cell. 2014. Vol. 55, no. 5. P. 678—693. pmid: 25132176; 3D Chromatin Architecture of Large Plant Genomes Determined by Local A/B Compartments / P. Dong [et al.] // Mol Plant. 2017. Vol. 10, no. 12. P. 1497—1509. pmid: 29175436; Prominent Topologically Associated Domains Differentiate Global Chromatin Packing in Rice from Arabidopsis / C. Liu [et al.] // Nat Plants. 2017. Sept. Vol. 3, no. 9. P. 742—748.

<sup>40</sup>Evolutionarily Conserved Principles Predict 3D Chromatin Organization / M. J. Rowley [et al.] // Mol Cell. 2017. Vol. 67, no. 5. 837—852.e7. pmid: 28826674.

<sup>41</sup>Systemic Loss and Gain of Chromatin Architecture throughout Zebrafish Development / L. J. Kaaij [et al.] // Cell Rep. 2018. Vol. 24, no. 1. 1—10.e4. pmid: 29972771. <sup>42</sup>Principles of 3D Chromosome

Folding and Evolutionary Genome Reshuffling in Mammals / L. Alvarez-Gonzalez [et al.] // Cell Rep. 2022. Vol. 41, no. 12. P. 111839. pmid: 36543130. <sup>43</sup>Chakraborty, A. The Role of 3D Genome Organization in Disease: From Compartments to Single Nucleotides / A. Chakraborty, F. Ay // Semin Cell Dev Biol. 2019. Vol. 90. P. 104—113. pmid: 30017907.

<sup>44</sup>Razin, S. V. Genome-Directed Cell Nucleus Assembly / S. V. Razin, S. V. Ulianov // Biology. 2022. May. Vol. 11, issue 5, no. 5. P. 708.

<sup>45</sup>Single-Cell Dynamics of Genome-Nuclear Lamina Interactions / J. Kind [et al.] // Cell. 2013. Vol. 153, no. 1. P. 178—192. pmid: 23523135.

периферии вблизи ламины позволяет поддерживать репрессивное состояние генов<sup>46</sup>. Таким образом, хромосомы, содержащие значительную долю ЛАДов могут конкурировать за локализацию на ядерной периферии, а дополнительная копия хромосомы при трисомии выступает еще одним конкурентом в этом взаимодействии.

Ведущей теорией формирования петель и ТАДов в настоящее время является экструзия петель<sup>47</sup>. В настоящей работе было показано, что такие взаимодействия (названные хроматиновыми петлями) представляют собой высококонсервативные между часами развития последовательно расположенные структуры. Кроме того, нами было показано, что асимметрия вытянутых хроматиновых петель может быть обусловлена различиями экспрессии генов вокруг оснований, аналогично асимметричным фонтанам в хроматине *Caenorhabditis elegans*<sup>48</sup>.

Одним из центральных результатов текущей работы является предположение о роли конвергентно ориентированных экспрессируемых генов в формировании и поддержании границы петель у *D. discoideum*. Для *D. discoideum* это наблюдение сделано впервые, однако, в литературе есть примеры петель и доменов, границы которых, по-видимому, поддерживаются теми же принципами, например, ТАДы в хроматине представителей динофлагеллят *Breviolum minutum*<sup>49</sup> или петли в хроматине почкующихся дрожжей *Saccharomyces cerevisiae*<sup>50</sup>.

В **заключении** приведены основные результаты работы, которые заключаются в следующем:

1. Хромосомы *D. discoideum* и *H. dujardini* образуют структуру Рэбля с обогащенными межтеломерными и межцентромерными контактам.
2. Хроматин *H. dujardini* компартментализован, и в нем присутствуют хроматиновые фонтаны.
3. Хроматин *D. discoideum* организован в преимущественно последовательные, неиерархические петли, внутренняя часть которых слабо изолирована от соседних локусов.
4. Основания петель представляют собой сайты конвергентной ориентации генов.

<sup>46</sup>Promoter-Intrinsic and Local Chromatin Features Determine Gene Repression in LADs / C. Leemans [et al.] // Cell. 2019. Vol. 177, no. 4. 852—864.e14.

<sup>47</sup>Emerging Evidence of Chromosome Folding by Loop Extrusion / G. Fudenberg [et al.] // Cold Spring Harbor Symposia on Quantitative Biology. 2017. Vol. 82. P. 45—55.

<sup>48</sup>Cohesin Forms Fountains at Active Enhancers in *C. Elegans* / B. Isiaka [et al.]. 2023. preprint.

<sup>49</sup>Transcription-dependent domain-scale three-dimensional genome organization in the dinoflagellate *Breviolum minutum* / G. K. Marinov [et al.] // Nature Genetics. 2021. May. Vol. 53, no. 5. P. 613—617.

