

Отчёт о проделанной работе с использованием оборудования ИВЦ НГУ

1. Аннотация

С помощью разработанного ПО C-InterSecure было проведено сравнение архитектуры хроматина разных таксономических групп на уровне отдельных пространственных контактов. Исследование показало, что консервативность контактов наблюдается как между разными представителями позвоночных (*Homo sapiens*, *Mus musculus* и *Gallus gallus*), так и между комарами рода *Anopheles*. Во всех случаях, большая консервативность наблюдается для наиболее сильных и наиболее слабых контактов, что соответствует представлениям о механизмах формирования пространственной организации хроматина. В тоже время, не подтвердилось предположение о консервативности измеряемых в ходе эксперимента частот контактов. Показано, что эволюционной консервативностью обладают относительные величины контактов, учитывающие расстояние между целевыми локусами и распределение частот контактов для данного расстояния.

2. Тема работы

Межвидовое сравнение пространственной организации хроматина

3. Состав коллектива

Нуридинов Мирослав Абдурахимович, аспирант, м.н.с., Институт Цитологии и Генетики СО РАН

Фишман Вениамин Сергеевич, канд. биол. наук, в.н.с. Институт Цитологии и Генетики СО РАН

4. Информация о грантах

РФФИ, №17-00-00147, Реконструкция хромосомной организации геномов млекопитающих и новый взгляд на хромосомную эволюцию, 2017-2020, руководитель — Ларкин Денис Михайлович.

РФФИ, №19-04-00840, Исследование роли конденсинов в поддержании трехмерной архитектуры интерфазного ядра плюрипотентных клеток, 2019-2022, руководитель — Юнусова Анастасия Маратовна.

РНФ, №19-74-20075, Топологически-ассоциированные домены хроматина и A/B-компарменты в геноме курицы: идентификация и визуализация с помощью технологии HiC и микроскопии сверхвысокого разрешения, 2019-2022, руководитель — Красикова Алла Валерьевна

5. Научное содержание работы

5.1. Постановка задачи

Разработка методов сравнения пространственной организации хроматина между разными видами. Проверка эффективности разработанных методов при сравнении архитектуры хроматина на разных таксономических групп (Vertebrata и Diptera).

5.2. Современное состояние проблемы

В настоящее время, благодаря развитию экспериментальных методик семейства захвата конформации хромосом показано, что хроматин в пространстве клеточного ядра уложен не произвольно, а формируя сложные, динамические, структуры [1]. Показано, что архитектура

хроматина демонстрирует эволюционную консервативность [2,3] для млекопитающих. Консервативным оказывается и организация хроматина и между разными типами клеток [2,4], при этом наблюдаемые различия соответствуют разнице в профиле экспрессии [5]. Однако остаются неизвестным, консервативна ли архитектура хроматина в целом для позвоночных и наблюдается ли такого рода консервативность в других таксономических группах.

Необходимо отметить, что основным способом сравнения архитектуры хроматина между видами является выделение с помощью разных алгоритмов тех или иных структур, и сравнения их положения в геноме в соответствии с картами синтении [2,3]. Данный метод обладает рядом существенных недостатков. Так он очень чувствителен к различиям в применяемых алгоритмах. Различия в описании архитектуры хроматин возникающие из-за использования разных алгоритмов оказываются сопоставимыми с различиями между разными типами клеток и разными видами. Кроме этого он не позволяет оценивать сходство или различия пространственной организации хроматина создаваемое отдельными контактами.

Это требует разработки методов, позволяющих непосредственно сравнивать архитектуру хроматина между разными видами.

1. Lieberman-Aiden E. et al. Comprehensive mapping of long range interactions reveals folding principles of the human genome. // *Science*. – 2009. – N. 326. – P. 289-293.

2. Dixon J. R. et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. // *Nature*. – 2012. – N. 485. – P. 376–380.

3. Vietri Rudan M. et al. Comparative Hi-C reveals that CTCF underlies evolution of chromosomal domain architecture // *Cell Rep*. – 2015. – N. 10. – P. 1297-1309.

4. Battulin N. et al. Comparison of the three-dimensional organization of sperm and fibroblast genomes using the Hi-C approach. // *Genome Biology*. – 2015. – N. 16. – S. 77.

5. Fraser J. et al. Hierarchical folding and reorganization of chromosomes are linked to transcriptional changes in cellular differentiation // *Molecular Systems Biology*. – 2015. – N. 11. – S 852.

5.3. Подробное описание работы, включая используемые алгоритмы

Для сравнения пространственной организации хроматина на уровне отдельных контактов нами было разработано программное обеспечение, названное C-InterSecture (**Computational tool for *InterSpecies* analysis of genome architecture**).

В основе данного ПО лежит алгоритм перекартирования индивидуальных контактов (или иной величины, вычисленной на их основе) между исследуемыми видами. В соответствии с этим степень консервативности архитектуры хроматина определяется сравнением величин перекартированных и наблюдаемых контактов, на основе чего делается вывод о сходстве или различии архитектуры хроматина данного локуса у сравниваемых видов.