<sup>50</sup>Cohesin relocation from sites of chromosomal loading to places of convergent transcription / A. Lengronne [et al.] // Nature. 2004. July. Vol. 430, no. 6999. P. 573—578.

5. Сила оснований петель, отражающая эффективность барьерной функции для экструзии, зависит от силы окружающей основания конвергентной транскрипции.
6. Конвергентные пары генов, расположенные в основания петель, образуют контактные домены, структура которых зависит от уровня транскрипции обоих генов в паре.
7. В геномах губок присутствуют гены типичных ферритинов (один или два паралога) и гены атипичного ферритина с аминокислотными заменами в консервативных доменах; основание хроматинового фонтана находится в непосредственной близости от генов консервативных ферритинов.
8. Трисомии человека по хромосомам 13, 16 и 18 приводят к изменению укладки хромосом как на уровне межхромосомных контактов, так и на уровне локальных взаимодействий.
9. При трисомии по трем изученным хромосомам значительно меняются взаимодействия внутри малых хромосом (хромосомы 16-22), а также взаимодействие малых и больших хромосом друг с другом по сравнению с нормальным кариотипом.
10. Локальные хроматиновые взаимодействия (на расстоянии не более 1 млн. п.н.) значительно меняются внутри малых хромосом по сравнению с контролем при трисомии по хромосомам 13 и 18, но не 16.

## Публикации автора по теме диссертации

A1. Convergent gene pairs restrict chromatin looping in *Dictyostelium discoideum*, acting as directional barriers for extrusion / I. Zhegalova, S. Ulianov, A. Galitsyna, I. Pletenev, O. Tsoy, A. Luzhin, P. Vasiluev, E. Bulavko, D. Ivankov, M. Logacheva, A. Gavrilov, E. Khrameeva, M. Gelfand, S. Razin // bioRxiv. – 2024. – doi: <https://doi.org/10.1101/2024.06.12.598618>

A2. Trisomies reorganize human 3D genome / I. Zhegalova, P. Vasiluev, I. Flyamer, A. Shtompel, E. Glazyrina, N. Shilova, M. Minzhenkova, Z. Markova, N. Petrova, E. Dashinimaev, S. Razin, S. Ulianov // International Journal of Molecular Sciences. – 2023. – Vol. 24, issue 22, no. 22. – P. 16044. – (Scopus, WoS).

A3. Domain model of eukaryotic genome organization: from DNA loops fixed on the nuclear matrix to TADs / S. Razin, I. Zhegalova, O. Kantidze // Biochemistry (Moscow). – 2022. – Vol. 87, no. 7. – P. 667–680. – (Scopus, WoS).

A4. Conservative and atypical ferritins of sponges / K. Adameyko, A.

Burakov, A. Finoshin, K. Mikhailov, O. Kravchuk, O. Kozlova, <...> I. Zhegalova, A. Georgiev, V. Mikhailov, N. Gogoleva, G. Gazizova, E. Shagimardanova, O. Gusev, Y. Lyupina // International Journal of Molecular Sciences. – 2021. – Vol. 22, no. 16. – P. 8635. – (Scopus, WoS).

A5. Consecutive chromatin loops in *Dictyostelium discoideum* / I. Zhegalova, S. Ulianov, A. Galitsyna, E. Khrameeva // Bioinformatics of Genome 31 Regulation and Structure/Systems Biology (BGRS/SB-2022). – Novosibirsk, Russia, 2022.

A6. Chromatin loops are involved in spatial organization of replication in budding yeast / K. Ulyanov, I. Zhegalova, E. Khrameeva, M. Gelfand // Bioinformatics of Genome Regulation and Structure/Systems Biology (BGRS/SB-2022). – Novosibirsk, Russia, 2022.

A7. Features of chromatin structure & gene expression during *D. discoideum* development / I. Zhegalova, S. Ulianov, A. Galitsyna, E. Khrameeva // 10th International Moscow Conference on Computational Molecular Biology (MCCMB-2021). – Moscow, Russia, 2021.

A8. Elongated chromatin dots in *D. discoideum* / I. Zhegalova, S. Ulianov, A. Galitsyna, E. Khrameeva, M. Gelfand // Информационные технологии и системы (ИТИС-2021). – онлайн, 2021.

A9. G-quadruplexes and loops in chromatin of *Dictyostelium discoideum* / I. Zhegalova, O. Tsoy, A. Galitsyna, S. Ulianov, E. Khrameeva // Информационные технологии и системы (ИТИС-2019). – Пермь, Россия, 2019.

*Жегалова Ирина Владимировна*

Анализ трехмерной структуры хроматина эукариот

Автореф. дис. на соискание ученой степени канд. биол. наук

Подписано в печать . . . . Заказ №

Формат 60×90/16. Усл. печ. л. 1. Тираж 100 экз.

Типография