Основная сложность проведения такого рода перекартирования является биновый характер данных Hi-C: для получения достоверных результатов необходимо объединять все контакты на протяжённом участке генома — бине, чья длина для большей части экспериментов составляет тысячи и десятки тысяч пар оснований.

Чтобы учесть эти особенности, синтенные блоки разбиваются на короткие гомологичные, у сравниваемых видов, фрагменты (~200 п.н.), которые не пересекают границы бинов и не перекрываются друг с другом. Предполагается, что каждый такой фрагмент вносит свой вклад в наблюдаемую частоту контактов. При этом, если архитектура хроматина консервативна, то и гомологичные фрагменты одинаковым образом формируют контакты.

Ещё одной сложностью было отсутствие в литературных источниках такого понятия как «консервативность контактов» в контексте межвидового сравнения. Это проблема была решена за счёт возможности использования в C-InterSecture разных моделей консервативности. Что позволяет, в том числе, и сравнить их предсказательную силу.

5.4. Полученные результаты

С помощью C-InterSecture было проведено сравнение пространственной организации хроматина *Homo sapiens*, *Mus musculus* и *Gallus gallus*. Результаты сравнения показали высокую консервативность архитектуры хроматина, прослеживаемую на уровне отдельных контактов. Так коэффициент корреляции при сравнении *Homo sapiens* и *Mus musculus* достигает 0.56 и 0.38 для сравнения *Homo sapiens* и *Gallus gallus*.

Сравнение показало, что в большей степени консервативность выражена для наиболее обогащённых и наиболее инсулированных контактов, разделённых расстоянием, не превышающим 1-2 млн п.о. В этом случае совпадения контактов наблюдается в 3-4 раза чаще, чем для случайных событий. Данный факт согласуется с представлениями о механизмах формирования архитектуры хроматина, а именно действием специальных архитектурных белков, способных как формировать специфические, сильные, взаимодействия между удалёнными локусами, так и ограничивать пространственные контакты между ними.

В то же время были получены и неожиданные результаты. Из представления о том, что в регуляции экспрессии генов участвует архитектура хроматина следует, что консервативна должна быть именно измеренная в ходе эксперимента частота контактов между локусами. Однако при сравнении разных видов было обнаружено, что частота контактов изменяется согласно тому, как изменилось расстояние между целевыми локусами и сравниваемых видов. При этом консервативным оказывается относительная частота контакта, показывающая степень обеднённости или обогащённости по сравнению со средним для расстояния разделяющего контактирующие локусы.

Аналогичные результаты были получены при исследовании архитектуры хроматина комаров рода *Anopheles*. Примечательно, что даже между самыми далёкими видами, эволюционное расстояние между которыми сравнимо с расстоянием между *Homo sapiens* и *Mus musculus*, коэффициент корреляции достигал 0.95, при этом при использовании абсолютной модели корреляция была лишь немногим ниже от 0.8 до 0.9, что, по всей видимости связано с меньшим изменением расстояний между целевыми локусами внутри синтенных блоков.

Таким образом, использование алгоритма сравнения пространственной организации хроматина на уровне отдельных контактов реализованных в рамках ПО C-InterSecture свою эффективность в разных таксономических группах и высокую консервативность архитектуры хроматина. Показана консервативность относительных частота контактов, что поднимает вопрос о том, в какой степени измеренные в ходе эксперименты частоты связаны

с функцией контактов. Эта проблема требует дальнейшего изучения.

6. Эффект от использования кластера в достижении целей работы

Обработка данных по пространственной организации хроматина требует больших объёмов физической и оперативной памяти, что делает её невозможной на обычных рабочих станциях. Использование ресурсов кластера является определяющим для достижения целей работы.

7. Перечень публикаций, содержащих результаты работы

1) Veniamin Fishman, Nariman Battulin, **Miroslav Nuriddinov**, Antonina Maslova, Anna Zlotina, Anton Strunov, Darya Chervyakova, Alexey Korablev, Oleg Serov, Alla Krasikova. 3D organization of chicken genome demonstrates evolutionary conservation of topologically associated domains and highlights unique architecture of erythrocytes' chromatin// *Nucleic Acids Research*. – 2019. – V. – 47. – I. 2. – P. 648–665. <https://doi.org/10.1093/nar/gky1103>

2) **Miroslav Nuriddinov**, Veniamin Fishman. C-InterSecture—a computational tool for interspecies comparison of genome architecture// *Bioinformatics*. – 2019. – V. 35. – I. 23. – P. 4912–4921. <https://doi.org/10.1093/bioinformatics/btz415>

3) Varvara Lukyanchikova, Miroslav Nuriddinov, Polina Belokopytova, Jiangtao Liang, Maarten J.M.F. Reijnders, Livio Ruzzante, Robert M. Waterhouse, Zhijian Tu, Igor V. Sharakhov, Veniamin Fishman. *Anopheles* mosquitoes revealed new principles of 3D genome organization in insects // *bioRxiv* – 2020. – 2020.05.26.114017. <https://doi.org/10.1101/2020.05.26.114017>